Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 16 (2014), Issue 9 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Short Paper

Reviews

Viewpoints

Original Paper

A Web-Based Sexual Violence Bystander Intervention for Male College Students: Randomized Controlled Trial

Laura F Salazar[1], PhD; Alana Vivolo-Kantor[2], MPH; James Hardin[3], PhD; Alan Berkowitz[4], PhD

[1]Georgia State University, School of Public Health, Atlanta, GA, United States

[2]Centers for Disease Control and Prevention, Division of Violence Prevention, Atlanta, GA, United States

[3]University of South Carolina, Department of Epidemiology and Biostatistics, Columbia, SC, United States

[4]Independent Consultant, Mount Shasta, CA, United States

Corresponding Author:
Laura F Salazar, PhD
Georgia State University
School of Public Health
PO Box 3984
Atlanta, GA, 30302-3984
United States
Phone: 1 404 413 1396
Fax: 1 404 413 2544
Email: lsalazar1@gsu.edu

Abstract

Background: Bystander intervention approaches offer promise for reducing rates of sexual violence on college campuses. Most interventions are in-person small-group formats, which limit their reach and reduce their overall public health impact.

Objective: This study evaluated the efficacy of RealConsent, a Web-based bystander approach to sexual violence prevention, in enhancing prosocial intervening behaviors and preventing sexual violence perpetration.

Methods: A random probability sample of 743 male undergraduate students (aged 18 to 24 years) attending a large, urban university located in the southeastern United States was recruited online and randomized to either RealConsent (n=376) or a Web-based general health promotion program (n=367). Participants were surveyed online at baseline, postintervention, and 6-months postintervention. RealConsent was delivered via a password-protected Web portal that contained six 30-minute media-based and interactive modules covering knowledge of informed consent, communication skills regarding sex, the role of alcohol and male socialization in sexual violence, empathy for rape victims, and bystander education. Primary outcomes were self-reported prosocial intervening behaviors and sexual violence perpetration. Secondary outcomes were theoretical mediators (eg, knowledge, attitudes).

Results: At 6-month follow-up RealConsent participants intervened more often (P=.04) and engaged in less sexual violence perpetration (P=.04) compared to controls. In addition, RealConsent participants reported greater legal knowledge of sexual assault (P<.001), greater knowledge of effective consent (P<.001), less rape myths (P<.001), greater empathy for rape victims (P<.001), less negative date rape attitudes (P<.001), less hostility toward women (P=.01), greater intentions to intervene (P=.04), less hyper-gender ideology (P<.001), less positive outcome expectancies for nonconsensual sex (P=.03), more positive outcome expectancies for intervening (P<.001), and less comfort with other men's inappropriate behaviors (P<.001).

Conclusions: Our results support the efficacy of RealConsent. Due to its Web-based format, RealConsent has potential for broad-based dissemination thereby increasing its overall public health impact on sexual violence.

Trial Registration: Clinicaltrials.gov: NCT01903876; http://clinicaltrials.gov/show/NCT01903876 (Archived by WebCite at http://www.webcitation.org/6S1PXxWKt).

(J Med Internet Res 2014;16(9):e203) doi:10.2196/jmir.3426

KEYWORDS

Internet; sex offenses; rape; universities; students; public health

Introduction

Sexual violence is a major public health problem with long-term mental, physical, and social effects for victims [1-3]. Sexual violence encompasses a breadth of personal violations ranging from minor, nonconsensual noncontact acts, such as exhibitionism or verbal sexual harassment to sexual coercion, up to severe acts, such as attempted or completed nonconsensual oral, genital, or anal penetration [4,5]. The 2010 National Intimate Partner and Sexual Violence Survey found that 18.3% of women in the United States have been raped in their lifetime [1]. Younger age has been shown to be a significant factor in sexual violence risk. Recent research has shown that 80% of female victims of sexual violence reported that their first rape occurred before the age of 25 years [1]. According to results from the National College Women Sexual Victimization study, 2.8% of female college students surveyed reported either a completed rape or attempted rape in the previous 6 months, which equates with an incidence rate of 5.6% per year [6]. Research has also shown that the majority of sexual victimizations occur by perpetrators the victim knows [6-8].

To combat the problem of sexual violence, most prevention and intervention programs have focused on college populations and have shifted efforts recently to target elements in the environment rather than solely targeting individual characteristics of perpetrators or victims. The bystander model is one such approach. It is a theoretical model of community-level change that targets community members to intervene actively in situations that may be harmful and engages them to be accountable and to take action [9-11]. In fact, bystander intervention programs applied to dating and/or sexual violence interventions have proliferated in the past 5 years [12-20]. Most of these have involved a small-group format (eg, workshops), whereas a few involved a localized social marketing campaign. Subsequent evaluations demonstrated that the bystander model, in some cases, is effective in promoting active bystander behaviors and in changing social norms [13].

As promising as these interventions are, because of their small-group in-person format, they are resource-intensive and limited in their reach and sustainability. Alternatively, the use of the Internet as an effective medium to deliver health-related interventions has emerged [21,22] and offers significant advantages over in-person interventions, such as lower cost of intervention delivery, greater reach, maintenance of fidelity, the possibility of delivery in a wide range of settings, and ability to tailor content to a variety of users [23-26]. There are Web-based treatment programs designed to ameliorate depression [27,28], obesity [29,30], eating disorders [31], alcohol abuse [32-36], smoking [37], sexual risk reduction [38], and posttraumatic stress disorder [39]. In a study of college students, results showed that a Web-based format was preferable over a practitioner-delivered intervention for a hazardous drinking prevention program [33].

To date, there have been no Web-based sexual violence prevention programs that target male college students and incorporate the bystander approach, which have also been tested using a true experimental design and with sexual violence as an outcome. In response, RealConsent, a Web-based sexual violence prevention program incorporating a bystander approach, was developed and tested for its efficacy. This paper reports primary and secondary outcomes from a randomized controlled trial (RCT) of RealConsent. It was hypothesized that participants randomized to receive the RealConsent program would report greater increases in self-reported prosocial intervening behaviors and fewer incidents of sexual violence perpetration in comparison to participants randomized to receive an attention-placebo comparison program.

Methods

Recruitment and Study Design

An RCT was implemented at a Georgia State University, a large, urban university located in Atlanta, GA. Study procedures were approved by the participating sites' Institutional Review Boards. Participants provided informed consent; however, documentation was waived because of online recruitment.

Eligible participants were male undergraduates aged 18 to 24 years, single, who self-reported being either heterosexual or bisexual; exclusion criteria were graduate student status and homosexual sexual orientation. Active recruitment began February 2010 and ended in April 2010, and was accomplished through email messages from the principal investigator's university email address sent to randomly selected students. The sampling frame to generate the random sample was a list of student names obtained from the university's Office of Legal Affairs. The list included only first and last names, email addresses, and year of birth for all undergraduate students enrolled during 2009-2010 school year (N=21,280). To pare down the sampling frame, the list was sorted by birth year and all students born on or before January 1, 1984 were deleted from the list. Next, the list was sorted alphabetically by first name and an online resource (Baby Name Guesser) was used to determine likelihood of a name belonging to a male. This process resulted in a final sampling frame of 8458 male students. A random sample function from SPSS version 18.0 was used to select groups for invitation to participate in the RCT. Over the course of 10 weeks, 5 groups of randomly selected email addresses ranging in size from 300-3000 were emailed an invitation to participate.

Email invitations included a link to a website that included a short description of the study that blinded participants to the research questions and a short screener to determine eligibility. Eligible students were directed to a webpage that contained the informed consent form. Students were told that the purpose of the study was to "test multimedia, Web-based interactive programs designed for male college students." They were also told that the study would be anonymous (ie, email addresses could not be linked to their user ID or to their responses on the survey). Participants who provided consent were then directed to the RealConsent Web portal to register and obtain a username and set their password. After registering, simple randomization was implemented using a customized algorithm that assigned participants to either the experimental condition (RealConsent) or to an attention-placebo comparison condition. Participants were then directed to complete the baseline online survey.

Following completion of the survey, they were directed to either the RealConsent program or to the comparison program where they could begin the modules immediately or at a later time. Participants were asked to complete a survey postintervention and at 6-month follow-up. Weekly email reminders were sent to participants reminding them to complete each survey. The Web portal also allowed participants to send emails to a project coordinator and post messages if they needed assistance or had questions. Participants were paid US $25 via PayPal for completing each online survey.

Interventions

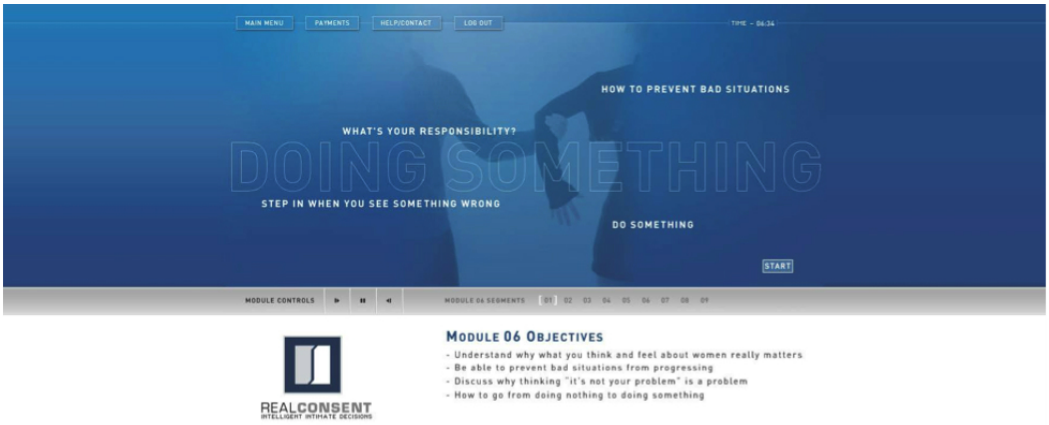
Experimental Intervention

The content for RealConsent was based on several complementary theoretical frameworks (social cognitive theory [40], social norms theory [41], and the bystander educational model [42]) as well as extensive formative research with the targeted population. RealConsent had two primary goals: (1) to increase prosocial intervening behaviors that reduce risk for sexual violence perpetration (eg, expressing disapproval when a peer is verbally disparaging toward women, attempting to stop a peer who tries to be coercive/violent) and (2) to prevent sexually violent behaviors toward women. These goals were to be achieved by affecting theoretically and empirically derived mediators, such as increasing knowledge of and skills for safely intervening, correcting misperceptions in normative beliefs, affecting negative attitudes toward date rape, increasing knowledge of the elements constituting effective consent for sex, affecting masculine gender roles, enhancing communication skills, and increasing empathy for victims of sexual assault.

RealConsent was delivered via a password-protected Web portal (see Figure 1 for screenshot) and consisted of six 30-minute modules with each module ranging in number of segments (1-14) and types of activities with diverse actors and appropriate language. Each of the modules involved interactivity, didactic activities, and episodes of a serial drama, which allowed us to model positive behaviors and illustrate both positive and negative outcome expectations for intervening and for perpetrating violence against women. Behaviors modeled in the serial drama included intervening, communicating with female sex partners, and obtaining effective consent for sex.

RealConsent was also programmed so that participants could not skip or click-through segments within each module. Participants were allowed to complete the modules at their own pace, but were encouraged via email to complete all modules within 3 weeks. Following the completion of each module, participants were immediately assessed and compensated US $10 for providing their acceptability and feedback on the module. These procedures ensured that each participant at follow-up had completed each module in its entirety.

Figure 1. Screenshot of RealConsent's bystander intervention module.



Attention-Placebo Comparison Intervention

The comparison condition involved a Web-based, general health promotion program titled Health Connection developed by the ISA Group [43]. The Health Connection program is an online, multimedia, health promotion program with 4 primary modules: stress management, fitness, weight management/nutrition, and substance abuse. Each program module is approximately 45 minutes long and audio narrated and approximated RealConsent in intensity, format, and time duration.

Measures

Overview

Primary outcome measures included prosocial intervening behaviors and sexual violence perpetration; secondary outcome measures were a host of theoretical mediating variables linked to the intervention activities. The primary behavioral outcomes, legal knowledge of sexual assault/rape and knowledge of effective consent for sex, are considered indexes and comprise items deemed "causal indicators," which indicates items are not necessarily correlated; thus, internal reliability was not calculated for these indexes because it is not an appropriate method to assess reliability [44,45]. The remaining theoretical mediators, considered "scales" and comprised of items deemed "effect indicators," should theoretically be highly intercorrelated. Cronbach alpha was calculated for the scale measures [44].

Primary Outcome Measures

Prosocial intervening behaviors were assessed with the Reactions to Offensive Language and Behavior (ROLB) index that measures whether or not men confronted inappropriate behaviors of other men [46,47]. We used the 7-item self-behavior subscale plus an additional 8 items [9], which directly reflected the content of RealConsent. A series of 15

potential intervening situations were presented and participants were asked to indicate whether they had experienced this situation (yes/no), whether they had "ever" intervened (at baseline), or whether they had intervened "within the past 6 months" (at 6-month follow-up). For each participant, a percent score was calculated that represented the total number of situations in which they intervened out of the total number of situations encountered multiplied by 100. For participants who indicated they had not encountered any of the 15 situations, their data were dropped from analyses.

Sexual violence was assessed with the sexual coercion subscale from the Revised Conflict Tactics Scale (CTS2) [48]. The CTS2 sexual coercion subscale is a 7-item questionnaire that assesses a range of sexually coercive behaviors in which participants are asked to indicate whether they had engaged in a sexually abusive tactic (eg, "I used force like hitting, holding down, or using a weapon to make a woman have sex") within a certain timeframe. We used "ever" at baseline and "within the past 6 months" at the 6-month follow-up.

Secondary Outcome Measures

The secondary outcome measures included legal knowledge of assault/rape [49], knowledge of effective consent for sex, self-efficacy to intervene [9,46], intentions to intervene, outcome expectancies for intervening behaviors [9], normative beliefs regarding sexual violence toward women [46,47], rape myths [50], gender-role ideology [51], empathy for rape victims [52], hostility toward women [53], attitudes toward date rape [54], and outcome expectancies for engaging in nonconsensual sex. Table 1 provides additional information about these secondary measures, including the mean and standard deviation, Cronbach alpha, the number of items, response options, and a sample item.

Table 1. Description of secondary outcomes/psychosocial mediators and means (SD) at baseline.

Psychosocial mediators	Mean (SD)	Range	Cronbach alpha	Items, n	Response options	Sample item
Legal knowledge assault/rape[a,b]	4.57 (1.75)	0-9	—	9	True/false	"In the state of Georgia, it is always legal to engage in sexual intercourse with a person under the age of 16 so long as he or she gives consent?"
Knowledge effective consent for sex[a,b]	11.58 (2.46)	0-14	—	14	True/false	"If a woman doesn't physically resist sex, she has given consent."
Self-efficacy to intervene	88.32 (22.27)	18-126	.95	18	1 (not at all confident) to 7 (extremely confident)	"Indicate my displeasure when I hear a sexist comment."
Intentions to intervene	52.39 (12.31)	15-75	.94	15	1 (not at all likely) to 5 (extremely likely)	"If I saw a man being verbally abusive toward a woman, I would intervene."
Outcome expectancies intervening	28.49 (4.60)	8-40	.80	8	1 (strongly disagree) to 5 (strongly agree)	"If I intervene, I can prevent someone from being hurt."
Self-comfort with men's inappropriate behaviors (normative beliefs)	33.11 (10.77)	8-56	.89	8	1 (not at all comfortable) to 7 (extremely comfortable)	Estimate how comfortable you feel with each of the following situations..."You are getting ready to go on a date when your roommate walks in with a bottle of tequila. He says to you, 'if you give her a couple shots of this, she'll loosen up.'"
Rape myth acceptance	36.28 (10.36)	17-85	.86	17	1 (not at all agree) to 5 (very much agree)	"Rape happens when a man's sex drive gets out of control."
Outcome expectancies engaging in rape	58.38 (10.96)	14-70	.87	14	1 (strongly disagree) to 5 (strongly agree)	If I engage in sex without clear consent..."I would feel more like a man."
Empathy for rape victims	68.87 (9.72)	19-95	.78	19	1 (strongly disagree) to 5 (strongly agree)	"In general, I feel that rape is an act that is not provoked by the rape victim."
Hostility toward women	3.66 (2.49)	0-10	.73	10	True/false	"I feel that many times women flirt with men just to tease them or hurt them."
Date rape attitudes	128.65 (24.95)	50-250	.93	50	1 (strongly disagree) to 5 (strongly agree)	"Most women don't understand that sexual jokes and innuendos are only for fun and are harmless."
Hyper-gender ideology	46.49 (11.28)	19-95	.89	19	1 (strongly disagree) to 5 (strongly agree)	"If men pay for a date, they deserve something in return."

[a] Mean represents the mean percent correct.

[b] Cronbach alpha not calculated for this index measure.

Data Analysis

Sample size calculations for the primary behavioral outcomes were estimated to guarantee that power would be at least .75 for the detection of small effect sizes [55]. Under a 2-group design and assuming 10% attrition over the 6-month follow-up period required enrolling at least 340 participants in each study condition to detect the specified effect size with power of .83.

Analyses were performed only on prespecified hypotheses using an intention-to-treat protocol in which participants were analyzed according to their original assigned study conditions [56,57]. At baseline, descriptive statistics were calculated to summarize sociodemographic variables, theoretical mediators, intervening, and sexual perpetration behaviors between study conditions. Differences between study conditions were assessed using *t* tests for continuous variables and chi-square analyses for categorical variables [58]. Variables in which differences between study conditions approached statistical significance ($P<.10$) or that were theoretically or empirically identified as potential confounders were included as covariates in the models. The effectiveness of RealConsent on primary outcomes was analyzed for the 6-month period examining changes in behavioral outcomes across 2 time points (baseline and 6 months); the postintervention time point was not examined because we did not expect behavioral changes immediately following completion of RealConsent. The effects analysis for primary outcomes used logistic regression to compute adjusted odd ratios (AORs) for dichotomous outcomes [59] and analyses of covariance (ANCOVA) [60] to compute adjusted means and mean differences for continuous outcomes. Each of these approaches included the corresponding baseline measure for

the specific outcome as a covariate in the analysis. Effect sizes with Cohen's *d* were also calculated using Practical Meta-Analysis Effect Size Calculator [61].

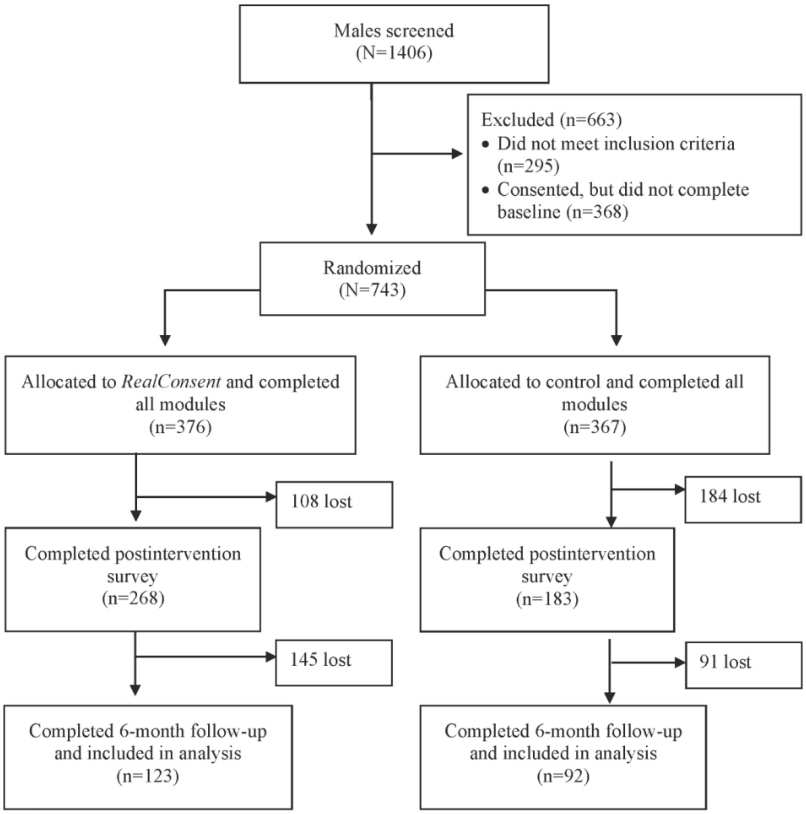
To assess effects of RealConsent on secondary outcomes, linear generalized estimating equation (GEE) regression models were estimated to control for repeated within-subject measurements and examined changes in outcomes across 3 time points (baseline, postintervention, and 6-months); these models assuming exchangeable correlation are equivalent to random-effects models which include a second variance component for participants. These models admitted a differential number of observations on study participants over the longitudinal course of observation and included a time-independent variable (treatment condition) as well as time-dependent variables (covariates and outcomes). The models adjusted for the corresponding baseline measure for each outcome and other covariates to obtain adjusted mean differences used to assess the effect of the intervention on each continuous outcome. Additionally, an indicator for the time period was included in the model to capture any unaccounted temporal effects [62,63]. The 95% confidence intervals (CIs) around the adjusted mean differences and the corresponding *P* values were also computed. To obtain adjusted means and mean differences, models were repeatedly estimated from the bootstrap samples in which samples were drawn with replacement at the level of the participant. For each model, adjusted means were calculated and standard errors were then calculated from the collection of bootstrap results [64]. Percentage of relative change for continuous variables was computed as the difference between the adjusted means for each condition divided by the adjusted mean for the comparison condition. Analyses were performed using Stata statistical software, version 13 (StataCorp LP, College Station, TX, USA), and SAS version 9.3 (SAS Institute Inc, Cary, NC, USA).

Results

Recruitment

The recruitment process resulted in 1406 male college students who were screened. Of those, 295 (20.98%) were not eligible and 1111 (79.02%) accepted and consented. There was some attrition (33.12%, 368/1111) between initially agreeing and subsequently completing the baseline survey. The final sample resulted in 743 eligible students who were randomized and completed baseline. At postintervention, 451 of 743 (60.7%) completed the follow-up survey. Attrition at this time point was loss to follow-up. At 6-month follow-up, 215 of 743 (28.9%) completed the survey and were included in analyses (Figure 2). Attrition at this final time point was because of loss to follow-up, but also because the trial ended prematurely. More information is provided in the Discussion section.

Figure 2. CONSORT diagram.



Participant Characteristics

Overview

The average age was 20.38 years (SD 1.67). Racial breakdown of participants was white (44.1%, 328/743), African American (22.3%, 166/743), Asian American (19.6%, 146/743), and Hispanic (10.8%, 80/743). Most participants were full-time students (92.1%, 684/743) and single (75.2%, 559/743); a small number were members of fraternities (12.1%, 90/743) or athletic teams (8.5%, 63/743). Overall, the prevalence of sexual violence perpetration at baseline was 32.2% (234/727). The mean percent prosocial intervening behaviors at baseline was 72% (SD 25%), meaning male participants self-reported that they intervened on average 72% of the time when exposed to a situation.

Table 2 provides data on the breakdown by sociodemographic variables, proposed mediators, and outcome behaviors for both conditions. Participants in the comparison condition reported higher levels of hostility toward women (mean 3.89, SD 2.54 vs mean 3.42, SD 2.43; *P*=.01), higher average sexual coercion tactics used (mean 0.76, SD 1.29 vs mean 0.53, SD 1.07; *P*=.01), and were more likely to ever report sexual coercion perpetration (37.0%, 134/362 vs 27.4%, 100/365; *P*=.006) than RealConsent participants. These variables at baseline were controlled for in subsequent analyses.

Table 2. Baseline equivalence of RealConsent and attention-placebo comparison condition participants (N=743).

Characteristics	RealConsent intervention (n=376)	Attention-placebo comparison (n=367)
Sociodemographics		
Age (years), mean (SD)	20.42 (1.69)	20.33 (1.66)
International student,[a] n (%)	9 (2.4)	18 (4.9)
Member of a fraternity,[a] n (%)	37 (9.9)	53 (14.6)
Member of an athletic team at college, n (%)	29 (7.8)	34 (9.4)
Race, n (%)		
White	170 (45.2)	158 (43.1)
African American or black	83 (22.1)	83 (22.6)
Asian or Pacific Islander	73 (19.4)	73 (19.9)
Hispanic or Latino	38 (10.1)	42 (11.4)
American Indian, Alaskan native, or native Hawaiian	12 (3.2)	11 (3.0)
Frequency of vaginal intercourse, mean (SD)[a]	7.55 (12.75)	8.38 (13.03)
Grade point average, [a] n (%)		
<2.0	8 (2.1)	14 (3.8)
2.0-2.9	112 (29.8)	124 (33.8)
3.0-3.4	163 (43.4)	150 (40.9)
3.5-4.0	93 (24.7)	79 (21.5)
Residence status, [a] n (%)		
Campus or residence hall	80 (21.5)	80 (22.0)
Fraternity house	2 (0.5)	6 (1.6)
Other university housing	0 (0.0)	4 (1.1)
Off-campus housing such as own apartment	141 (37.9)	123 (33.8)
Parent’s or guardian’s home	144 (38.7)	149 (40.9)
Other	5 (1.3)	2 (0.5)
Relationship status, [a] n (%)		
Single	289 (77.7)	270 (74.2)
Married or has domestic partner	1 (0.3)	6 (1.6)
Engaged or in committed relationship	82 (22.0)	87 (23.9)
Separated	0 (0.0)	0 (0.0)
Divorced	0 (0.0)	1 (0.3)
Primary behavioral outcomes		
Sexual coercion perpetration, mean (SD)[a]	0.53 (1.07)	0.76 (1.29)
Sexual coercion perpetration, dichotomized[a] n(%)	100 (27.4)	134 (37.0)
Prosocial intervening, mean percent (SD)	72 (25)	72 (24)
Secondary outcomes (mediators), mean (SD)		
Legal knowledge of assault/rape	4.54 (1.73)	4.61 (1.78)
Knowledge of effective consent for sex	11.69 (2.44)	11.46 (2.49)
Self-efficacy to intervene	88.74 (20.93)	87.88 (23.60)
Intentions to intervene	52.68 (11.85)	52.09 (12.78)
Outcome expectancies for intervening	28.72 (4.33)	28.26 (4.86)

Characteristics	RealConsent intervention (n=376)	Attention-placebo comparison (n=367)
Self-comfort with men's inappropriate behaviors (normative beliefs)	32.68 (10.82)	33.55 (10.71)
Rape myth acceptance	35.88 (10.43)	36.69 (10.29)
Outcome expectancies for engaging in rape	58.71 (10.76)	58.02 (11.17)
Empathy for rape victims	68.96 (9.66)	68.78 (9.79)
Hostility toward women[a]	3.42 (2.43)	3.89 (2.54)
Date rape attitudes	128.85 (24.07)	128.45 (25.84)
Hyper-gender ideology	46.59 (11.28)	46.38 (11.30)

[a] Denotes differences in baseline responses between conditions (P<.15).

Differences Between Completers and Noncompleters

An analysis of sample attrition was conducted to better understand differences between participants who completed the 6-month follow-up and those who did not. A total of 528 participants were lost to follow-up (71.1%), 275 of 367 (74.9%) in the comparison condition versus 253 of 376 (67.3%) in the RealConsent condition (P=.02). Comparisons of sociodemographic variables and primary outcomes on baseline responses indicated that completers were more likely to have higher grade point averages (GPAs) than noncompleters (P=.01). Completers and noncompleters did not differ on primary outcomes. Inference of GEE model results carries with it an assumption that missing data are missing completely at random (MCAR; more conservative assumption that "missingness" is independent of observed and missing outcomes), whereas our data were missing at random (MAR; less conservative assumption that missingness is only independent of observed outcomes). To investigate whether inference changed as a result of attrition (MAR vs MCAR), missing outcomes were imputed and no change in inference was found; thus, results using original data are presented.

Primary Outcomes

The effectiveness of RealConsent on prosocial intervening behaviors and sexual coercion was estimated with ANCOVA; covariates included baseline scores for the behaviors and those correlated sociodemographic variables. Unadjusted means were graphed across all 3-survey time points and are presented in Figures 3 and 4. RealConsent participants (adjusted mean 0.81, SE 0.03) reported significantly more prosocial intervening behaviors at 6-month follow-up than comparison participants (adjusted mean 0.72, SE 0.03; $F_{1,123}$=4.128, P=.04). Additionally, RealConsent participants reported significantly less sexual violence at 6-month follow-up (adjusted mean 0.26, SE 0.08) than comparison participants (adjusted mean 0.50, SE 0.09; $F_{1,193}$=4.18, P=.04). Cohen's d effect sizes for prosocial intervening behaviors and sexual coercion were 0.37 and 0.29, respectively.

Using logistic regression, we assessed the primary prevention effect (ie, comparing those who had not perpetrated to those who had) of RealConsent on prevalence of perpetrating sexual violence. The odds for perpetrating among RealConsent participants were 73% lower than participants in the comparison condition (AOR 0.27, 95% CI 0.11-0.70, P=.007).

Figure 3. Unadjusted means for sexual violence perpetration across 3 time points for RealConsent and attention-placebo comparison conditions.

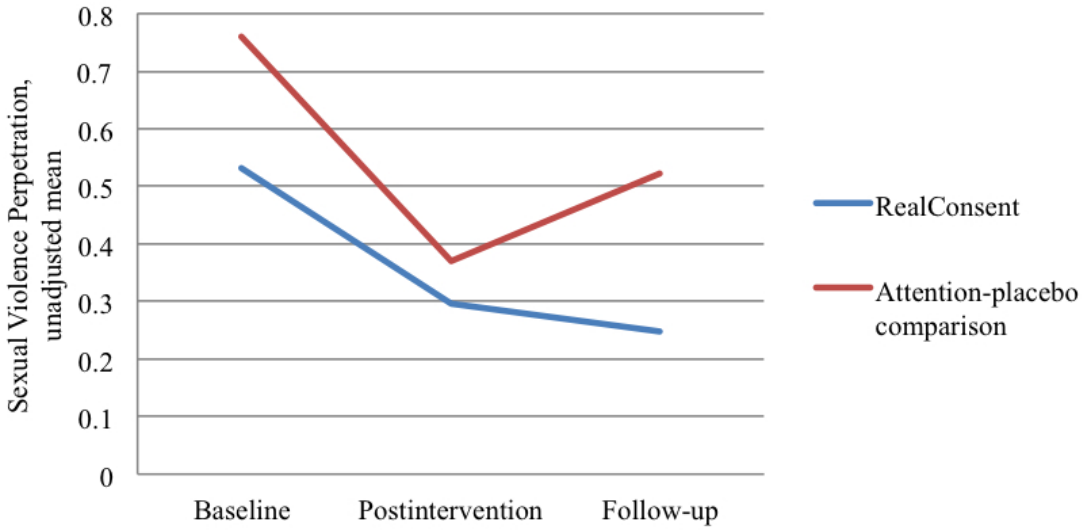
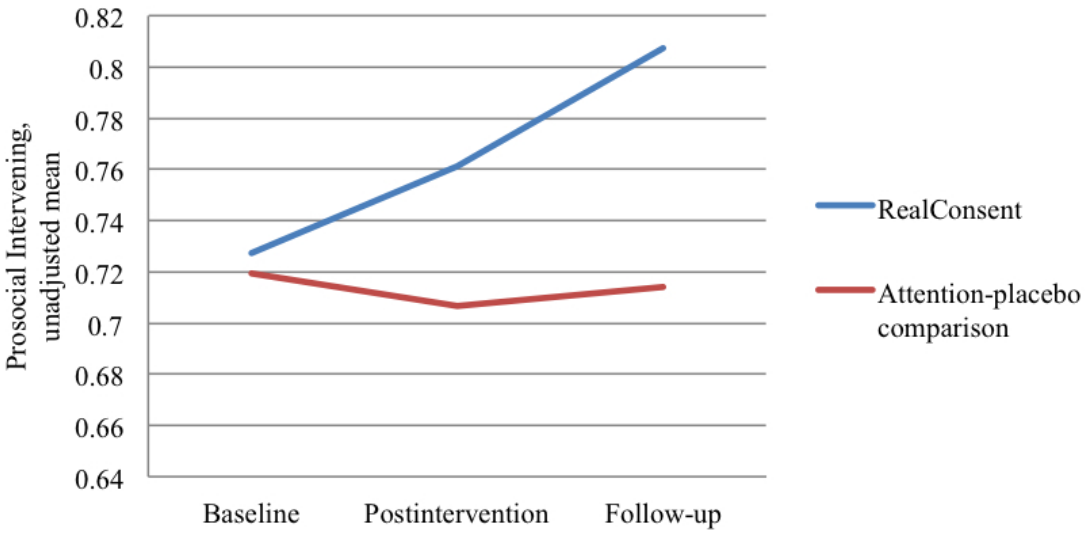


Figure 4. Unadjusted means for prosocial intervening behavior across 3 time points for RealConsent and attention-placebo comparison conditions.



Secondary Outcomes/Mediators

Separate GEE analyses were conducted to examine the effect of RealConsent on the 12 proposed mediators. Similar to the previous analyses, we controlled for baseline scores for each mediator and sociodemographic variables and compared the adjusted means of each group over the entire 6-month follow-up period. Of the 12 mediators, only self-efficacy to intervene was not significant (see Table 3).

Table 3. Generalized estimating equation (GEE) regression models of intervention effects on psychosocial mediators.

Psychosocial mediators	Estimate (SE)	*P*	RealConsent, adjusted mean (95% CI)[a]	Comparison, adjusted mean (95% CI)[a]
Legal knowledge of assault/rape	1.99 (0.18)	<.001	6.64 (6.41, 6.86)	4.65 (4.38, 4.92)
Knowledge effective consent for sex	0.80 (0.19)	<.001	12.83 (12.59, 13.07)	12.02 (11.74, 12.31)
Self-efficacy to intervene	1.45 (1.57)	.36	91.01 (89.07, 92.95)	89.56 (87.21, 91.92)
Intentions to intervene	1.87 (0.87)	.04	54.72 (53.61, 55.82)	52.84 (51.52, 54.17)
Outcome expectancies for intervening	1.25 (0.38)	.001	29.08 (28.61, 29.55)	27.83 (27.26, 28.40)
Self-comfort with men's inappropriate behaviors (normative beliefs)	–2.85 (0.83)	.001	30.87 (29.83,31.91)	33.72 (32.49, 34.95)
Rape myth acceptance	–5.48 (0.77)	<.001	31.14 (30.18, 32.10)	36.62 (35.48, 37.76)
Outcome expectancies for rape	1.25 (0.58)	.03	55.12 (54.42, 55.82)	53.86 (52.98, 54.74)
Empathy for rape victims	3.51 (0.73)	<.001	72.04 (71.14, 72.93)	68.52 (67.43, 69.62)
Hostility toward women	–.46 (0.18)	.01	2.74 (2.51, 2.97)	3.20 (2.93, 3.47)
Date rape attitudes	–13.67 (2.08)	<.001	112.36 (109.75, 114.97)	126.03 (122.94, 129.12)
Hyper-gender ideology	–3.67 (0.78)	<.001	42.49 (41.53, 43.46)	46.17 (44.99, 47.34)

[a] All GEE models included the covariates international student status, fraternity membership, frequency of vaginal intercourse, GPA, residence status, and relationship status.

Discussion

Principal Findings

This study was the first Web-based, sexual violence prevention program that incorporated a bystander approach and demonstrated significant changes in behavior for an ethnically diverse sample of male college students. Over the course of the 6-month follow-up period, RealConsent participants were significantly less likely to engage in sexual violence perpetration and significantly more likely to engage in prosocial intervening behavior when they encountered a situation in which they could intervene. It was also observed that these primary behavioral outcomes might have been achieved through hypothesized effects on a host of the program's theoretical mediators. We found significant changes in all but 1 of the mediating variables and all in the hypothesized direction. This success is noteworthy given that many previous evaluations of sexual violence prevention programs for male college students have measured mostly behavioral intentions rather than actual behavior as outcomes [65,66]; even when behaviors were measured, rarely were significant effects observed [13]. Further, the online

administration of the RealConsent program, during which no face-to-face interaction with study participants occurred, suggests that RealConsent's mode of delivery is equally effective as other approaches involving face-to-face interaction. Thus, given a Web-based modality, RealConsent provides significant advantages by facilitating dissemination thereby increasing reach [22].

Several mechanisms could explain how RealConsent not only increased prosocial intervening behaviors, but also showed effects on sexual violence perpetration. Research has shown theory to be a significant factor in contributing to behavior change among Web-based interventions [67]. RealConsent incorporated social cognitive theory, social norms theory, and the bystander educational model as an overarching framework for developing activities and interactive segments that putatively supported behavior change and identified relevant constructs to be targeted, such as knowledge and self-efficacy, and suggested the correcting of inaccurate perceived norms. Current research documents that misperceptions in norms are a major barrier to bystander intervention and also that perpetrators overestimate other men's support for what they do [13]. Correcting misperceptions along with teaching bystander intervention combines 2 evidence-based approaches that together may have produced the observed behavioral outcomes. In addition, this framework provided specific behavior change techniques (eg, evoking vicarious learning of targeted behaviors). Because we found significant effects in the hypothesized direction in all but 1 of the theoretical mediators that represented these constructs, it is plausible that several of these theoretical mediating variables would explain the observed behavioral effects. Future research should determine the specific theoretical mediating mechanisms underlying the effects of RealConsent on the primary outcomes.

The efficacy of RealConsent may also be partly attributable to the extensive formative qualitative research conducted by the authors with the targeted population to inform the relevance of content and messages and the style of the Web portal. This formative research, in turn, provided the authors with the necessary resources to build each module of RealConsent. The research team also sought to gain insight from multiple disciplines in the development of the modules; thus, the team included experts from public health, the social sciences, Web design, and marketing. This integration of expertise contributed to the use of proven behavior change techniques that span multiple fields, such as educational entertainment, didactic methods, and problem-based learning with user-face interactivity [68,69] to evoke targeted behavioral outcomes.

Furthermore, because most previous prevention programs have been tested with mostly white participants [65,66], it is important to note that RealConsent showed significant effects among a racially diverse sample of male undergraduates recruited from a large, urban university in the southeast United States. In developing RealConsent, although given no budget constraints, it would have been preferable to reflect deep structure in terms of cultural sensitivity in the content and messages [70]; we were at least able to provide surface structure by using diverse actors and appropriate language in several segments, including the serial drama as well as still images. Appealing to a diverse population provides enhanced generalizability of the program and suggests that RealConsent may be effective with a diverse audience.

These results highlight an important contribution to the literature on evidence-based public health approaches to prevent sexual violence. A review by researchers at the Centers for Disease Control and Prevention (CDC) identified the lack of evidence-based sexual violence prevention approaches at all levels of the social ecology [71]. In particular, they found few effective community-level approaches that seek to alter the characteristics of settings by changing community-level norms. A main tenet of the bystander approach is to provide resources to individuals that may influence their willingness to intervene in situations either actively or passively and to change perceptions of inaccurate norms. Although RealConsent is mostly targeted at changing individual norms and behavior, the effects on intentions to intervene and prosocial intervening behaviors may aggregate to community-level change. Future research could potentially test the effects at the college level to determine whether RealConsent's effects would reach beyond the male participants enrolled.

Limitations and Strengths

This study is not without limitations. First, there was a significant amount of attrition mostly because of funding issues and some loss to follow-up. Developing the content, producing the content, and programming the Web portal and the online recruitment platform, in addition to implementing a RCT with a 6-month follow-up took much longer than the anticipated 3-year time frame. Unfortunately, the trial ended prematurely resulting in significant and unforeseen "forced" attrition at 6-month follow-up. Second, we experienced some loss to follow-up. It is unclear what the potential reasons were for this loss to follow-up because data for noncompletion were not collected; however, previous research has shown that attrition in Web-based trials may be higher compared to in-person trials [72,73] and high dropout rates may be considered "a natural and typical feature" [73]. Most important, however, is that our statistical analyses comparing completers and noncompleters on baseline responses indicated 1 minor difference (GPA), which had no influence on study outcomes, and we imputed missing outcomes and no change in inference was found suggesting attrition bias is not a significant threat to the results observed.

Significant differences at baseline on several primary outcomes suggest that randomization was not perfect; however, we have no reason to believe that these differences were not due to chance because participants were assigned by computer algorithm. Nonetheless, we controlled for baseline scores in all analyses and still found significant differences between groups. Lastly, the sample selected for this trial was specific to a large, urban university in the Southeast so results may not generalize to other populations. Future research should replicate these results with populations in both urban and rural universities.

Nevertheless, there were some inherent strengths. The theoretically derived intervention was developed using extensive formative research with the study population and pilot tested with much success and praise. In addition, the RCT coupled with random probability sampling and the use of an

attention-placebo comparison group represent significant methodological strengths.

Conclusions

Given that colleges and universities that receive federal aid are mandated via the Clery Act (20 USCA § 1092) to inform students about crime statistics as well as policies and procedures that are in place to prevent sexual offenses, evidence-based sexual violence prevention programs that are cost-effective, easily implemented, and appeal to a diverse population are much needed. Prior studies have presented numerous sexual violence prevention programs that are effective in changing negative attitudes, rape myths, and behaviors for some; however, none involved an empirically tested Web-based program [20,65,66]. Recent reviews of sexual violence prevention programs describe the importance of engaging men to be women's allies in preventing sexual violence as an important element, as in a bystander approach [66]; however, equally important is the ability of interventions to reach large populations rather than only the men who volunteer [66]. RealConsent is a scalable intervention that with its Web-based approach, behavioral outcomes, and additional conclusive research holds potential to reach large segments of male undergraduate populations while also engaging young men to intervene so that sexual violence will be prevented.

Acknowledgments

This study was funded by the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, Grant #R49 CE000892. The authors also wish to acknowledge the significant contributions and talents of Jody Kaufman, Nichole Daluga-Guenther, and Kira McGroarty Koon to intervention development, survey development, and preliminary testing of the program.

Conflicts of Interest

Laura Salazar was the developer of RealConsent, but she did not derive financial income from the Web-based program. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [74].

[PDF File (Adobe PDF File), 1004KB - jmir_v16i9e203_app1.pdf]

References

1. Black MC, Basile KC, Breiding MJ. The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary report. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2011. URL: http://www.cdc.gov/violenceprevention/pdf/nisvs_report2010-a.pdf [accessed 2014-08-28] [WebCite Cache ID 6SAPUDoWz]
2. Casey EA, Lindhorst TP. Toward a multi-level, ecological approach to the primary prevention of sexual assault: prevention in peer and community contexts. Trauma Violence Abuse 2009 Apr;10(2):91-114. [doi: 10.1177/1524838009334129] [Medline: 19383629]
3. Amar AF, Gennaro S. Dating violence in college women: associated physical injury, healthcare usage, and mental health symptoms. Nurs Res 2005;54(4):235-242. [Medline: 16027566]
4. Basile KC, Saltzman LE. Sexual violence surveillance: Uniform definitions and recommended data elements. Atlanta, GA: National Center for Injury Control and Violence Prevention; 2009. URL: http://www.cdc.gov/violenceprevention/pdf/sv_surveillance_definitionsl-2009-a.pdf [accessed 2014-08-28] [WebCite Cache ID 6SAcqfEHi]
5. Abbey A, McAuslan P. A longitudinal examination of male college students' perpetration of sexual assault. J Consult Clin Psychol 2004 Oct;72(5):747-756. [doi: 10.1037/0022-006X.72.5.747] [Medline: 15482033]
6. Fisher BS, Cullen FT, Turner MG. The Sexual Vicitimization of College Women. Washington, DC: United States Department of Justice, Office of Justice Programs; 2000 Dec. URL: https://www.ncjrs.gov/pdffiles1/nij/182369.pdf [accessed 2014-08-28] [WebCite Cache ID 6SAmozzsB]
7. Forke CM, Myers RK, Catallozzi M, Schwarz DF. Relationship violence among female and male college undergraduate students. Arch Pediatr Adolesc Med 2008 Jul;162(7):634-641. [doi: 10.1001/archpedi.162.7.634] [Medline: 18606934]
8. Banyard VL. Sexual violence: current perspectives on prevention and intervention. J Prev Interv Community 2008;36(1-2):1-4. [doi: 10.1080/10852350802022225] [Medline: 19042459]
9. Banyard VL, Plante EG, Moynihan MM. Bystander education: Bringing a broader community perspective to sexual violence prevention. J Community Psychol 2003 Jan;32(1):61-79. [doi: 10.1002/jcop.10078]
10. Ward KJ. Mentors in Violence Prevention: Evaluation 1999/2000. Boston: Northeastern University; 2001. URL: http://www.mvpnational.org/wp-content/uploads/2011/12/MVP-HS-Eval-Report-1999-2000.pdf [accessed 2014-08-28] [WebCite Cache ID 6SAnPKi8K]

11. Tabachnick J. Engaging bystanders in sexual violence prevention. Enola, PA: National Sexual Violence Resource Center; 2009. URL: http://www.nsvrc.org/sites/default/files/Publications_NSVRC_Booklets_Text-Only_Engaging-Bystanders-in-Sexual-Violence-Prevention.pdf [accessed 2014-08-28] [WebCite Cache ID 6SAnbWdwV]
12. Foubert JD, Langhinrichsen-Rohling J, Brasfield H, Hill B. Effects of a rape awareness program on college women: increasing bystander efficacy and willingness to intervene. J Community Psychol 2010 Aug 03;38(7):813-827. [doi: 10.1002/jcop.20397]
13. Gidycz CA, Orchowski LM, Berkowitz AD. Preventing sexual aggression among college men: an evaluation of a social norms and bystander intervention program. Violence Against Women 2011 Jun;17(6):720-742. [doi: 10.1177/1077801211409727] [Medline: 21571742]
14. Potter SJ. Using a multimedia social marketing campaign to increase active bystanders on the college campus. J Am Coll Health 2012;60(4):282-295. [doi: 10.1080/07448481.2011.599350] [Medline: 22559087]
15. Potter SJ, Stapleton JG. Bringing in the target audience in bystander social marketing materials for communities: suggestions for practitioners. Violence Against Women 2011 Jun;17(6):797-812. [doi: 10.1177/1077801211410364] [Medline: 21727157]
16. Ahrens CE, Rich MD, Ullman JB. Rehearsing for real life: the impact of the InterACT Sexual Assault Prevention Program on self-reported likelihood of engaging in bystander interventions. Violence Against Women 2011 Jun;17(6):760-776. [doi: 10.1177/1077801211410212] [Medline: 21628338]
17. Langhinrichsen-Rohling J, Foubert JD, Brasfield HM, Hill B, Shelley-Tremblay S. The Men's Program: does it impact college men's self-reported bystander efficacy and willingness to intervene? Violence Against Women 2011 Jun;17(6):743-759. [doi: 10.1177/1077801211409728] [Medline: 21571743]
18. Coker AL, Cook-Craig PG, Williams CM, Fisher BS, Clear ER, Garcia LS, et al. Evaluation of Green Dot: an active bystander intervention to reduce sexual violence on college campuses. Violence Against Women 2011 Jun;17(6):777-796. [doi: 10.1177/1077801211410264] [Medline: 21642269]
19. Moynihan MM, Banyard VL, Arnold JS, Eckstein RP, Stapleton JG. Sisterhood may be powerful for reducing sexual and intimate partner violence: an evaluation of the Bringing in the Bystander in-person program with sorority members. Violence Against Women 2011 Jun;17(6):703-719. [doi: 10.1177/1077801211409726] [Medline: 21628337]
20. Banyard VL, Moynihan MM, Crossman M. Reducing sexual violence on campus: the role of student leaders as empowered bystanders. Journal of College Student Development 2009;50(4):446-457. [doi: 10.1353/csd.0.0083]
21. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
22. Christensen H, Calear AL, Andersson G, Thorndike FP, Tait RJ. Beyond efficacy: the depth and diversity of current internet Interventions. J Med Internet Res 2012;14(3):e92 [FREE Full text] [doi: 10.2196/jmir.2206] [Medline: 22743601]
23. Cassell MM, Jackson C, Cheuvront B. Health communication on the Internet: an effective channel for health behavior change? J Health Commun 1998;3(1):71-79. [Medline: 10947375]
24. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13:e17. [Medline: 21320854]
25. Evers KE, Prochaska JM, Prochaska JO, Driskell MM, Cummins CO, Velicer WF. Strengths and weaknesses of health behavior change programs on the internet. J Health Psychol 2003 Jan;8(1):63-70. [doi: 10.1177/1359105303008001435] [Medline: 22113901]
26. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res 2010;12(5):e60 [FREE Full text] [doi: 10.2196/jmir.1463] [Medline: 21169162]
27. Christensen H, Leach LS, Barney L, Mackinnon A, Griffiths K. The effect of web based depression interventions on self reported help seeking: randomised controlled trial [ISRCTN77824516]. BMC Psychiatry 2006;6:13 [FREE Full text] [doi: 10.1186/1471-244X-6-13] [Medline: 16595018]
28. 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]
29. Zabinski MF, Pung MA, Wilfley DE, Eppstein DL, Winzelberg AJ, Celio A, et al. Reducing risk factors for eating disorders: targeting at-risk women with a computerized psychoeducational program. Int J Eat Disord 2001 May;29(4):401-408. [Medline: 11285577]
30. Maes L, Vereecken CA, Gedrich K, Rieken K, Sichert-Hellert W, De Bourdeaudhuij I, HELENA Study Group. A feasibility study of using a diet optimization approach in a web-based computer-tailoring intervention for adolescents. Int J Obes (Lond) 2008 Nov;32 Suppl 5:S76-S81. [doi: 10.1038/ijo.2008.186] [Medline: 19011657]
31. Celio AA, Winzelberg AJ, Wilfley DE, Eppstein-Herald D, Springer EA, Dev P, et al. Reducing risk factors for eating disorders: comparison of an Internet- and a classroom-delivered psychoeducational program. J Consult Clin Psychol 2000 Aug;68(4):650-657. [Medline: 10965640]
32. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]

33. Kypri K, Saunders JB, Williams SM, McGee RO, Langley JD, Cashell-Smith ML, et al. Web-based screening and brief intervention for hazardous drinking: a double-blind randomized controlled trial. Addiction 2004 Nov;99(11):1410-1417. [doi: 10.1111/j.1360-0443.2004.00847.x] [Medline: 15500594]
34. Bewick BM, Trusler K, Mulhern B, Barkham M, Hill AJ. The feasibility and effectiveness of a web-based personalised feedback and social norms alcohol intervention in UK university students: a randomised control trial. Addict Behav 2008 Sep;33(9):1192-1198. [doi: 10.1016/j.addbeh.2008.05.002] [Medline: 18554819]
35. Postel MG, de Haan HA, ter Huurne ED, Becker ES, de Jong CA. Effectiveness of a web-based intervention for problem drinkers and reasons for dropout: randomized controlled trial. J Med Internet Res 2010;12(4):e68 [FREE Full text] [doi: 10.2196/jmir.1642] [Medline: 21163776]
36. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
37. Lenert L, Muñoz RF, Stoddard J, Delucchi K, Bansod A, Skoczen S, et al. Design and pilot evaluation of an internet smoking cessation program. J Am Med Inform Assoc 2003;10(1):16-20 [FREE Full text] [Medline: 12509354]
38. Danielson CK, McCauley JL, Jones AM, Borkman AL, Miller S, Ruggiero KJ. Feasibility of delivering evidence-based HIV/STI prevention programming to a community sample of African American teen girls via the internet. AIDS Educ Prev 2013 Oct;25(5):394-404 [FREE Full text] [doi: 10.1521/aeap.2013.25.5.394] [Medline: 24059877]
39. Mouthaan J, Sijbrandij M, de Vries GJ, Reitsma JB, van de Schoot R, Goslings JC, et al. Internet-based early intervention to prevent posttraumatic stress disorder in injury patients: randomized controlled trial. J Med Internet Res 2013;15(8):e165 [FREE Full text] [doi: 10.2196/jmir.2460] [Medline: 23942480]
40. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
41. Fabiano PM, Perkins HW, Berkowitz A, Linkenbach J, Stark C. Engaging men as social justice allies in ending violence against women: evidence for a social norms approach. J Am Coll Health 2003;52(3):105-112. [doi: 10.1080/07448480309595732] [Medline: 14992295]
42. Banyard VL, Moynihan MM, Plante EG. Sexual violence prevention through bystander education: An experimental evaluation. J Community Psychol 2007 May;35(4):463-481. [doi: 10.1002/jcop.20159]
43. The ISA Group. URL: http://www.isagroup.com/ [accessed 2014-08-22] [WebCite Cache ID 6S1RBJQL7]
44. Streiner DL. Being inconsistent about consistency: when coefficient alpha does and doesn't matter. J Pers Assess 2003 Jun;80(3):217-222. [doi: 10.1207/S15327752JPA8003_01] [Medline: 12763696]
45. Hennessy M, Bleakley A, Fishbein M. Measurement Models for Reasoned Action Theory. Ann Am Acad Pol Soc Sci 2012 Mar;640(1):42-57 [FREE Full text] [doi: 10.1177/0002716211424709] [Medline: 23243315]
46. Kilmartin C, Conway A, Friedberg A. Social conformity and sexism in all-male peer groups. In: Virginia Psychological Association Spring Convention. 1999 Presented at: Virginia Psychological Association Spring Convention; Apr 1999; Virginia Beach, VA.
47. Loh C, Gidycz CA, Lobo TR, Luthra R. A prospective analysis of sexual assault perpetration: risk factors related to perpetrator characteristics. J Interpers Violence 2005 Oct;20(10):1325-1348. [doi: 10.1177/0886260505278528] [Medline: 16162492]
48. Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The Revised Conflict Tactics Scales (CTS2): development and preliminary psychometric data. Journal of Family Issues 1996 May 01;17(3):283-316. [doi: 10.1177/019251396017003001]
49. Maxwell CD, Robinson AL, Post LA. The nature and predictors of sexual victimization and offending among adolescents. Journal of Youth and Adolescence 2003;32:465-477. [doi: 10.1023/A:1025942503285]
50. Payne DL, Lonsway KA, Fitzgerald LF. Rape myth acceptance: exploration of its structure and its measurement using the Illinois Rape Myth Acceptance Scale. Journal of Research in Personality 1999 Mar;33(1):27-68. [doi: 10.1006/jrpe.1998.2238]
51. Hamburger ME, Hogben M, McGowan S, Dawson LJ. Assessing hypergender ideologies: development and initial validation of a gender-neutral measure of adherence to extreme gender-role beliefs. Journal of Research in Personality 1996 Jun;30(2):157-178. [doi: 10.1006/jrpe.1996.0011]
52. Deitz SR, Blackwell KT, Daley PC, Bentley BJ. Measurement of empathy toward rape victims and rapists. J Pers Soc Psychol 1982 Aug;43(2):372-384. [Medline: 7120042]
53. Check JVP. The Hostility Toward Women Scale [dissertation]. Winnipeg, MB: University of Manitoba; 1985.
54. Burgess GH. Assessment of rape-supportive attitudes and beliefs in college men: development, reliability, and validity of the rape attitudes and beliefs scale. J Interpers Violence 2007 Aug;22(8):973-993. [doi: 10.1177/0886260507302993] [Medline: 17709805]
55. Cohen J. Statistical Power for the Behavioral Sciences. 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
56. Piantadosi S. Clinical Trials: A Methodologic Perspective. New York: John Wiley & Sons; 1997.
57. Pocock SJ. Clinical Trials. New York: John Wiley & Sons; 1993.
58. Fleiss JL, Levin B, Paik MC. Statistical Methods for Rates and Proportions. 3rd ed. New York: John Wiley & Sons; 2003.
59. Hosmer DW, Lemeshow S. Applied Logisitic Regression. New York: John Wiley & sons; 1989.

60. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Applied Regression Analysis and Other Multivariable Methods. 3rd ed. Belmont, CA: Thomson Brooks/Cole Publishing Co; 1998.
61. Wilson DB. Practical meta-analysis effect size calculator. 2013. URL: http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD-main.php [accessed 2014-08-28] [WebCite Cache ID 6SApXjxtA]
62. Hardin JW, Hilbe JM. Generalized Estimating Equations. Boca Raton, FL: Chapman & Hall/CRC Press; 2003.
63. Liang K, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73(1):13-22. [doi: 10.1093/biomet/73.1.13]
64. Efron B. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrika 1981;68:589-599.
65. Vladutiu CJ, Martin SL, Macy RJ. College- or university-based sexual assault prevention programs: a review of program outcomes, characteristics, and recommendations. Trauma Violence Abuse 2011 Apr;12(2):67-86. [doi: 10.1177/1524838010390708] [Medline: 21196436]
66. Garrity SE. Sexual assault prevention programs for college-aged men: a critical evaluation. J Forensic Nurs 2011 Mar;7(1):40-48. [doi: 10.1111/j.1939-3938.2010.01094.x] [Medline: 21348933]
67. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
68. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
69. Bandura A. Social cognitive theory of mass communication. In: Bryant J, Zillmann D, editors. Media Effects: Advances in Theory and Research. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994:61-90.
70. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
71. DeGue S, Holt MK, Massetti GM, Matjasko JL, Tharp AT, Valle LA. Looking ahead toward community-level strategies to prevent sexual violence. J Womens Health (Larchmt) 2012 Jan;21(1):1-3. [doi: 10.1089/jwh.2011.3263] [Medline: 22185587]
72. Peels DA, Bolman C, Golsteijn RH, De Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between Web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res 2012;14(6):e179 [FREE Full text] [doi: 10.2196/jmir.2229] [Medline: 23246790]
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. 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

AOR: adjusted odd ratios
CTS2: Revised Conflict Tactics Scale
GEE: generalized estimating equation
GPA: grade point average
MAR: missing at random
MCAR: missing completely at random
RCT: randomized controlled trial
ROLB: Reactions to Offensive Language and Behavior

Edited by G Eysenbach; submitted 25.03.14; peer-reviewed by A Hattery, C Allen, S McMahon; comments to author 02.07.14; revised version received 12.07.14; accepted 16.08.14; published 05.09.14.

Please cite as:
Salazar LF, Vivolo-Kantor A, Hardin J, Berkowitz A
A Web-Based Sexual Violence Bystander Intervention for Male College Students: Randomized Controlled Trial
J Med Internet Res 2014;16(9):e203
URL: http://www.jmir.org/2014/9/e203/
doi:10.2196/jmir.3426
PMID:25198417

Short Paper

Facebook Apps for Smoking Cessation: A Review of Content and Adherence to Evidence-Based Guidelines

Megan A Jacobs[1*], MPH; Caroline O Cobb[1*], PhD; Lorien Abroms[2*], ScD; Amanda L Graham[1,3*], PhD

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

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

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

[*]all authors contributed equally

Corresponding Author:
Megan A Jacobs, MPH
Schroeder Institute for Tobacco Research and Policy Studies, Legacy
1724 Massachusetts Avenue, NW
Washington, DC, 20036
United States
Phone: 1 202 454 5753
Fax: 1 202 454 5785
Email: mjacobs@legacyforhealth.org

Abstract

Background: Facebook is the most popular social network site, with over 1 billion users globally. There are millions of apps available within Facebook, many of which address health and health behavior change. Facebook may represent a promising channel to reach smokers with cessation interventions via apps. To date, there have been no published reports about Facebook apps for smoking cessation.

Objective: The purpose of this study was to review the features and functionality of Facebook apps for smoking cessation and to determine the extent to which they adhere to evidence-based guidelines for tobacco dependence treatment.

Methods: In August 2013, we searched Facebook and three top Internet search engines using smoking cessation keywords to identify relevant Facebook apps. Resultant apps were screened for eligibility (smoking cessation-related, English language, and functioning). Eligible apps were reviewed by 2 independent coders using a standardized coding scheme. Coding included content features (interactive, informational, and social) and adherence to an established 20-item index (possible score 0-40) derived from the US Public Health Service's Clinical Practice Guidelines for Treating Tobacco Use and Dependence.

Results: We screened 22 apps for eligibility; of these, 12 underwent full coding. Only 9 apps were available on Facebook. Facebook apps fell into three broad categories: public pledge to quit (n=3), quit-date–based calculator/tracker (n=4), or a multicomponent quit smoking program (n=2). All apps incorporated interactive, informational, and social features except for two quit-date–based calculator/trackers apps (lacked informational component). All apps allowed app-related posting within Facebook (ie, on self/other Facebook profile), and four had a within-app "community" feature to enable app users to communicate with each other. Adherence index summary scores among Facebook apps were low overall (mean 15.1, SD 7.8, range 7-30), with multicomponent apps scoring the highest.

Conclusions: There are few smoking cessation apps available within Facebook. Among those available, adherence to cessation treatment guidelines was low. Smoking cessation interventions provided via the Facebook platform are a unique and as yet untapped treatment strategy that can harness existing social support and social networks for quitting. Research is needed to examine whether apps that adhere to clinical practice guidelines for tobacco dependence treatment are more effective in promoting cessation than those that do not.

(J Med Internet Res 2014;16(9):e205) doi:10.2196/jmir.3491

KEYWORDS

Facebook; smoking cessation/methods; social media; apps

Introduction

Facebook use has become nearly universal and continues to increase. As of December 31, 2013, the social network site has over 1 billion global users [1]. In 2013, 57% of all American adults and 73% of those aged 12-17 years used Facebook [2]. Intensity of use is also escalating: in 2014, 64% of Facebook users visited the site on a daily basis, up from 51% in 2010 [2]. A primary channel through which users interact with Facebook is through millions of third-party software applications (apps) [3]. Facebook apps are available on personal computers and some are also accessible on mobile and/or smartphones. The broad reach and intensive use of Facebook worldwide represents a potentially powerful opportunity to deliver health-related behavior change interventions.

Facebook cessation apps may represent a unique approach to treatment that can leverage the power of social network ties [4,5]. Upon installing an app, their personal profile data and social network graph are typically made available to the app. Real-time access to an individual's social network may facilitate the provision of social support from network members and the spread of an intervention through social networks.

To date, there have been no reviews of the content or quality of Facebook cessation apps. Two reviews of mobile apps by Abroms et al [6,7] found that they were heavily downloaded but most did not adhere to clinical guidelines. Choi et al [8] found cessation mobile apps may be limited as autonomous interventions. The goals of this study were to (1) assess the availability of Facebook apps for smoking cessation, (2) describe their approach and features, and (3) examine their adherence to an index based on the US Public Health Service's Clinical Practice Guideline for Treating Tobacco Dependence (the Guideline).

Methods

Identifying Facebook Apps

Two strategies were used to ensure we located apps likely to be encountered by the typical smoker searching Facebook for cessation assistance [9,10]. We searched for apps within Facebook using the general search toolbar with "Apps/Games" selected and using the App Center general search toolbar using cessation-related keywords (eg, "smoking", "quit", "cessation", "cigarette cessation"). We used these same keywords in combination with "Facebook" in search engine queries on Google, Bing, and Yahoo! Since Internet searchers typically review only the first page of search engine results [8], the first two pages of results for each keyword were reviewed for any links or mention of a Facebook app. Links that led to a Facebook App Center page, directly to an app, or to an app page were saved for eligibility review. Searches were conducted in August 2013.

To determine eligibility for full review, two Master's level coders catalogued basic information for each app link retrieved. This approach to app screening has been used in other reviews [6,7]. To be eligible for full review, apps had to be in English and include text related to cessation either in the Facebook App Center overview or within the app.

Coding of Apps

Eligible apps were installed using the native Facebook platform on a personal computer or the iPhone platform for mobile-only apps. Coders used apps over 3 days with at least 3 logins to ensure all features were utilized and coded. (Co-authors MAJ and ALG were involved in the development of one of the Facebook apps that was eligible for full review. Coders of this app were research staff not directly involved in app development or the associated research grant [R01 CA155369-01A1]). Apps were coded for publisher/developer type, cost, and content features (interactive, informational, and social). Operational definitions of content features were developed prior to coding and were noted as present or absent. Interactivity was defined as "any content-related user input that results in feedback from the website" [10]. Informational features were those specific to smoking cessation, withdrawal symptoms, triggers to smoke/cravings, and/or ways to deal with cravings. Social features included (1) a within-app community with communication features (eg, public posting wall, personal message function), and (2) the ability to post updates or information to a user's Facebook wall to share information with existing social network ties. Using an existing scheme [6,7], apps were classified as public pledge (eg, choose a reason for quitting and post to Facebook wall), calculator/tracker (eg, money saved/cigarettes not smoked calculated based on quit date), or multicomponent (eg, contained a calculator/tracker feature and quitting guide).

Apps were then coded for adherence to the Guideline [11] using a modified version of an established index [6,7]. Modifications included (1) removal of "Recommend counseling and medicines" and "Refer to recommended treatment" because they duplicated other "ASSIST with a quit plan" items, (2) editing of "Text alerts" to "Notifications (any type including text)" given alternative notification options within Facebook, and (3) addition of "Recommend talking to your health care provider about quitting" as key feature of cessation treatment (Table 1). Items were scored on a 3-point scale (0=not at all present, 1=partially present, or 2=fully present). For example, for the guideline to "ARRANGE for follow-up", apps that did not mention any follow-up or send an invitation to return to the app received a 0; apps that either mentioned follow-up or sent an invitation to return received a score of 1; and apps that did both received a score of 2. For items such as "Enhance motivation: rewards", apps that referred to specific rewards (eg, whiter teeth) received a 2; apps that referred only to generic rewards (eg, "quitting has lots of benefits") received a 1; and apps that did not refer to any rewards received a 0.

Similar to Abroms et al [7], when scores between coders differed by 2 points (ie, 0 vs 2), coders discussed the discrepancy and recoded the item. After recoding, if item scores differed by 1 point or less, two scores were averaged; all item scores for each app were summed (maximum adherence score possible=40).

Data Analysis

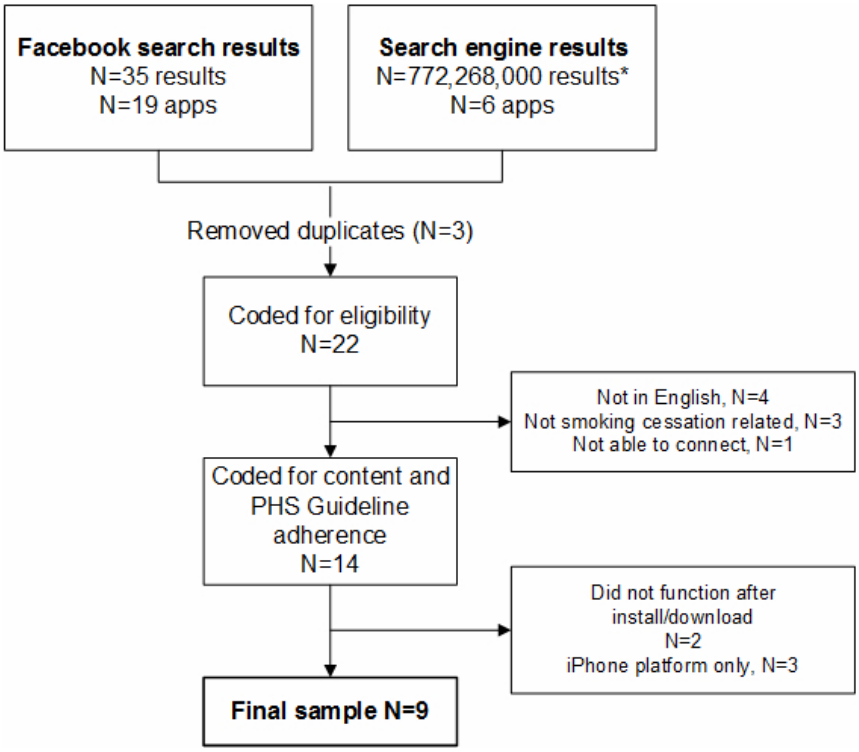
Frequencies and descriptive statistics were used to characterize apps. For each adherence index item, the proportion of apps receiving a 2 (fully present) by at least one coder was calculated (average score≥1.5; as in [7]). The relationship between publisher/developer source and app type on the adherence index summary score was examined. All analyses were performed in SPSS Version 21.

Results

Summary

The search resulted in 19 apps from within Facebook's internal search platform and 6 apps from Internet search engines (Figure 1). After removing duplicates (n=3), those available only on iPhone (n=3), and ineligible (n=8) and non-functional (n=2) apps, the final sample included 9 apps that were available on Facebook (see Multimedia Appendix 1).

Figure 1. Study flow diagram from search results to final sample of apps reviewed (total number of results from Yahoo! search engine query were not available).



Facebook App Characteristics

Three apps were sponsored by a pharmaceutical company, three were funded/sponsored by a government entity, and three originated from an individual/private company or were unclear. As required by Facebook [12], all apps were free to install. All apps were interactive, and most contained informational features (7/9). Only four apps had a within-app "community" feature to enable app users to communicate with each other. All apps allowed app-related posting within Facebook (ie, on self/other Facebook profile). Four apps were categorized as calculator/tracker apps, three were related to a public pledge to quit, and two were coded as multicomponent (see Multimedia Appendix 2 for representative screenshots from each category).

Adherence Index Results

Reviewer agreement was 67% or better for 17 of 20 items. Agreement for the remaining 3 items fell between 44% and 56%: "ASSIST with a quit plan: supplementary information", "Enhance motivation: risk", and "Enhance motivation: roadblocks". Table 1 displays the proportion of apps that received a "fully present" rating by at least one coder (ie, average score≥1.5 for a given item).

Only 3 of 20 adherence index items were fully present in the majority of apps: "Specific to smoking", "Enhance motivations: Rewards", and "Interactive". The average adherence index score was 15.1 (SD 7.8; N=9). Government sponsored/funded apps scored higher than other publisher/developer types, and multicomponent cessation apps scored higher than calculator/tracker and public pledge to quit apps (Table 2).

Table 1. Adherence index item endorsement across Facebook apps (N=9).

Adherence index item	n (%) of apps[a]
1. Specific to smoking	9 (100)
2. ASK for smoking status	4 (44.4)
3. ASSESS willingness to quit	2 (22.2)
4. ADVISE every user to quit	3 (33.3)
5. ADVISE every user to quit: personalized	2 (22.2)
6. ADVISE every user to quit: clear	2 (22.2)
7. ADVISE every users to quit: strong	2 (22.2)
8. ASSIST with a quit plan: overall plan	2 (22.2)
9. ASSIST with a quit plan: practical counseling	3 (33.3)
10. ASSIST with a quit plan: recommend approved meds	3 (33.3)
11. ASSIST with a quit plan: intra-treatment social support	4 (44.4)
12. ASSIST with a quit plan: supplementary information	1 (11.1)
13. ARRANGE for follow-up	1 (11.1)
14. Enhance motivation: rewards	7 (77.8)
15. Enhance motivation: risks	3 (33.3)
16. Enhance motivation: roadblocks	3 (33.3)
17. Connect to a quitline	1 (11.1)
18. Interactive	9 (100)
19. Notifications (any type including text)	1 (11.1)
20. Talk with healthcare provider re: quitting	1 (11.1)

[a]Proportion of apps with at least one coder indicating item was “fully present” (total N=9).

Table 2. Facebook app adherence index summary score by publisher/developer and category (N=9).

Review category	Mean (SD)[a]
By publisher/developer	
Individual/other (n=3)	13.3 (9.8)
Pharmaceutical company (n=3)	15.5 (1.3)
Government group (n=3)	16.3 (11.9)
By category	
Calculator/Tracker (n=4)	11.6 (8.7)
Public pledge to quit (n=3)	14.5 (1.8)
Multicomponent cessation program (n=2)	22.8 (9.5)

[a]Summary score across all adherence index items; absolute range 0-40.

Discussion

Principal Findings

We identified only 9 English-language Facebook smoking cessation apps. Overall, apps scored poorly on adherence to clinical recommendations for tobacco dependence treatment. Referrals to quitlines and health care providers were infrequent. Consistent with earlier reviews [6,7], treatment components completely absent from most Facebook apps included personalized advice to quit and notifications. These omissions are noteworthy given the ease with which Facebook makes it possible to personalize app experiences and send notifications.

All apps included in-app social support and social media integration. Facebook’s social environment may make it easier for apps to leverage these tools relative to other kinds of mobile app platforms. Future research should explore the use and effectiveness of app-related social support and examine the extent to which users have concerns about sharing smoking/quitting information with their Facebook network [13]. Mobile and Facebook app designs for smoking cessation or

other health behavior change could benefit from knowing which components correlate with effectiveness.

Limitations

A limitation of the study is that apps were scored using an adherence index based on clinical interventions delivered by health care providers [7], which may not be the most suitable criteria for technology-based approaches to smoking cessation. Also, the use of social networking within apps was not explored. Future app reviews should specifically consider social networking features and the unique social network elements of Facebook apps. A strength of the study was the use of apps for a minimum of 3 days by two independent coders to ensure that all content and functionality in the app were experienced, including elements that revealed themselves after several uses. Our methods add important information to other studies [14,15] that have not employed this approach.

Conclusions

Facebook apps for cessation are few in number and generally lack core components of evidence-based cessation treatment. Research is needed to determine whether adherence to treatment guidelines is related to effectiveness in promoting cessation.

Acknowledgments

MAJ, COC, and ALG participated in study concept and drafting of the manuscript. MAJ, COC, ALG, and LA participated in study design and made comments on the manuscript. MAJ and COC participated in acquisition of data, COC participated in statistical analysis, and ALG participated in obtaining funding.

Primary funding for this work was from the National Cancer Institute at the National Institutes of Health (R01 CA155369-01A1). The funding agency had no involvement in the conduct of the study or preparation of this manuscript.

We thank Sarah Cha, MSPH, and Allison Glasser, MPH, for their work in coding the UbiQUITous Facebook app.

Conflicts of Interest

L Abroms/The George Washington University has licensed the Text2Quit program to Voxiva, Inc. Text2Quit is a text messaging program for smoking cessation. C Cobb, M Jacobs, and A Graham are employees of Legacy, a non-profit public health foundation that runs BecomeAnEX.org, an online tobacco cessation intervention. M Jacobs and A Graham are funded in part through R01 CA155369-01A1, which studies the viral spread of a cessation intervention through Facebook.

Multimedia Appendix 1

Final sample of Facebook and iPhone apps coded.

[PDF File (Adobe PDF File), 2KB - jmir_v16i9e205_app1.pdf]

Multimedia Appendix 2

Representative screenshots from apps in each category.

[PDF File (Adobe PDF File), 762KB - jmir_v16i9e205_app2.pdf]

References

1. Facebook Inc. Facebook Company Info. URL: http://newsroom.fb.com/company-info/ [accessed 2014-04-23] [WebCite Cache ID 6P390owud]
2. Smith A. 6 new facts about Facebook.: Pew Research Center URL: http://www.pewresearch.org/fact-tank/2014/02/03/6-new-facts-about-facebook/ [accessed 2014-06-18] [WebCite Cache ID 6QQKqQY1Z]
3. Facebook Inc. Form 10-K Annual Report. 2014. URL: http://files.shareholder.com/downloads/AMDA-NJ5DZ/2301311196x0xS1326801-13-3/1326801/1326801-13-3.pdf [accessed 2014-04-09] [WebCite Cache ID 6Oi6MlOpC]
4. Cobb NK, Jacobs MA, Saul J, Wileyto EP, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomised controlled trial study protocol. BMJ Open 2014;4(1):e004089 [FREE Full text] [doi: 10.1136/bmjopen-2013-004089] [Medline: 24448847]
5. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13(4):e119 [FREE Full text] [doi: 10.2196/jmir.1911] [Medline: 22182518]
6. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285 [FREE Full text] [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]
7. 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]
8. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]

9. Bock B, Graham A, Sciamanna C, Krishnamoorthy J, Whiteley J, Carmona-Barros R, et al. Smoking cessation treatment on the Internet: content, quality, and usability. Nicotine Tob Res 2004 Apr;6(2):207-219. [doi: 10.1080/14622200410001676332] [Medline: 15203794]
10. Bock BC, Graham AL, Whiteley JA, Stoddard JL. A review of web-assisted tobacco interventions (WATIs). J Med Internet Res 2008;10(5):e39 [FREE Full text] [doi: 10.2196/jmir.989] [Medline: 19000979]
11. Fiore M, Jaén C, Baker T. Treating tobacco use and dependence: 2008 Update: Clinical Practice Guideline. Darby, PA: DIANE Publishing; 2008.
12. Brady A. Facebook Developer Blog. 2014. Introducing the App Center URL: https://developers.facebook.com/blog/post/2012/05/09/introducing-the-app-center/ [accessed 2014-04-09] [WebCite Cache ID 6Oi6Yh3f5]
13. Ramo DE, Liu H, Prochaska JJ. A Mixed-Methods Study of Young Adults' Receptivity to Using Facebook for Smoking Cessation: If You Build It, Will They Come? Am J Health Promot 2014 Feb 27:-. [doi: 10.4278/ajhp.130326-QUAL-128] [Medline: 24575728]
14. 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;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
15. O'Neill S, Brady RR. Colorectal smartphone apps: opportunities and risks. Colorectal Dis 2012 Sep;14(9):e530-e534. [doi: 10.1111/j.1463-1318.2012.03088.x] [Medline: 22646729]

Edited by G Eysenbach; submitted 23.04.14; peer-reviewed by D Ramo, J Choi, M Zhang; comments to author 15.05.14; revised version received 01.07.14; accepted 26.08.14; published 09.09.14.

Please cite as:
Jacobs MA, Cobb CO, Abroms L, Graham AL
Facebook Apps for Smoking Cessation: A Review of Content and Adherence to Evidence-Based Guidelines
J Med Internet Res 2014;16(9):e205
URL: http://www.jmir.org/2014/9/e205/
doi:10.2196/jmir.3491
PMID:25205129

Original Paper

Norwegian General Practitioners' Perspectives on Implementation of a Guided Web-Based Cognitive Behavioral Therapy for Depression: A Qualitative Study

Maja Wilhelmsen[1], MD; Ragnhild Sørensen Høifødt[2], PsyD; Nils Kolstrup[1], MD, PhD; Knut Waterloo[2], PhD; Martin Eisemann[2], PhD; Richard Chenhall[3], PhD; Mette Bech Risør[1], PhD

[1]Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

[2]Department of Psychology, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

[3]Center for Health Equity, School of Population and Global Health, University of Melbourne, Melbourne, Australia

Corresponding Author:
Maja Wilhelmsen, MD
Department of Community Medicine
Faculty of Health Sciences
UiT The Arctic University of Norway
Breivika
Tromsø, 9037
Norway
Phone: 47 99001559
Fax: 47 77645291
Email: maja.wilhelmsen@uit.no

Abstract

Background: Previous research suggests that Internet-based cognitive behavioral therapy (ICBT) has a positive effect on symptoms of depression. ICBT appears to be more effective with therapist support, but it is unclear what this support should comprise. General practitioners (GPs) have positive attitudes toward ICBT. However, ICBT is rarely used in regular care in general practice. More research is warranted to integrate the potential of ICBT as part of regular care.

Objective: The aim of this study was to explore aspects perceived by GPs to affect the implementation of guided ICBT in daily practice. Understanding their perspectives may contribute to improving the treatment of depression in the context of general practice.

Methods: A training package (3-day course) introducing a Norwegian translation of the ICBT program MoodGYM was developed and presented to GPs in Norway. Following training, GPs were asked to include guided ICBT in their regular care of patients with symptoms of depression by providing brief, face-to-face follow-up consultations between modules. We interviewed 11 GPs who had taken the course. Our interview guide comprised open questions that encouraged GPs to frame their responses using examples from their experiences when implementing ICBT. Thematic analysis was chosen to explore patterns across the data.

Results: An overall belief that ICBT would benefit both the patients' health and the GPs' own work satisfaction prompted the GPs to take the ICBT course. ICBT motivated them to invest time and effort in improving treatment. The most important motivating aspects in MoodGYM were that a program based on cognitive behavioral therapy could add a structured agenda to their consultations and empower depressed patients. Organizational aspects, such as a lack of time and varied practice, inhibited the use of ICBT. Inadequate knowledge, recalling the program, and changing own habits were also challenging. The GPs were ambivalent about whether ICBT had a negative impact on the doctor–patient interaction in the module follow-ups. Generally, GPs made an effort to recommend MoodGYM, but the expected module follow-ups were often not provided to patients and instead the GPs returned to standard treatment.

Conclusions: GPs' feedback in the present study contribute to our understanding of the challenges of changing treatment for depression. Our findings indicated that recommending ICBT could add to the GP's toolkit. Offering training and highlighting the following aspects may increase recommendation of ICBT by GPs: (1) ICBT is theory-based and credible, (2) ICBT increases the GPs' work satisfaction by having a tool to offer, and (3) ICBT facilitates empowerment of patients in their own health. In

addition, the present study also indicated that complex aspects must be accommodated before module follow-ups can be incorporated into GPs' treatment of depression.

(J Med Internet Res 2014;16(9):e208) doi:10.2196/jmir.3556

KEYWORDS

mental health; Internet; telemedicine; qualitative research; primary health care; cognitive therapy; depression

Introduction

Overview

Every year, it is estimated that more than one-third of the European population suffers from mental disorders, with depression and anxiety being the most frequent [1,2]. Depression imposes tremendous emotional, financial, and social burdens for patients, their families, and society [3]. Mental disorders have been suggested as one of the biggest health challenges because of deficiencies in available treatment and poor service provision. Rethinking our provision of mental health care is needed [1,4]. Mental health patients tend to prefer consultations with a therapist to prescribed medication [5,6]. Standard cognitive behavioral therapy (CBT) in consultations in routine specialized mental health services is effective but time consuming [7], making this therapy inaccessible to many.

Internet-Based Treatment of Depression

In multiple trials, the use of Internet-based CBT (ICBT) has shown promising results in treating depression (eg, [8-11]). ICBT can be self-administered or supported by either minimal-contact follow-ups or by guided follow-ups that focus on process issues. It is unclear what the support should comprise [12]. However, guided ICBT appears to be more effective than self-administered ICBT [8,9,13]. Guided ICBT may also be as effective as standard face-to-face therapy [9,14]. Furthermore, patients value being active agents using ICBT, although they also emphasize the helpfulness of a relationship with a trusted clinician [15]. In Norway, an increasing proportion of the population uses the Internet for health purposes, and in 2010 the proportion was 77% [16]. This may indicate that the population is amenable to supported online supplements, such as guided ICBT.

Treatment of Depression in Primary Care

Research has recommended that mental illness should be detected and treated early, before more severe expressions can occur [1,17]. GPs are often the first point of professional contact for individuals with depression, and most of these patients are currently treated in primary care [1,18]. In Norway, the mental health provider in primary care is most often the GP [19]. In guidelines from the National Institute for Health and Clinical Excellence [20] and in Norwegian national guidelines for treatment of depression [19], it is recommended initially to apply low-intensity, non-pharmacological approaches such as CBT-based self-help or online interventions in the treatment of mild to moderate depression. Studies have found that GPs have positive attitudes toward such interventions [5,21,22]. However, other studies have shown that GPs rarely recommend evidence-based self-management or eHealth programs to patients with depressive symptoms [18,22].

ICBT has been suggested to be effective in primary care even if the provider (ie, the GP) lacks extensive specialized training [6,9,23-25]. MoodGYM is an ICBT program developed at the Centre for Mental Health Research at the Australian National University. It has been proven in trials to be effective in alleviating depression as a self-administered self-help program [26] and as guided ICBT for patients from primary care [10]. However, there is a gap in the research evidence related to the conclusions of trials and the knowledge about how to implement new methods in regular care [27,28]. More research is needed to reduce this gap and to explore the potential of ICBT deployed in everyday clinical settings [21,23]. Little is known about trained GPs' experiences of implementing guided ICBT into regular care.

A wealth of theories have been developed to explain aspects affecting the implementation of innovations in health care [29]. We chose normalization process theory (NPT), developed by May and Finch [27,30], as a framework when investigating implementation of ICBT because NPT is derived from multiple qualitative studies exploring the implementation of complex interventions and eHealth contextualized in regular health practice. NPT suggests that, for the health professional, successful implementation depends on a complex interplay of the following four main components of "work": (1) coherence, whereby the participants make sense of an intervention, (2) cognitive participation, which requires the participants to engage in the intervention, (3) collective action, which requires that efforts are made to enable the intervention to happen, and (4) reflexive monitoring, comprising formal and informal appraisals of the benefits and costs of the intervention. Small improvements in the treatment of depression in general practice can have positive consequences for many patients. The aim of this study was to explore those aspects perceived by GPs to affect their implementation of guided ICBT in daily practice. Understanding their perspectives may contribute to improving the treatment of depression in the context of general practice.

Methods

Study Context

A training package based on a Norwegian translation of the ICBT program MoodGYM was developed and presented by a GP (NK) and two psychologists (RSH and Kjersti Lillevoll) as a 3-day course for GPs in spring 2011. MoodGYM is a free Internet-based self-help program comprising five interactive modules introducing cognitive behavioral principles through online exercises. MoodGYM demonstrates the relationship between thoughts and emotions and teaches relaxation techniques. It also includes sections on managing relationships and increasing engagement in positive activities. A sample screenshot from MoodGYM is presented in Figure 1. The course

consisted of an introduction to CBT principles, presentation of and a group session on MoodGYM's content, and presentation of a manual for follow-ups. The manual was also supplied online and contained a short summary of each module and suggestions for follow-up questions. In the final session, a patient described his experiences with guided ICBT.

The course recommended guided ICBT as follows: (1) to recommend MoodGYM to patients with symptoms of depression and subsequently (2) to provide module follow-ups comprising brief and structured face-to-face consultations between the online modules (see Figure 2). The follow-ups were meant to be of a motivating nature, by offering a dialogue on process issues and allowing some time for reflection. Process issues were related to working with and understanding principles presented in the online modules such as what was difficult and what was useful. Follow-ups were suggested as 20-minute consultations every second week. GPs who currently worked in general practice and who completed the course all signed up for further research on implementation of the guided ICBT.

Figure 1. Sample screenshot from MoodGYM.

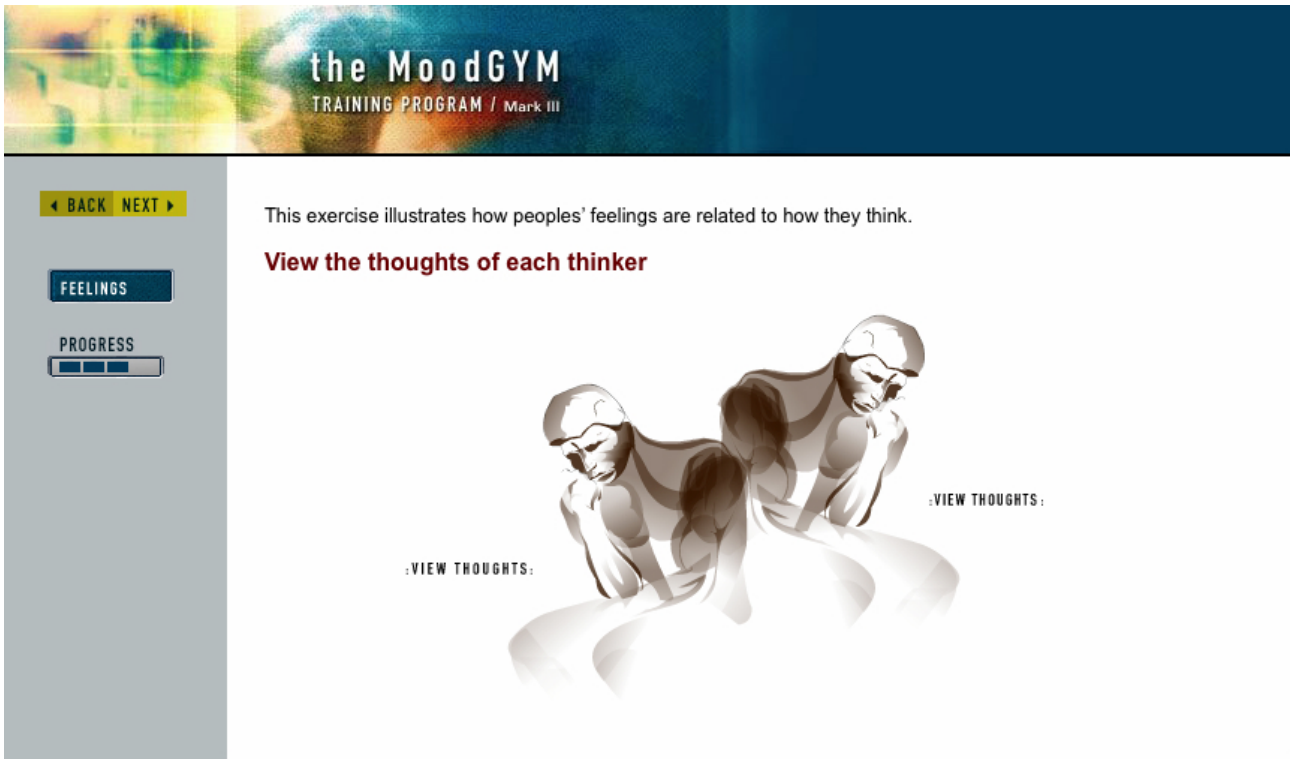
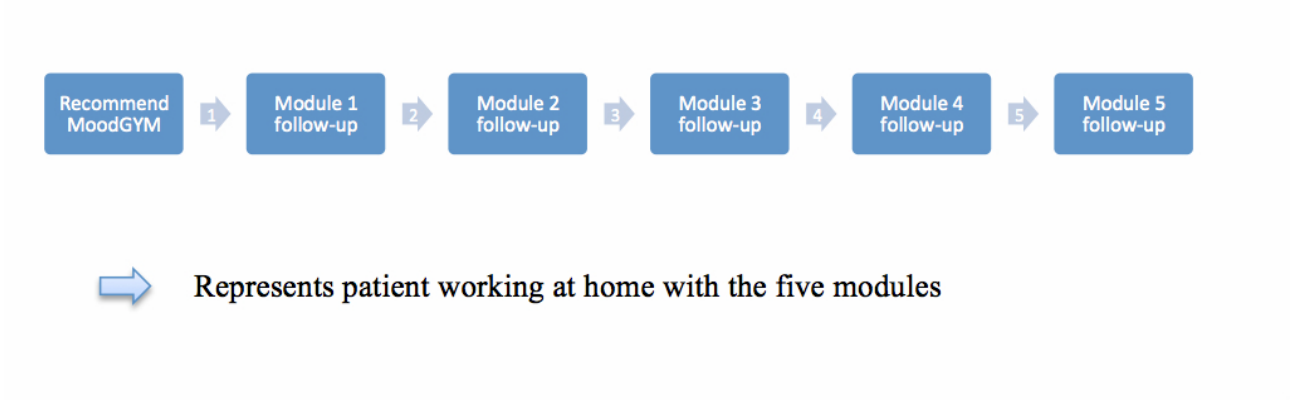


Figure 2. ICBT with module follow-ups.



Participants

A purposive sample [31] of 11 GPs from northern Norway was included, and interview arrangements were made over the telephone. The study originally included GPs who had enrolled voluntarily in the guided ICBT course. In the process of recruiting, we had the opportunity to include two additional GPs who had attended a 3-hour presentation of this treatment model given by one of the GPs who had attended the course. These two GPs were included in the analysis primarily for comparative purposes, to add an extra dimension to the analysis of the GPs' experiences with ICBT. Participants were both men and women, of various ages, and with various lengths of experience as GPs (Table 1).

Table 1. Participant characteristics.

GP	Age	Date of Interview	Duration of interview, minutes	Experience as a GP, years	Recommend MoodGYM	Gender	3-day course	Interview venue
1	52	06.02.13	40	21	yes	male	yes	office
2	35	14.02.13	50	3.5	yes	female	yes	home
3	53	20.02.13	33	21	yes	male	no	office
4	33	25.02.13	35	3.5	yes	female	yes	home
5	38	25.05.13	47	8	yes	female	yes	home
6	58	12.03.13	70	28	no	female	yes	office
7	42	14.03.13	85	3	yes	female	yes	office
8	54	20.03.13	75	10	yes	female	yes	university
9	45	08.04.13	76	12	yes	female	yes	office
10	51	12.09.13	60	12	yes	female	no	office
11	34	19.09.13	83	3.5	yes	female	yes	office

Interview Schedule and Data Collection

The female interviewers (MW and RSH) conducted semistructured in-depth interviews designed to gather information about (1) the GPs' general views on their work with depressed patients, (2) motivational aspects for learning ICBT, (3) experiences from implementing guided ICBT, and (4) implications of the use of ICBT for consultation quality and patient-doctor interaction. The interview guide (see Multimedia Appendix 1) was developed by the interdisciplinary author group combining prior experience within eHealth research, clinical experience from general practice, and long experience from medical anthropology. A pilot was not conducted; however, adjustments were made after three interviews. The venue was flexible, and GPs chose their home, office, or the University of Tromsø. Individual interviews lasted 33-85 minutes and were audio recorded. The interviews were conducted as face-to-face dialogues, and open-ended questions were used to evoke descriptions of situations grounded contextually in everyday events from their clinical practice. The emphasis was on working together in the interview to reach an understanding of the GPs' experiences that reflected their perceptions and attitudes [31]. The GPs were not invited to give feedback later in the research process. Field notes were recorded immediately after interviews.

Data Analysis

Interviews were transcribed into NVivo 10 data analysis software and anonymized. Thematic analysis is an accessible stepwise method often used to investigate patterns across datasets [32]. It is also a flexible method compatible with an inductive, data-driven approach. We chose thematic analysis as a method because it is suitable when exploring patterns in stories that describe experiences and meanings. Our aim was to explore all aspects perceived by the GPs to have affected their implementation of ICBT, and thus, an inductive approach was chosen. The initial analysis involved repeated readings of each transcript to obtain an overall impression. The next phase involved coding the entire dataset on a semantic level. Events, thoughts, and actions were coded as themes on the basis of their ability to capture content with meanings that were important in relation to the overall research question. Subsequently, we identified and interpreted overarching themes in a constant process of moving between the data, potential themes, and maps made for visualization. A storyline extending from what initially attracted the GPs to take the course to how they applied ICBT in the everyday clinical setting was identified. We worked toward elucidating themes that were internally homogeneous and externally heterogeneous and had explanatory power. Finally, the themes were considered if they were coherent and represented the meanings found in the interviews [32]. Throughout the analytical process, all findings were discussed and validated with an experienced qualitative researcher (MBR) and two GPs (MW and NK). Inconsistencies were resolved through discussion and further reflection. All transcripts have been checked against the audio recordings. Inspired by NPT [27,30], the researchers were sensitized to components in interplay that are suggested as important during implementation of the complex interventions. Further, findings were discussed in light of existing literature.

Ethics Approval

Written informed consent was obtained from all participants. Ethical approval was given by the Regional Ethical Committee, Tromsø 2011/2163.

Results

Key Themes

We identified four key themes and present here the common storyline extending from what initially attracted the GPs to take the course to how they applied ICBT in the everyday clinical setting. The four key themes are: "GPs wanted to improve treatment and work satisfaction", "The value of a structured agenda and the patient as an active agent in treatment", "Constraints of hectic practice, inadequate knowledge, and competing tasks", and "GPs recommended it, but few module follow-ups were provided".

GPs Wanted to Improve Treatment and Work Satisfaction

GPs reflected on what initially attracted them to take the course and how they could make sense of a change in the treatment of depression they normally provided, in the following referred to as standard treatment. GPs noted that symptoms of depression were common and treatment was considered part of their job. They had inadequate access to specialist care, and thus referral was often difficult. GPs hoped that implementation of an online intervention would benefit patients with mental health issues. Comments that implied a need to improve GPs' own treatment were prominent:

> *[We give] what we call supporting consultations in general practice, but actually it is just talking without any…I wouldn't say meaning, but no concrete agenda in a way. I mean, there is no…It is just: How is it going? What have you been doing lately? It is more the patient talking in a free way about how they are and stuff. But no actual therapy really.* [Participant 11]

> *Many times I feel that I don't have an agenda or further program other than these supportive consultations, which I often feel it [depression treatment] ends up with.* [Participant 5]

GPs used devaluating words such as "only supportive consultations" or "just talking" when they referred to standard treatment, and some described the situation with frustration, indicating concerns about the quality of care. They hoped that supplementing depression treatment would result in improved health and quicker recovery for their patients.

Standard treatment was described as self-taught, informal, and unstructured. The frustration identified might also imply a sense of lack of control and vulnerability. Others expressed that their treatment had improved with experience and was helpful, but they felt they needed something more to offer their patients: "And I think many can feel the frustration about what should we do. How can we help these patients [with depression]? In what way?...To me it felt very good [to learn ICBT]! Because now I finally felt I had some treatment I could try" [Participant 11].

The word "tool" was often used to describe how implementing ICBT could supplement treatment. This indicated another important aspect of the GPs' reasons for enrollment; that is, they wanted to improve their work satisfaction by increasing their competence and confidence by adding a concrete instrument to their treatment of depression: "So it is, it is nice to have a tool to offer people. That in itself makes you feel better as a caregiver. You feel more professional!" [Participant 1].

According to NPT, coherence is experienced if the participants find an intervention to be a good idea and they are able to define the intervention [27,30]. The GPs showed coherence as they found recommending an online intervention to be a good idea and felt that competence in delivering this type of intervention would be helpful for the patient and of benefit to the GPs. When they described the use of MoodGYM, they talked about it as a supplemental tool, but they did not explicitly define what module follow-ups should comprise and why module follow-ups would improve treatment. This might indicate limited coherence for the module follow-ups.

The Value of a Structured Agenda and the Patient as an Active Agent in Treatment

Internet self-help was regarded as a good idea. The GPs commented further on what aspects of MoodGYM engaged them, enabling them to invest time and effort. Overall, the GPs valued that MoodGYM added a structured agenda to treatment in two ways. First, GPs described MoodGYM's educational content enthusiastically because it is systematic and based on CBT and therefore added a theoretical basis to the treatment: "I find it [MoodGYM] to be…it is very systematic and built up in a good way with the different modules...Yes, I think it is really great and it helps…it is a good help for me as a GP to have something to recommend so that the patients can be educated" [Participant 7] and "To in a way get theory in and...That you [patient] have an opportunity to go into a webpage where there are concrete tasks. I found that positive" [Participant 11].

GPs had confidence in the developers and felt more skillful when able to recommend an evidence-based Web intervention, which implied they felt more professional in their ability to address depression. Compared with the others, the two GPs (Participants 3 and 10) who had not taken the course but had only been to a presentation were more reluctant to tell patients and colleagues that they acknowledged the content as credible and appreciated the structure of ICBT. This may indicate that the course strengthened confidence and the perceived credibility of MoodGYM.

GPs were familiar with referring to the Internet for other conditions. Convenience was mentioned as a positive aspect of recommending MoodGYM because it is free and available, which means it can fit easily into depressed patients' lives. GPs commented that they could help solve their patients' mental health issues only through the efforts of the patients themselves. Patients were regarded as their own best helpers, and the GPs found that ICBT could facilitate self-help. This implied a second aspect of what the GPs found valuable: viewing MoodGYM as a structured agenda to empower patients as active agents to contribute to their own health and recovery. "You give the responsibility to the patient, the responsibility to become better. You place the responsibility. It doesn't work like [the patient can] just come to me and be cured of his depression. He [the patient] needs to do it himself" [Participant 4].

Evaluation was needed to find patients suitable for ICBT, and different characteristics were considered to contribute positively to the success of ICBT. Patients who accepted that their problem was psychological and who could reflect on their own condition were considered suitable for ICBT. A young age was considered positive because young people were assumed to be more computer literate and because the style of MoodGYM appealed to young people. Motivation and initiative were essential: "The people [who] I think will be able to make MoodGYM useful are those who are self-motivated. Those who really want it. And…I mean, they need to have some determination in them. Then I think it will be useful" [Participant 4].

The severity of depression also had to be considered because depression by itself has negative influences on initiative, concentration, and motivation. GPs considered the program best suited to people with mild to moderate depression.

Theory-based structured agenda in the treatment and empowerment of patients were the most important aspects to promote engagement. This engagement relates to the NPT component of cognitive participation [27,30].

Constraints of Hectic Practice, Inadequate Knowledge, and Competing Tasks

Engaging aspects of the program were reported; however, GPs also discussed several challenges in using ICBT. One challenge mentioned was patients being negative to ICBT or not adhering. In the organizational context of general practice, lack of time in the consultation was also a recurrent constraint. The structure of a typical day comprised a tight appointment schedule, no time for preparation between patients, and 20-minute consultations. GPs commented that this made it difficult to take the time needed to think about new treatment methods and remember to recommend MoodGYM to patients: "You already have so little time in the consultation in general practice that everything has to be ready when the patient comes in…It is just opening the book to check what they are coming for. And then things must be ready…It [guided ICBT] isn't done in 20 minutes" [Participant 5].

Subsequent to recommending MoodGYM, the GPs were asked to provide module follow-ups; however, they described this as a difficult task. Lack of time worsened as a patient often has several issues to discuss in follow-up consultations. Given that the GPs saw a variety of patients, it could also be a long time between each encounter with a suitable patient. The lack of continuity made it difficult for the GPs to become familiar with and knowledgeable about MoodGYM, and thus it was demanding to change their habits. Inadequate knowledge about the modules and prioritizing time to go through modules themselves was a challenge because of other competing tasks and lack of incitement to do so. A need for detailed knowledge indicated that they had high standards for the competence needed to understand and apply MoodGYM's content and to talk about the modules with their patients. Lack of practical training of module follow-ups in the course was noted as an element that made it difficult to integrate it to a clinical setting. The hectic day also made it difficult to find and apply the manual for module follow-ups: "I should have done it all [entire MoodGYM] so I would know what they [the patients] have been through each time…But I haven't had the time yet" [Participant 4] and "You had made such nice manuals for each module…I felt it was a mess in my things…And it is about me as well, but I think I should have known it [each module] better!" [Participant 5].

Ambivalent findings were identified on how module follow-ups might affect the interaction with the patients. Dialogue about the programs' content could give a common platform to talk about. On the other hand, a focus on process issues from working with MoodGYM was thought to inhibit other tasks perceived to be important in standard treatment, such as the open dialogue:

> *I find it a barrier [to give module follow-ups]…Yes, it breaks in a way with how you normally work in a way. So I think that is one of the main reasons for not.* [Participant 9]
>
> *Interviewer: What kinds of elements are most important in the treatment you give?*
>
> *No, it must be to listen to the patients…They must feel safe…must be safe to meet me. I think that is the most important to have a relationship. And you must have a relationship to get a dialogue, which enables reflection of the patients themselves…The fact that they get a recommendation to MoodGYM and that, it doesn't influence the actual conversation [in further consultations]*. [Participant 9]

Overall, many inhibiting aspects on engagement were identified: limited time, lack of detailed knowledge about the modules, the variety of patients seen in the practice, and competing tasks in the consultation. These constraints applied strongly to the module follow-ups. The value of discussing process issues in the consultations was considered problematic. This may indicate little cognitive participation similar to limited coherence [27,30] in this part of the treatment.

GPs Recommended It, But Few Module Follow-Ups Were Provided

All GPs tried to implement MoodGYM as best as they could remember and apply it within a hectic day. Many had the Internet address ready to give to patients. To various extents, they found suitable patients and took the time to recommend MoodGYM. One exception was a GP (Participant 6) who chose not to recommend MoodGYM and explained that it was too demanding to change her treatment habits, although she felt that this would have been different if she had been younger. Other feedback included "If I were to estimate [how many patients I have recommended it to] I would think 25" [Participant 1] and "And I have recommended it to some patients. Especially young people who are on the Internet all the time" [Participant 3].

GPs applied strategies to promote the use of MoodGYM. Motivation was essential and to increase patients' motivation, GPs emphasized the importance of recommending it in a convincing way: "You need to sell it like this; it is made in Australia, I say. It has had great success. It is translated into Norwegian and YES! You need to try this! I really say it like that" [Participant 1].

By "convincing", GPs meant communicating with confidence, emphasizing aspects such as that they acknowledged the program as credible and that they had competence in MoodGYM because they had taken a course. They had positive experiences making the patient feel safe by suggesting strategies to ensure anonymity and noting that it was an evidence-based program. The availability and user-friendly nature of the program were also positive aspects mentioned by the GPs. They seemed to provide a rationale for MoodGYM when they recommended it, in the context of a trusting relationship with the patient.

After recommending MoodGYM, follow-ups were conducted in three different ways. One way was to recommend MoodGYM

and not mention the program in further consultations. This way was explained as not being conscious about the intervention, and sometimes they did not record the recommendation making it impossible for them to remember to follow up: "But I haven't, I just realized, I haven't got any feedback from any of the patients…And I believe I haven't written in the records that I have given the recommendation. So, I wouldn't know to ask if they have tried it" [Participant 10].

A second way, described only by one GP (Participant 7) was to deploy the module follow-up manual in consultations in a structured motivating way. When the patients replied to questions about process issues, this GP noted that they told about episodes in life where principles from MoodGYM were applied. However, both the GP and the patient became impatient because time shortage and role conflicts were a problem. The module follow-ups revealed a struggle in the dialogue with patients because both the GP and patients wanted to explore problems from the patient's everyday life as a starting point instead of discussing process issues related to MoodGYM. When problems in life were the starting point for the dialogue, neither the patients nor the GP were able to link it back to MoodGYM:

> *My role [with guided ICBT] was then in a way to motivate people and explore how it is going and…Like have you had…In the concrete program, right…Because first we were meant to talk about the modules and the things around that, but automatically we started talking about…It was totally unnatural for me not to ask: "How is it going at work?...She [the patient] was very interested in telling me…or wanted to talk about it [her mother and work] and…it was a very difficult role to have…and she became a little impatient…it didn't work very well, that's my opinion.* [Participant 7]

A third and most common way was to recommend MoodGYM but then limit the follow-up of the program to asking the patients if they had done it or not. Neither GPs nor the patients initiated talking about process issues apart from a few patients who gave brief feedback on whether they liked the program. GPs appreciated and encouraged the independent work undertaken by patients using MoodGYM. They admitted to not giving the program much attention in further consultations. As GPs made little effort to obtain feedback on MoodGYM from their patients, they also had little knowledge about how the patients perceived working at home with ICBT:

> *And to ask if they have gone into it [MoodGYM]. If it [depression] expanded in time…should they have sick leave? Have I done, not missed anything, right? And then…I ask about how it is going.* [Participant 9]
>
> *Interviewer: Do any of you…do you bring up or do you discuss anything from the program [MoodGYM] in your consultations?*
>
> *No, I haven't done it…I don't get to know anything, but I just get to know if they liked it or not.* [Participant 9]

Despite GPs' acknowledgement of the insufficiencies of standard treatment and regarding guided ICBT as a way to achieve improvement, paradoxically they returned to standard treatment instead of module follow-ups in subsequent consultations. Time constraints were also a challenge with standard treatment, although efforts to be flexible were made. Central to standard treatment was to support the patients telling their stories. GPs could recount these stories in detail in the interviews. Supportive tasks, such as being available for the patient, responding as a human being, and listening and acknowledging the patient's problems were perceived to be most important in the treatment of depressed patients. These aspects of interaction were important in sustaining a trusting relationship and were part of standard treatment. No successful experiences to integrate process issues working with MoodGYM together with these supportive tasks were made explicit: "I guess it is [most important] to let the patient tell their story. It feels wrong to interrupt that type of story, about such things [when suffering from depression]. It is not so easy to cut to the chase—where in your stomach does it hurt?" [Participant 4] and "to strengthen the faith they might have in being able to become better. That there is hope. And…of course, to listen to them. To be willing to give more consultations and things like that" [Participant 2].

Confidence and viewing the patient as an active agent in treatment promoted GPs' efforts to remember, find the time, and motivate patients to log on and work independently with ICBT. These efforts to recommend MoodGYM may relate to the NPT component of collective action or efforts to make an intervention doable [27,30]. Again, components of work to implement the module follow-ups were not successful or prioritized, and to accommodate constraints GPs either chose not to mention the program or briefly ask patients about adherence before returning to standard treatment in further consultations.

Discussion

Principal Findings

In this paper, we address aspects of a gap between the evidence from multiple trials that have found that ICBT can reduce the symptoms of depression (eg, [8-11]) and the lack of knowledge about implementing guided ICBT in the everyday clinical setting in general practice from the perspective of the GPs themselves. In the present study, a storyline was investigated starting from what prompted GPs to learn about ICBT, to how they applied the guided ICBT in their treatment of depression. Our findings imply that guided ICBT in practice was perceived to consist of two steps: (1) a consultation recommending use of MoodGYM and subsequently (2) module follow-up consultations. NPT suggests four main components of "work" to implement interventions: coherence (make sense), cognitive participation (engagement), collective action (efforts), and reflexive monitoring (feedback) [27,30]. In our study, the interplay between wanting to improve treatment, engaging by adding structured agenda, and empowering the patient—being efforts to the first step of recommendation—reflects three of the NPT components: coherence, cognitive participation, and collective action. Overall, GPs' experiences with ICBT in our study demonstrated positive attitudes, as consistent with the literature

[5,21,22]. On the other hand, the second step of module follow-ups was inhibited by various aspects, such as a hectic and varied practice, inadequate knowledge of the content of modules, and competing tasks of standard treatment and thus did not generate engagement. Mohr suggests the rationale of "what, why, when and how" must be defined in depth to enable development of implementable technical interventions to change behavior [33], and we argue that a health worker must be able to some extent answer these questions to integrate interventions into treatment. What module follow-ups should comprise, why they would improve treatment, and how they should be deployed was not made explicit by the GPs in our study. This incomplete rationale may contribute to the insufficient coherence and cognitive participation and might explain the little effort or little collective action in this part of treatment. NPT here contributes by providing a structure through which to understand how complex the challenges are that affect the GPs' implementation of the module follow-ups.

A first step towards improvement was taken by GPs in our study because they made efforts to recommend MoodGYM; previous literature has found that GPs rarely recommend evidence-based self-management programs to patients with depressive symptoms [18,22]. An important aspect of promoting coherence and cognitive participation in MoodGYM was that it was based on CBT, a theory the GPs acknowledged as credible. Gunn and Palmer's study emphasized that improved treatments for depression that are theory-based are more likely to promote cognitive participation and therefore be implemented [34]. Earlier findings show that GPs do not recommend online interventions because they lack faith in, and knowledge of, the content of the interventions [22]. Another study showed that GPs who know of trusted sites incorporate the Internet into their role as a GP [35]. In the present study, GPs engaged with MoodGYM and could convey information about the specific theory MoodGYM is based on, especially those who had taken the 3-day course. This indicated that knowledge about evidence and the theoretical base of an online intervention may promote use by GPs.

Competence was a recurring aspect that affected the GPs' use of ICBT. Wanting to acquire competence promoted coherence, and experiencing competence gave the GPs' work satisfaction and indicated cognitive participation. The GPs' sense of competence was conveyed in encounters with patients, leading the GPs to make the effort to recommend the MoodGYM program, and it indicated collective action. The literature suggests that there is a relationship between the medical practitioner's competence in the treatment of depression and patient outcomes [36,37]. Patients who receive guided ICBT also emphasize the importance of therapist's competence [38]. The efforts or collective action GPs made to recommend MoodGYM to suitable patients in a motivating manner were often based on information they had learned at the course. Sinclair et al suggested redefining the role of the GPs to more of a consultant to enable proper use of ICBT during the treatment of depression [22]. Our findings suggest that with gained competence, GPs acted as consultants when they convincingly recommended MoodGYM and because they viewed their patients as active agents. This may indicate that it is consistent with the GPs' role to recommend Internet self-help and that a course can enable GPs to gain the necessary competence and confidence to recommend ICBT to their patients.

In the present study, the second step was problematic; that is, the GPs were not successful in providing module follow-ups following initial recommendation of MoodGYM to their patients. Two important aspects were inadequate module knowledge and time shortages. It is possible that the course did not sufficiently emphasize the follow-ups to build competence for this part of treatment. Confidence could be enhanced by more practical training and highlighting the rationale for the module follow-ups (eg, that guided ICBT is more effective than self-administered [8,9,13]). It is well documented that organizational constraints, such as lack of time and a varied practice, can be unsupportive to implementing eHealth in general [39] and CBT or ICBT in general practice [40-42]. In our opinion, time is also flexible, and constraints may be facilitated with incitements.

Paradoxically, although GPs expressed the inadequacy of standard treatment, they still preferred this approach in follow-ups. Despite their enthusiasm for the structure of guided ICBT, GPs found it difficult to encourage patients to talk about MoodGYM process issues or to use process issues as a trigger to talk about everyday life concerns. NPT claims that efforts to implement an intervention occur within interactions with others [27,30]. Previous studies have shown that clinicians [22] and patients [43] prefer online interventions to be used as adjuncts rather than alternatives to standard treatment. Perhaps the patients' expectations and preferences contributed to this choice in our study. An observational study from general practice found that when dealing with their patients' mental health issues, GPs intuitively chose to listen to their patients' problems contextualized in their everyday life, instead of being in control of the dialogue [44]. Both patients and GPs were more satisfied when the focus was on the patient as a whole person [34,44]. In our study, one GP (Participant 7) made explicit a role conflict she experienced between module follow-ups and her regular supporting role, indicating she saw the two types of care as dichotomous. Overall, our findings may imply that the focus on process issues working with modules is perceived as an instrumental approach, whereas when the depressed patients set the agenda by recounting their stories, a more patient-centered approach could be taken. This indicates a tension between the provisions of both structured therapeutic content of a program with a supportive patient-centered dialogue. Previous research from general practice has demonstrated this struggle between instrumental and patient-centered approaches [45], and the implementation of instrumental follow-ups does not fit with the GP's role in depression treatment [46]. If GPs in our study had better knowledge of the modules and the rationale for module follow-ups and could link back to MoodGYM instead of having it as a starting point of the dialogue, coherence, cognitive participation, and efforts could be strengthened to enable implementation of module follow-ups combined with a patient-centred approach.

Given that the current evidence is not clear, it is not a surprise that GPs in our study lacked knowledge about the composition of support in guided ICBT. A review has suggested that, when studying depression, the effect size in symptom reduction appears to be greater for minimal-contact follow-ups where the patient is provided with a rationale (but no support focusing on process issues) for the use of self-help materials, compared with guided follow-ups with a focus on process issues [12]. In our study, guided ICBT was incomplete. On the other hand, the GPs made an effort to recommend and provide a rationale for their patients' use of MoodGYM. The GPs also reported that they had offered follow-ups without process issues, during which they aimed to be available to their patients and to listen attentively. Such adjunct use to standard treatment may coincide with ICBT with minimal-contact follow-ups and perhaps is more compatible with the role of the GP.

Strengths and Limitations

A methodological strength of this study was the use of in-depth interviews. The GPs had experienced these change processes and were thus able to elaborate on their experiences implementing ICBT. This study explored not only their attitudes but also the work involved in embedding the approach into regular practice. The second author (RSH) is a psychologist and was a therapist in a randomized controlled trial exploring guided ICBT and was a presenter of the course for the GPs in the present study, and the first author works as a GP. Both interviewers had prior understanding of ICBT from clinical settings. We hoped this would improve our understanding of the field and of the terminology; however, we tried not to presume that the GPs included in this study shared our perceptions [47]. RSH's involvement in delivering the training might have led to a response bias if the GPs felt reluctant to be critical of the training or program. To reduce this possibility, we made it clear before each interview that we were not there to defend the treatment but that we wanted to better understand the GPs' experiences. Another strength of the study was that the experienced third author (MBR) gave continuous feedback about our interviews to ensure sufficiently high quality and depth.

Although the sample size was appropriate for the needs of the study, we interviewed only 11 GPs from northern Norway, and our findings should be interpreted as only a partial description of the full range of GPs' experiences. Participants were self-selected GPs enrolled in the course of the blended ICBT, and we cannot exclude the possibility of selection bias since participants may have been more interested in mental health or online interventions than the average GP. However, our main aim in this study was to explore aspects of the GPs' experiences when they made the effort to implement ICBT in everyday practice and, accordingly, motivated GPs were acceptable. There was an overrepresentation of women at the course and therefore reflects participants interviewed. More male participants might have produced other stories, and a balanced gender analysis would have been possible. Only one GP (female) declined an interview invitation. NPT encourages exploring all stakeholders' involvement [30], but time constraints and the necessary resources put this beyond the scope of this study.

Conclusions and Implications

Perceptions of GPs in the present study contribute to our understanding of the challenges associated with changing the treatment of depression in general practice. A need to supplement standard treatment was prominent, and GPs endorsed the principles of guided ICBT. Guided ICBT was seen as involving two steps. The first step was to recommend MoodGYM and efforts were made to integrate this step into treatment. This indicates that recommending ICBT can add a valuable tool to GP's toolkit. Offering training and highlighting the following aspects may increase recommendation of ICBT by GPs: (1) ICBT is theory-based and credible, (2) ICBT increases GPs' work satisfaction having a tool to offer, and (3) ICBT facilitates empowerment of patients in their own health.

The second step was to integrate module follow-ups into treatment. GPs expressed that they had difficulties with this step and instead returned to standard treatment. A number of reasons and paradoxes were identified when exploring this incomplete implementation. Our study indicates that recommending a theory-based Internet self-help program is acceptable within the role of a GP, however, unclear for module follow-ups. More practical training and providing incentives to enable GPs to prioritize time to complete the online program themselves to obtain knowledge may improve inadequate knowledge of the modules. Our findings imply that it is important to have a patient-centered approach in the follow-ups. More research is needed to explore what the support of ICBT should comprise when deployed in the context of general practice. A key question is to investigate if GPs can combine patient-centered follow-ups with process issues, or if adjunct use of ICBT to standard treatment is more suitable.

Acknowledgments

The authors gratefully acknowledge the help of the 11 participants. We confirm that all personal identifiers have been removed or disguised so the patient/person(s) described are not identifiable and cannot be identified through the details of the story.

This study was funded by Research Council of Norway (196423/V50).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[PDF File (Adobe PDF File), 33KB - jmir_v16i9e208_app1.pdf]

References

1. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]
2. Kringlen E, Torgersen S, Cramer V. A Norwegian psychiatric epidemiological study. Am J Psychiatry 2001 Jul;158(7):1091-1098. [Medline: 11431231]
3. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, CDBE2010Study Group. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Oct;21(10):718-779. [doi: 10.1016/j.euroneuro.2011.08.008] [Medline: 21924589]
4. Nutting PA, Rost K, Dickinson M, Werner JJ, Dickinson P, Smith JL, et al. Barriers to initiating depression treatment in primary care practice. J Gen Intern Med 2002 Feb;17(2):103-111 [FREE Full text] [Medline: 11841525]
5. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264. [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]
6. Proudfoot J, Goldberg D, Mann A, Everitt B, Marks I, Gray JA. Computerized, interactive, multimedia cognitive-behavioural program for anxiety and depression in general practice. Psychol Med 2003 Feb;33(2):217-227. [Medline: 12622301]
7. Hans E, Hiller W. Effectiveness of and dropout from outpatient cognitive behavioral therapy for adult unipolar depression: a meta-analysis of nonrandomized effectiveness studies. J Consult Clin Psychol 2013 Feb;81(1):75-88. [doi: 10.1037/a0031080] [Medline: 23379264]
8. 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]
9. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-9; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
10. Høifødt R, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, et al. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. J Med Internet Res 2013;15(8):e153 [FREE Full text] [doi: 10.2196/jmir.2714] [Medline: 23916965]
11. Spek V, Cuijpers P, Nyklícek I, Smits N, Riper H, Keyzer J, et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. Psychol Med 2008 May;38(5):635-639. [doi: 10.1017/S0033291707002590] [Medline: 18205965]
12. Farrand P, Woodford J. Impact of support on the effectiveness of written cognitive behavioural self-help: a systematic review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2013 Feb;33(1):182-195. [doi: 10.1016/j.cpr.2012.11.001] [Medline: 23238023]
13. 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]
14. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
15. Lillevoll KR, Wilhelmsen M, Kolstrup N, Høifødt R, Waterloo K, Eisemann M, et al. Patients' experiences of helpfulness in guided internet-based treatment for depression: qualitative study of integrated therapeutic dimensions. J Med Internet Res 2013;15(6):e126 [FREE Full text] [doi: 10.2196/jmir.2531] [Medline: 23786763]
16. Wangberg S, Andreassen H, Kummervold P, Wynn R, Sørensen T. Use of the internet for health purposes: trends in Norway 2000-2010. Scand J Caring Sci 2009 Dec;23(4):691-696. [doi: 10.1111/j.1471-6712.2008.00662.x] [Medline: 19473317]
17. Kessler RC, Merikangas KR, Berglund P, Eaton WW, Koretz DS, Walters EE. Mild disorders should not be eliminated from the DSM-V. Arch Gen Psychiatry 2003;60(11):1117. [doi: 10.1001/archpsyc.60.11.1117]
18. Hermens ML, Muntingh A, Franx G, van Splunteren PT, Nuyen J. Stepped care for depression is easy to recommend, but harder to implement: results of an explorative study within primary care in the Netherlands. BMC Fam Pract 2014;15:5 [FREE Full text] [doi: 10.1186/1471-2296-15-5] [Medline: 24400701]
19. Helsedirektoratet. Report number IS-1561.: Helsedirektoratet (The Norwegian Directorate of Health); 2009. Nasjonale retningslinjer for diagnostisering og behandling av voksne med depresjon i primær- og spesialisthelsetjenesten, (National guidelines for diagnosing and treatment of adults with depression in primary and specialized care) URL: http://www.helsedirektoratet.no/publikasjoner/nasjonale-retningslinjer-for-diagnostisering-og-behandling-av-voksne-med-depresjon-i-primer--og-spesialisthelsetjenesten/Sider/default.aspx [accessed 2014-09-05] [WebCite Cache ID 6SMZCNmMU]
20. Pilling S, Anderson I, Goldberg D, Meader N, Taylor C. Depression in adults, including those with a chronic physical health problem: summary of NICE guidance. BMJ 2009 Oct 27;339(oct27 1):b4108-b4108. [doi: 10.1136/Bmj.B4108]

21. 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]
22. Sinclair C, Holloway K, Riley G, Auret K. Online mental health resources in rural Australia: clinician perceptions of acceptability. J Med Internet Res 2013;15(9):e193 [FREE Full text] [doi: 10.2196/jmir.2772] [Medline: 24007949]
23. Høifødt R, Strøm C, Kolstrup N, Eisemann M, Waterloo K. Effectiveness of cognitive behavioural therapy in primary health care: a review. Fam Pract 2011 Oct;28(5):489-504 [FREE Full text] [doi: 10.1093/fampra/cmr017] [Medline: 21555339]
24. Proudfoot J, Ryden C, Everitt B, Shapiro DA, Goldberg D, Mann A, et al. Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. Br J Psychiatry 2004 Jul;185:46-54 [FREE Full text] [Medline: 15231555]
25. Williams AD, Andrews G. The effectiveness of Internet cognitive behavioural therapy (iCBT) for depression in primary care: a quality assurance study. PLoS One 2013;8(2):e57447 [FREE Full text] [doi: 10.1371/journal.pone.0057447] [Medline: 23451231]
26. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. J Med Internet Res 2002;4(1):e3 [FREE Full text] [doi: 10.2196/jmir.4.1.e3] [Medline: 11956035]
27. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010;8:63 [FREE Full text] [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
28. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012;1(2):e3 [FREE Full text] [Medline: 25075231]
29. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. Med J Aust 2004 Mar 15;180(6 Suppl):S57-S60. [Medline: 15012583]
30. May C, Finch T. Implementing, Embedding, and Integrating Practices: An Outline of Normalization Process Theory. Sociology 2009 Jun 15;43(3):535-554. [doi: 10.1177/0038038509103208]
31. Kvale S. Det kvalitative forskningsintervju 2. In: utgave Sage. Oslo, Norway: Gyldendal Norsk Forlag AS; 2009.
32. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
34. Gunn JM, Palmer VJ, Dowrick CF, Herrman HE, Griffiths FE, Kokanovic R, et al. Embedding effective depression care: using theory for primary care organisational and systems change. Implement Sci 2010;5:62 [FREE Full text] [doi: 10.1186/1748-5908-5-62] [Medline: 20687962]
35. Ahluwalia S, Murray E, Stevenson F, Kerr C, Burns J. 'A heartbeat moment': qualitative study of GP views of patients bringing health information from the internet to a consultation. Br J Gen Pract 2010 Feb;60(571):88-94 [FREE Full text] [doi: 10.3399/bjgp10X483120] [Medline: 20132702]
36. Muse K, McManus F. A systematic review of methods for assessing competence in cognitive-behavioural therapy. Clin Psychol Rev 2013 Apr;33(3):484-499. [doi: 10.1016/j.cpr.2013.01.010] [Medline: 23454222]
37. Salmon P, Peters S, Clifford R, Iredale W, Gask L, Rogers A, et al. Why do general practitioners decline training to improve management of medically unexplained symptoms? J Gen Intern Med 2007 May;22(5):565-571 [FREE Full text] [doi: 10.1007/s11606-006-0094-z] [Medline: 17443362]
38. Wilhelmsen M, Lillevoll K, Risør M, Høifødt R, Johansen ML, Waterloo K, et al. Motivation to persist with internet-based cognitive behavioural treatment using blended care: a qualitative study. BMC Psychiatry 2013;13:296 [FREE Full text] [doi: 10.1186/1471-244X-13-296] [Medline: 24199672]
39. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
40. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. BMJ 2002 Apr 20;324(7343):947-950 [FREE Full text] [Medline: 11964340]
41. Wiebe E, Greiver M. Using cognitive behavioural therapy in practice: qualitative study of family physicians' experiences. Can Fam Physician 2005 Jul;51:992-993 [FREE Full text] [Medline: 16926952]
42. Aschim B, Lundevall S, Martinsen EW, Frich JC. General practitioners' experiences using cognitive behavioural therapy in general practice: a qualitative study. Scand J Prim Health Care 2011 Sep;29(3):176-180 [FREE Full text] [doi: 10.3109/02813432.2011.595582] [Medline: 21861599]

43. Schneider J, Sarrami Foroushani P, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]
44. Barry CA, Stevenson FA, Britten N, Barber N, Bradley CP. Giving voice to the lifeworld. More humane, more effective medical care? A qualitative study of doctor-patient communication in general practice. Soc Sci Med 2001 Aug;53(4):487-505. [Medline: 11459399]
45. Reeve J, Dowrick CF, Freeman GK, Gunn J, Mair F, May C, et al. Examining the practice of generalist expertise: a qualitative study identifying constraints and solutions. JRSM Short Rep 2013 Dec;4(12):2042533313510155 [FREE Full text] [doi: 10.1177/2042533313510155] [Medline: 24475347]
46. Franx G, Oud M, de Lange J, Wensing M, Grol R. Implementing a stepped-care approach in primary care: results of a qualitative study. Implement Sci 2012;7:8 [FREE Full text] [doi: 10.1186/1748-5908-7-8] [Medline: 22293362]
47. Maxwell JA. Qualitative Research Design: An Interactive Approach (Applied Social Research Methods). Thousand Oaks, CA: Sage Publications, Inc; 2013.

Abbreviations

CBT: cognitive behavioral therapy
GP: general practitioner
ICBT: Internet-based cognitive behavioral therapy
NPT: normalization process theory

Edited by G Eysenbach; submitted 23.05.14; peer-reviewed by C Sinclair, L Donkin, T Donker; comments to author 18.06.14; revised version received 11.08.14; accepted 28.08.14; published 10.09.14.

Please cite as:
Wilhelmsen M, Høifødt RS, Kolstrup N, Waterloo K, Eisemann M, Chenhall R, Risør MB
Norwegian General Practitioners' Perspectives on Implementation of a Guided Web-Based Cognitive Behavioral Therapy for Depression: A Qualitative Study
J Med Internet Res 2014;16(9):e208
URL: http://www.jmir.org/2014/9/e208/
doi:10.2196/jmir.3556
PMID:25208886

Original Paper

A Web-Based Program Improves Physical Activity Outcomes in a Primary Care Angina Population: Randomized Controlled Trial

Reena Devi[1], BSc, MSc, PhD; John Powell[2], MBBChir, MSc, PhD, FFPH; Sally Singh[1,3], FCFP, PhD

[1]Coventry University, Applied Research Centre in Health and Lifestyle Interventions, Coventry, United Kingdom

[2]University of Oxford, Nuffield Department of Primary Care Health Sciences, Oxford, United Kingdom

[3]Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM), University Hospitals of Leicester, Glenfield Hospital, Centre for Exercise and Rehabilitation Science, Leicester, United Kingdom

Corresponding Author:
Reena Devi, BSc, MSc, PhD
Coventry University
Applied Research Centre in Health and Lifestyle Interventions
Priory Street
Coventry, CV1 5FB
United Kingdom
Phone: 44 24 7688 7688
Fax: 44 24 7765 5987
Email: reena.devi@coventry.ac.uk

Abstract

Background: Angina affects more than 50 million people worldwide. Secondary prevention interventions such as cardiac rehabilitation are not widely available for this population. An Internet-based version could offer a feasible alternative.

Objective: Our aim was to examine the effectiveness of a Web-based cardiac rehabilitation program for those with angina.

Methods: We conducted a randomized controlled trial, recruiting those diagnosed with angina from general practitioners (GPs) in primary care to an intervention or control group. Intervention group participants were offered a 6-week Web-based rehabilitation program ("ActivateYourHeart"). The program was introduced during a face-to-face appointment and then delivered via the Internet (no further face-to-face contact). The program contained information about the secondary prevention of coronary heart disease (CHD) and set each user goals around physical activity, diet, managing emotions, and smoking. Performance against goals was reviewed throughout the program and goals were then reset/modified. Participants completed an online exercise diary and communicated with rehabilitation specialists through an email link/synchronized chat room. Participants in the control group continued with GP treatment as usual, which consisted of being placed on a CHD register and attending an annual review. Outcomes were measured at 6-week and 6-month follow-ups during face-to-face assessments. The primary outcome measure was change in daily steps at 6 weeks, measured using an accelerometer. Secondary outcome measures were energy expenditure (EE), duration of sedentary activity (DSA), duration of moderate activity (DMA), weight, diastolic/systolic blood pressure, and body fat percentage. Self-assessed questionnaire outcomes included fat/fiber intake, anxiety/depression, self-efficacy, and quality of life (QOL).

Results: A total of 94 participants were recruited and randomized to the intervention (n=48) or the usual care (n=46) group; 84 and 73 participants completed the 6-week and 6-month follow-ups, respectively. The mean number of log-ins to the program was 18.68 (SD 13.13, range 1-51), an average of 3 log-ins per week per participant. Change in daily steps walked at the 6-week follow-up was +497 (SD 2171) in the intervention group and –861 (SD 2534) in the control group (95% CI 263-2451, *P*=.02). Significant intervention effects were observed at the 6-week follow-up in EE (+43.94 kcal, 95% CI 43.93-309.98, *P*=.01), DSA (–7.79 minutes, 95% CI –55.01 to –7.01, *P*=.01), DMA (+6.31 minutes, 95% CI 6.01-51.20, *P*=.01), weight (–0.56 kg, 95% CI –1.78 to –0.15, *P*=.02), self-efficacy (95% CI 0.30-4.79, *P*=.03), emotional QOL score (95% CI 0.01-0.54, *P*=.04), and angina frequency (95% CI 8.57-35.05, *P*=.002). Significant benefits in angina frequency (95% CI 1.89-29.41, *P*=.02) and social QOL score (95% CI 0.05-0.54, *P*=.02) were also observed at the 6-month follow-up.

Conclusions: An Internet-based secondary prevention intervention could be offered to those with angina. A larger pragmatic trial is required to provide definitive evidence of effectiveness and cost-effectiveness.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 90110503; http://www.controlled-trials.com/ISRCTN90110503/ISRCTN90110503 (Archived by WebCite at http://www.webcitation.org/6RYVOQFKM).

(J Med Internet Res 2014;16(9):e186) doi:10.2196/jmir.3340

KEYWORDS

stable angina; cardiac rehabilitation; Web-based interventions; secondary prevention; primary health care; physical activity

Introduction

The impact of angina is significant to both the individual [1] and to the health service [2]. Cardiac rehabilitation is recommended for individuals with angina in many international guidelines [3], but capacity to accommodate these individuals is limited and those with a recent cardiac event take priority. Recent data suggest that angina patients constitute only 4% of the referrals to rehabilitation, and almost 20% of programs do not accept those with angina [4].

There is a broad spectrum of interventions that may constitute a rehabilitation program, from fully supervised sessions to more remote home-based services. It is largely these self-directed home-based programs that have been tested in the angina population. A meta-analysis of 7 trials [5] demonstrated that psychoeducational interventions delivered via a trained professional significantly reduced medication use, physical limitations, and disease perception in angina populations. In addition, the angina population have been considered previously with a manual-based approach, The Angina Plan [6,7]; however, this has not been widely adopted [4]. A small number of trials have studied the effectiveness of secondary prevention interventions for coronary heart disease (CHD) delivered via the Internet. A recent Canadian study [8] evaluated a 6-month Web-based physical activity program for patients who had undergone percutaneous coronary revascularization. The study did not report baseline scores, but the authors reported higher levels of physical activity in the intervention group compared to the control group. The change in physical activity was reported from the 6- to 12-month follow-ups in the intervention group and this was significant compared to the control group. Recently a study conducted in Norway assessed the effectiveness of an Internet- and mobile phone-based intervention for physical activity as an extension of face-to-face cardiac rehabilitation [9]. The study demonstrated significantly higher physical activity levels in the intervention group compared to a control group at 3-month follow-up. However, the study is somewhat limited by the small sample size at follow-up (n=7) and the self-reported measure of physical activity. The value of Web-based interventions and physical activity promotion has also been investigated by Van den Berg et al in a systematic review [10]. Van den Berg et al reviewed 10 articles and reported online interventions are effective in improving physical activity levels [10]. This review emphasized the need to measure physical activity using objective measures. Positive findings have been reported from research measuring physical activity objectively when evaluating Web-based physical activity interventions [11,12].

We developed an interactive password-protected website specifically for individuals with CHD, which is a comprehensive educational package that aims to improve health behaviors related to CHD. The Internet allows for the delivery of a standard intervention that is not geographically or time restrained. It is intended that this intervention could be offered to those not routinely included within traditional cardiac rehabilitation, such as those with stable angina. The purpose of this study was to assess the clinical effectiveness of this independent Internet-delivered self-managed “rehabilitation” program in a population with chronic stable angina in a primary care setting. Because we were studying the efficacy of a novel intervention for which we had only limited previous data, our primary hypothesis was nondirectional and was “users of a Web-based cardiac rehabilitation program would alter their coronary risk factors compared to those receiving treatment as usual (control group).”

Methods

Study Design and Randomization

A randomized controlled trial with 2 parallel group arms was conducted (ISRCTN 90110503). The 2 groups consisted of the intervention group and the treatment-as-usual control group. A computerized block randomization list was produced by our departmental statistician. Allocation concealment was achieved by sequentially numbered sealed envelopes, opened after baseline data collection for each participant by the researcher carrying out the fieldwork (RD). Participants and the outcome assessor were not blinded to group allocation.

Recruitment and Participants

Participants were recruited offline from 9 primary care general practitioners (GPs) in 1 region of England. Participants were selected from CHD registers by a GP or practice nurse. Individuals were invited to participate if they had a confirmed diagnosis of stable angina, were able to read and speak fluent English, had regular access to the Internet, were computer literate, and had not had conventional cardiac rehabilitation within the previous year. Individuals were excluded if they had unstable angina, significant cardiac arrhythmia, any comorbidities preventing physical activity, or were severely anxious/ depressed. Severely anxious/depressed patients were excluded by eliminating anyone with a history of being prescribed medication for either anxiety or depression. Participants were not banned from attending conventional rehabilitation; however, if a participant was offered rehabilitation or any other secondary prevention intervention during the course of the study, they were excluded because this was considered a breach of study design. At each study

follow-up, this was asked and recorded in the study notes. Participant recruitment and outcome follow-ups were carried out from September 2008 to February 2010.

Outcome Measures and Data Collection

Both primary and secondary outcome measures were collected at baseline, 6 weeks after randomization, and then 6 months after the 6-week follow-up by a researcher (RD) visiting the participants at home. Participant follow-up continued until October 2010.

Primary Outcome Measure

The primary outcome measure was daily average step count change at 6-week follow-up. This was measured using Sensewear Pro 3 accelerometer technology, a nondisplay multisensor monitor. This monitor uses physiological signals, bodily movement, and in-built algorithms to estimate physical activity. Participants wore the monitor on the right upper arm for 2 weekdays (12 hours per day) at baseline and at the 6-week and 6-month follow-ups. Reliability and accuracy of this technology has been established in healthy individuals [13] and unhealthy individuals [14,15].

Secondary Outcome Measures

Secondary outcome measures included energy expenditure (EE), duration of sedentary activity (DSA), and duration of moderate activity (DMA); these were measured using the same accelerometer that measured the primary outcome. Participants wore the accelerometer on the right upper arm for 2 weekdays (12 hours per day) at baseline and at the 6-week and 6-month follow-ups. Weight, diastolic (DBP) and systolic blood pressure (SBP), and body fat percentage were measured using conventional instruments. Other outcomes were fat and fiber intake, anxiety and depression, self-efficacy, and health-related quality of life (QOL). Fat and fiber intake was measured using the Dietary Instrument for Nutritional Evaluation [16], which is a validated measure to assess fat and fiber intake [16]. This measure contained 19 groups of foods representing fat and fiber in a typical UK diet, and involved participants choosing the frequency of food groups consumed from multiple-choice answers. Anxiety and depression was assessed using the Hospital Anxiety and Depression Scale (HADS) [17], which is a 14-item validated measure of anxiety and depression [18]. This measure is also a reliable and valid instrument for use in a stable coronary population [19]. Self-efficacy was measured using The General Self-Efficacy Scale, a reliable and valid measure of self-efficacy [20], which is comprised of 10 items scored on a 4-point scale. The developers acknowledge it is a general scale and, therefore, suggest additional specific items can be added [21]. In this study, self-efficacy of exercise (3 items), knowledge of heart disease (1 item), and eating a healthy diet (1 item) were added as extra items to the scale. The final score of all items was used to describe the overall self-efficacy of participants; higher scores reflected greater self-efficacy. Health-related QOL was assessed using The MacNew questionnaire [22] and The Seattle Angina Questionnaire (SAQ) [23], of which both are CHD specific. The MacNew questionnaire consists of 27 items measuring perceived quality of emotional, physical, and social health. Each item was scored on a 7-point scale with lower scores corresponding to impaired QOL. This has been reported to be a valid and reliable measure, sensitive to changes in health-related QOL [22], and reliable/valid for use in angina patients [24]. The SAQ questionnaire comprises 19 questions that constitute 5 subscales: physical limitations, angina stability, angina frequency, treatment satisfaction, and disease perception. Lower scores indicate poorer health status and higher scores indicate better health. This measure has undergone validity and reliability testing [23]. In the intervention group, we also monitored the number of log-ins to the online program. This information was available from the administration side of the intervention.

Procedure

Eligible individuals were sent a postal invitation and those who replied with an interest in participating were contacted. Prospective participants were telephoned by the study researcher to check trial suitability and to arrange an initial home visit. The initial home visit was arranged at a time most convenient for the participant. During the home visit, the researcher (RD) explained that the purpose of the study was to investigate the effectiveness of a Web-based intervention, described study details, took participant consent, and carried out the physical baseline outcome measures (weight, blood pressure, and body fat percentage). Because the initial home visit was arranged at a time most convenient for the participant, it was not possible to control for factors such as time of day, whether the participant felt rested, nor whether the participant was alone or not during the time of measurement. During this initial meeting, participants were also given an accelerometer and a questionnaire pack. Each participant was instructed to wear the monitor for 2 weekdays (12 hours per day) and to complete the questionnaires (paper-based questionnaires). This initial meeting lasted approximately 40 minutes. After all baseline measures were collected, the researcher (RD) randomized each participant, telling each participant which group they had been allocated to. Those in the Web-based cardiac rehabilitation group received a face-to-face introductory session from the researcher (RD). This involved registering the individual, creating a unique username/password, and demonstrating how to use the program. Intervention group participants were told to log in to the program daily to record their daily physical activity. The control group did not receive any intervention and continued care as usual. Study outcome measures were repeated at the 6-week and 6-month follow-ups. Participants were not paid to take part in this trial.

Intervention

The intervention was delivered via the Internet and called “ActivateYourHeart” [25], a secure and password-protected site designed for participants to use at home. The program was developed at the University Hospitals of Leicester NHS Trust and coproduced with health care professionals, a software development team (HARK2), and a group of patients/members of the public. Development of the site was an interactive and iterative process, involving patients providing input and feedback on different versions of the website, including feedback on website content, layout, visual features, and ease of website navigation. The program aimed to improve patients’ cardiac

risk profile within 4 stages and was designed to be completed within 6 weeks. The intervention used the following behavior change techniques [26]: setting/reviewing behavioral goals, self-monitoring, feedback on behavior, graded tasks, social reward, providing information about health consequences, and reducing negative emotions.

At the beginning of the program, each user completed an online form providing information about their medical history and their current cardiac risk factors (Multimedia Appendices 1 and 2). This information was used to set individualized tailored goals focused on exercise (eg, being physically active for 30 minutes 5 times a week), diet (eg, eating more fruit/vegetables and reducing salt intake), emotions (eg, managing stress and other negative emotions), and smoking (eg, reduce cigarette smoking if relevant) (Figure 1).

Compliance to these goals was regularly assessed (using a short set of questions) and feedback on performance provided (Multimedia Appendix 3). Users making progress were congratulated when set goals were achieved. Throughout the program, goals were reset/modified depending on previous performance. As the user progressed through the program, goals set were made increasingly difficult.

Each user also kept an online exercise diary, recording details of their daily exercise (Multimedia Appendix 4). Feedback on the users' physical activity levels was also provided as they progressed through the program. Users who smoked cigarettes were provided with feedback regarding the amount of money they had spent/saved by smoking/reducing smoking. The program also contained written information about the health consequences of heart disease and a vast amount of information about CHD-related risk factors (exercise, diet, sexual activity, driving, returning to work, hobbies, holidays, benefits, smoking, anxiety, and emotions).

In addition, the programme aimed to reduce negative emotions by providing advice about stress/anxiety management skills (see Multimedia Appendix 5). The program also contained information to help users understand heart disease (Figure 2). Program users could initiate contact with cardiac rehabilitation nurses for advice and support via an online email link (see Multimedia Appendix 6) or by joining a scheduled synchronized chat room held on a weekly basis. The cardiac nurses were based at University Hospitals of Leicester. All participants in the intervention group used the program from home and were encouraged to log in to the program 3-4 times per week.

Figure 1. Program goal setting.

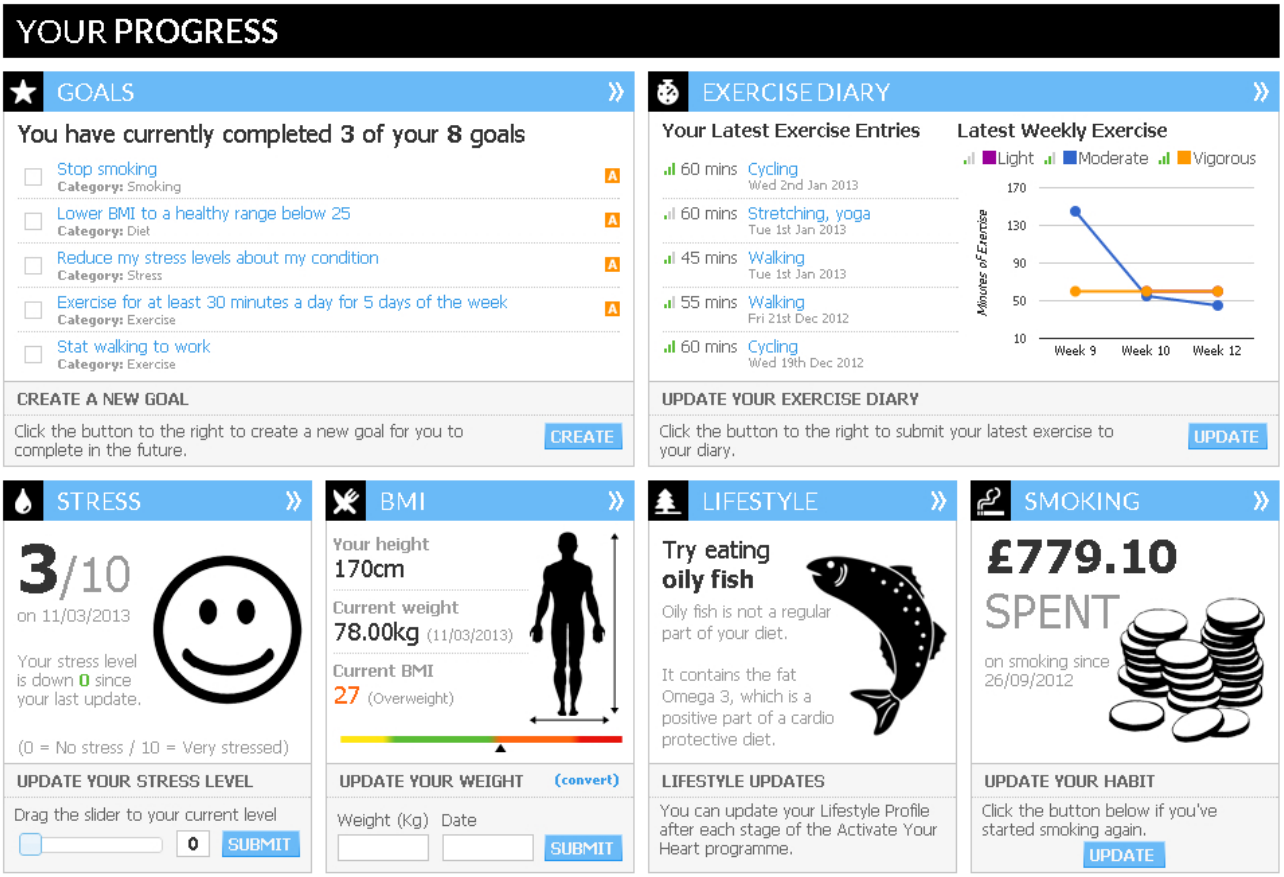
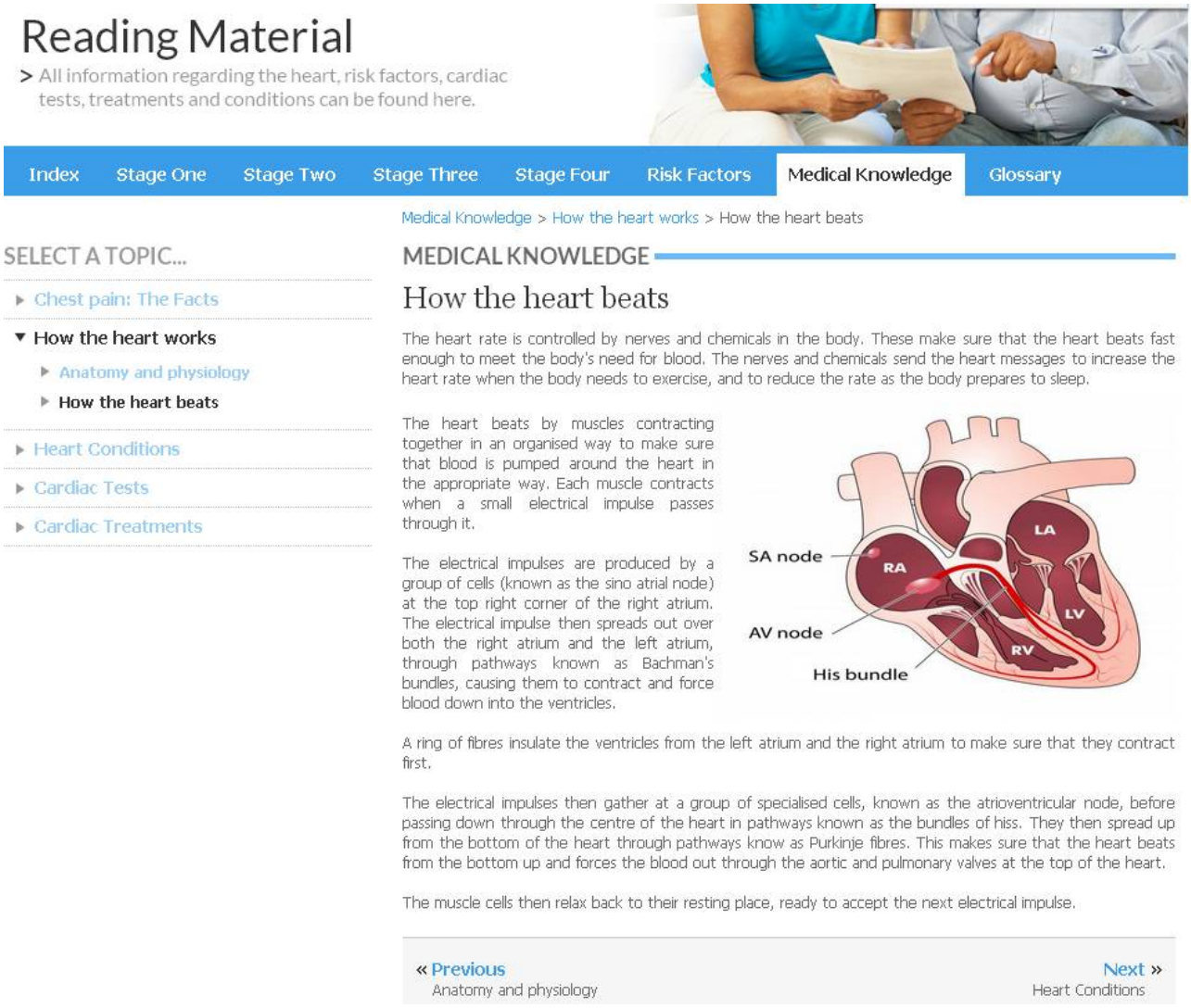


Figure 2. Information about heart disease contained in the program.



Control

Participants in the control group continued with treatment as usual from their GP and received no further contact from the researcher until the 6-week follow-up. Usual care in primary care for this population in the United Kingdom constitutes being placed on a CHD register and attending an annual check of risk factor management, usually with a practice nurse.

Sample Size Calculation

Sample size was based on detecting a significant change in the number of steps walked by participants at the 6-week follow-up. Using previous data, our sample size calculation was based on detecting a difference in means of 3501 steps walked between the intervention and control group [27]. This would require 24 (total 48) participants in each group (with 90% power and .05 significance). We recruited more than this (N=94, 96% more participants) to allow for dropout (often high in studies of Web-based interventions) and to allow for the detection of differences between secondary measures.

Statistical Methods

Demographic characteristics and baseline measures were compared at baseline using Pearson chi-square tests (categorical variables), independent samples *t* tests (continuous, normally distributed data), and Mann-Whitney *U* tests (nonnormally distributed data). Fisher exact test was used when chi-square test assumptions were violated. Baseline outcome measures in trial completers and trial dropouts were also compared. Change from baseline to follow-up time points in both primary and secondary outcome variables were calculated (follow-up score or value – baseline score or value). The change values in each group were then compared using an independent sample *t* test (normally distributed data) or Mann-Whitney *U* test (nonnormally distributed data). We chose to examine the change in primary/secondary outcome measures at 6-week and 6-month follow-ups and compare this value between groups. This approach to the analysis ensured that all participants' available data could be used irrespective of study completion level.

All statistical analyses were carried out using SPSS version 22 (IBM Corp, Armonk, NY, USA). Data were analyzed using intention-to-treat analyses; all participants with data available were included in the data analysis according to the group first assigned at randomization regardless of intervention compliance or adherence. Attrition was low; therefore, we did not use any imputation techniques to deal with attrition. Two-tailed findings were reported.

Ethics

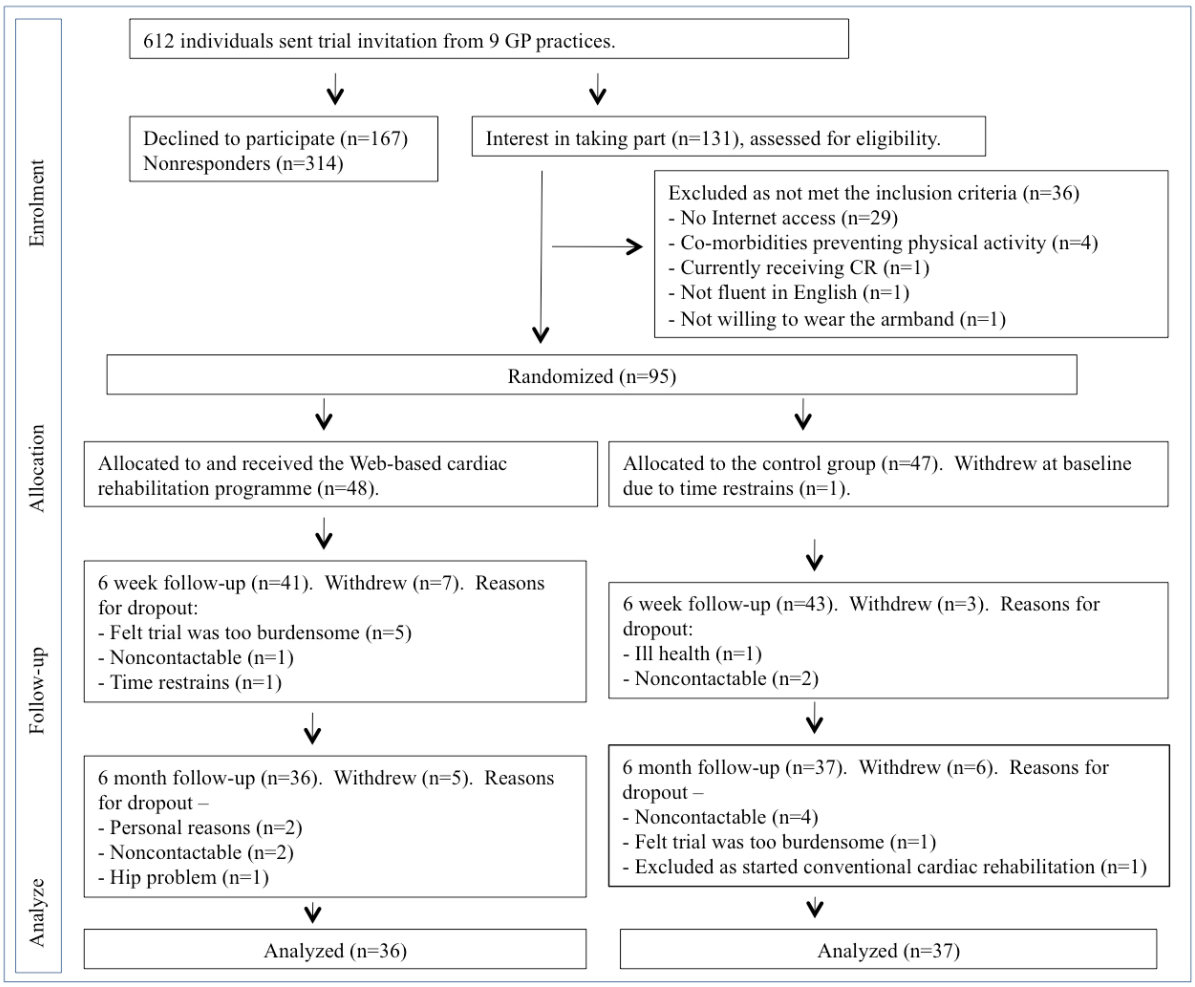
The study protocol gained ethical approval granted by the National Health Service Research Ethics Service (ref: 08/H1210/84) and by Coventry University.

Results

Participation Rates

A total of 612 patients were invited to take part; 481 (78.6%) declined/did not respond and 131 (21.4%) expressed an interest in the study, of which 95 (15.5%) consented to the trial (Figure 3). A total of 94 participants (99%) completed the baseline measures, 84 (89%) completed the 6-week follow-up (11% attrition), and 73 (78%) completed the 6-month follow-up (22% attrition). At baseline, SBP was higher in those who dropped out (mean 145.19 mm Hg, SD 12.53) compared to those who completed the study (mean 132.95 mm Hg, SD 16.28; *P*=.002). There were no other statistically significant differences between trial completers and trial dropouts in demographic characteristics or baseline outcome measures. Participant flow throughout the trial is shown in Figure 3.

Figure 3. Participant flow through the trial.



Demographic Characteristics and Baseline Measures

Participant demographic details are outlined in Table 1. There were no significant differences between the intervention and control group in demographic characteristics or baseline measures.

Short-Term Intervention Effects

Table 2 outlines baseline and 6-week follow-up values, and change values for all outcomes.

Table 1. Demographic characteristics of participants.

Demographic characteristic	Intervention group (n=48)	Control group (n=46)
Age (years), mean (SD)	66.27 (8.35)	66.20 (10.06)
Gender, n (%)		
Male	34 (71)	36 (78)
Female	14 (29)	10 (22)
Employment, n (%)		
Retired	29 (60)	21 (46)
Full-time	13 (27)	18 (39)
Part-time	4 (8)	7 (15)
Unemployed	2 (4)	0 (0)
Ethnicity, n (%)		
White British	44 (92)	42 (91)
Other	4 (8)	4 (9)
Years since diagnosis, mean (SD)	7.98 (4.53)	9.44 (5.81)
Angina treatment, n (%)		
Medication only	19 (44)	16 (37)
Stent(s)	15 (35)	21 (49)
Coronary artery bypass graft	9 (21)	6 (14)
Previous cardiac rehabilitation, n (%)		
No	34 (76)	35 (81)
Yes	11 (24)	8 (19)
Current smoking status, n (%)		
No	46 (96)	40 (87)
Yes	2 (4)	6 (13)

Table 2. Short-term intervention effects at baseline (T0) and 6-week follow-up (T1) for the intervention and control groups, including within- and between-group differences (D).

Outcome[a]	Intervention group, mean (SD)				Control group, mean (SD)							
	n[b]	T0	T1	D	n[c]	T0	T1	D	D	*P* [d]	ES	95% CI
Physical activity												
Daily steps	35	6716 (3060)	7212 (3188)	+497 (2171)	40	6624 (3189)	5763 (2533)	–861 (2534)	1357	.02	0.58	263, 2451
Daily EE (kcal)	35	1902.47 (392.32)	1946.41 (351.79)	+43.94 (271.90)	40	2055.05 (431.80)	1922.04 (306.47)	–133.01 (302.01)	176.96	.01	0.62	43.93, 309.98
DSA (min)[f]	35	675.00 (45.00)	671.50 (55.50)	–7.79[e] (40.14)	40	663.25 (103.25)	672.25 (61.75)	+23.23[e] (62.78)	–31.01	.01	0.59	–55.01, –7.01
DMA (min)[f]	35	43.50 (43.00)	48.50 (50.00)	+6.31[e] (34.37)	40	55.50 (96.25)	47.75 (61.38)	–22.29[e] (61.34)	28.60	.01	0.58	6.01, 51.20
Physiological measures												
Weight (kg)	41	82.80 (13.49)	82.24 (13.30)	–0.56 (2.00)	42	79.52 (14.36)	79.93 (14.74)	+0.40 (1.71)	–0.97	.02	0.52	–1.78, –0.15
Body fat (%)	39	38.78 (10.80)	38.36 (11.52)	–0.42 (7.67)	41	36.34 (8.01)	37.01 (7.07)	+0.68 (6.39)	–1.09	.49	0.16	–4.23, 2.04
SBP (mm Hg)	40	131.35 (15.34)	130.80 (14.70)	–0.55 (12.03)	42	137.55 (16.51)	128.55 (14.88)	–9.00 (12.77)	8.45	.003	0.68	2.99, 13.91
DBP (mm Hg)	39	72.92 (9.95)	69.00 (9.57)	–3.92 (8.75)	42	72.52 (10.73)	68.52 (9.16)	–4.00 (8.27)	0.08	.97	0.01	–3.69, 3.84
Diet												
Fat score	33	38.76 (8.46)	35.55 (9.18)	–3.21 (7.98)	32	40.88 (11.63)	39.38 (10.38)	–1.50 (11.89)	–1.71	.50	0.17	–6.72, 3.29
Fiber score	35	36.40 (9.84)	36.51 (8.77)	+0.11 (6.88)	33	35.09 (12.46)	33.79 (12.24)	–1.30 (12.14)	1.42	.55	0.14	–3.33, 6.16
Psychological												
Anxiety score	36	5.61 (3.57)	4.14 (3.50)	–1.47 (3.19)	39	5.51 (3.42)	4.87 (3.73)	–0.64 (2.27)	–0.83	.20	0.30	–2.10, 0.44
Depression scores[f]	37	3.00 (4.00)	2.00 (2.00)	–0.43[e] (2.15)	42	2.00 (3.00)	2.00 (4.25)	+0.10[e] (2.30)	–0.53	.30	0.24	–1.53, 0.48
Self-efficacy score	37	49.03 (6.55)	51.70 (6.37)	+2.68 (5.92)	39	49.79 (7.56)	49.92 (7.76)	+0.13 (3.49)	2.55	.03	0.52	0.30, 4.79
MacNew QOL												
Emotional score[f]	36	5.89 (1.21)	6.25 (1.04)	+0.31[e] (0.67)	40	5.96 (1.45)	6.32 (1.21)	+0.04[e] (0.44)	0.27	.04	0.48	0.01, 0.54
Physical score[f]	33	6.50 (0.71)	6.50 (0.92)	+0.04[e] (0.69)	41	6.50 (1.42)	6.58 (1.33)	+0.11[e] (0.57)	–0.07	.62	0.11	–0.37, 0.22
Social score[f]	34	6.54 (0.85)	6.73 (0.50)	+0.21[e] (0.66)	40	6.54 (1.17)	6.62 (1.19)	+0.07[e] (0.57)	0.14	.34	0.23	–0.15, 0.42
SAQ [g]												
Physical limitations score	37	64.19 (21.55)	62.16 (25.43)	–2.03 (19.20)	42	63.49 (25.40)	63.69 (27.03)	+0.20 (15.19)	–2.23	.57	0.13	–9.94, 5.49
Angina stability score[f]	33	42.86 (57.14)	33.33 (66.67)	–9.74[e] (39.81)	37	42.86 (57.14)	33.33 (66.67)	–9.97[e] (33.63)	0.23	.98	0.01	–17.29, 17.53
Angina frequency score	33	43.56 (31.58)	53.79 (30.70)	+10.23 (26.78)	41	44.51 (32.36)	32.93 (28.74)	–11.59 (29.63)	21.81	.002	0.77	8.57, 35.05
Treatment satisfaction score[f]	35	100.00 (0.00)	100.00 (0.00)	+4.04[e] (23.38)	36	100.00 (28.57)	100.00 (22.22)	–1.90[e] (30.52)	5.93	.36	0.22	–6.97, 18.83

Outcome[a]	Intervention group, mean (SD)				Control group, mean (SD)							
	n[b]	T0	T1	D	n[c]	T0	T1	D	D	*P* [d]	ES	95% CI
Disease perception score[f]	36	83.33 (33.33)	80.00 (40.00)	+0.97[e] (20.15)	40	83.33 (39.58)	80.00 (40.00)	–2.13[e] (17.54)	3.10	.48	0.16	–5.52, 11.71

[a]Daily steps was the primary outcome measure. DBP: diastolic blood pressure; DMA: duration of moderate activity; DSA: duration of sedentary activity; EE: energy expenditure; QOL: quality of life; SAQ: Seattle Angina Questionnaire; SBP: systolic blood pressure.

[b]Number of participants in the intervention group with complete baseline and 6-week follow-up data.

[c]Number of participants in the control group with complete baseline and 6-week follow-up data.

[d]Independent samples *t* test comparing change scores.

[e]The change values were normally distributed; therefore, mean (SD) values reported.

[f]Baseline and 6-week follow-up values were not normally distributed; therefore, median (IQR) values reported.

[g]Higher scores on this questionnaire represent better functioning.

Primary Outcome Measure

At 6 weeks, the intervention group had greater improvements in step count (+497 steps), whereas the control group had decreased level of steps (–861 steps), yielding an overall medium weight mean effect of 0.58 (95% CI 263-2451, *P*=.02).

Secondary Outcome Measures

Table 2 outlines the significant improvements in EE (ES=0.62, 95% CI 43.93-309.98, *P*=.01), DSA (ES=0.59, 95% CI –55.01 to –7.01, *P*=.01), DMA (ES=0.58, 95% CI 6.01-51.20, *P*=.01), weight (ES=0.52, 95% CI –1.78 to –0.15, *P*=.02), self-efficacy (ES=0.52, 95% CI 0.30-4.79, *P*=.03), emotional QOL score (ES=0.48, 95% CI 0.01-0.54, *P*=.04), and angina frequency (ES=0.77, 95% CI 8.57-35.05, *P*=.002) in the intervention group compared to the control group at the 6-week follow-up. Unexpectedly, there was also a significantly greater reduction in SBP in the control group compared to the Web-based cardiac rehabilitation group (ES=0.68, 95% CI 2.99-13.91, *P*=.001).

Medium-Term Intervention Effects

There were significantly lower levels of angina frequency (ES=0.63, 95% CI 1.89-29.41, *P*=.03) and increased social QOL score (ES=0.60, 95% CI 0.05-0.54, *P*=.02) favoring the intervention group at the 6-month follow-up. In contrast, there were no significant medium-term intervention effects in daily steps (ES=0.24, 95% CI–358 to 2324, *P*=.15), daily EE (EE=0.38, 95% CI –35.17 to 250.47, *P*=.14), DSA (ES=0.55, 95% CI 0.190-0.205, *P*=.20), DMA (ES=0.55, 95% CI 0.244-0.261, *P*=.24), weight (ES=0.35, 95% CI –2.46 to 0.34, *P*=.14), body fat percentage (ES=0.00, 95% CI –3.81 to 3.81, *P*>.99), SBP (ES=0.15, 95% CI –4.84 to 9.29, *P*=.53), DBP (ES=0.03, 95% CI –4.80 to 4.29, *P*=.91), fat intake (ES=0.30, 95% CI –6.12 to 1.80, *P*=.28), fiber intake (ES=0.29, 95% CI –2.23 to 8.53, *P*=.25), depression (ES=0.35, 95% CI –2.11 to 0.34, *P*=.15), anxiety (ES=0.47, 95% CI –2.60 to 0.04, *P*=.06), self-efficacy (ES=0.09, 95% CI –2.32 to 3.34, *P*=.72), physical QOL score (ES=0.29, 95% CI –0.11 to 0.43, *P*=.24), emotional QOL score (ES=0.46, 95% CI –0.02 to 0.62, *P*=.06), physical limitations (ES=0.08, 95% CI –7.20 to 10.50, *P*=.71), angina stability (ES=0.13, 95% CI –13.72 to 24.18, *P*=.58), treatment satisfaction (ES=0.08, 95% CI –15.31 to 10.69, *P*=.72), or disease perception (ES=0.17, 95% CI –8.41 to 14.99, *P*=.58). Although there were no significant intervention effects present for many of the outcome measures, it should be acknowledged that at the 6-month follow-up the intervention group showed trends of improved levels of baseline daily steps, EE, DSA, DMA, and weight, whereas the control group declined at the 6-month follow-up.

Usage of and Adherence to the Rehabilitation Program

Of the 48 intervention group participants, 19 (40%) completed the intervention and 29 (60%) did not progress past stage 3. The mean number of log-ins to the program was 18.68 (SD 13.13, range 1-51), an average of 3 log-ins per week per participant.

Discussion

Principal Findings

This study demonstrated daily physical activity improved as identified by step counts (our primary outcome). We also found significant improvements in a range of secondary outcome measures derived from the monitor, most importantly a reduction in sedentary time and an increase in the time spent being moderately active. This change in activity is an important outcome for this study because an important component of the website is to encourage daily exercise, most commonly walking. Although the changes were not significantly better at 6 months, there was a trend for the intervention group to remain improved compared to the control group, of which the effect sizes ranged from small to medium. This is in the absence of continued access to the site or any ongoing support. At the 6-week follow-up, we also observed important changes in weight, self-efficacy, emotional QOL score, and angina symptoms. We also observed significant changes at 6 months in angina symptoms and social QOL score.

Comparison With Previous Research

The use of technology and telehealth has been described previously in the literature to support individuals with CHD, but most of the studies have used telephone support as the technology [28]. The Internet has been used in a few projects examining a similar type of intervention. Antypas et al [9] recently assessed the effectiveness of an Internet- and mobile phone-based intervention for physical activity in a CHD population and reported increased physical activity levels in the intervention group compared to a control group at the 3-month follow-up; however, the study only had 7 participants at

follow-up and measured physical activity using self-reported measures. Southard et al [29] reported on an Internet-based intervention conducted in a mixed population across primary and secondary care; their data suggested that the intervention was effective in some areas, but failed to change levels of self-reported physical activity. Reid et al [8] reported significant group effects in physical activity and QOL following a Web-based physical activity intervention given to patients who had undergone percutaneous coronary revascularization. However, the intervention was not a comprehensive CHD secondary prevention package and targeted physical activity only.

Our current study has demonstrated significant improvements at the 6-week follow-up in walking, DSA, and DMA in comparison to a control group. We also observed improvements in weight, emotional QOL score, self-efficacy, and angina symptoms. At the 6-month follow-up, we were also able to demonstrate lowered angina symptoms, increased social QOL scores, and trends for physical activity to remain improved in the intervention group compared to the control group at the 6-month follow-up. This was in the absence of continued access to the site or any ongoing support because participants did not receive any support between the 6-week and 6-month follow-up assessment. This may not reflect what would happen in practice if this intervention was adopted in the health service.

Previous trials of a manual-based, self-management approach, The Angina Plan, reported significantly increased self-reported physical activity postintervention and at the 6-month follow-up [6,7]. Interestingly, the current study recruited participants with an established diagnosis of angina, whereas previous angina trials recruited those with a new diagnosis, a stage when motivation to adopt a healthier lifestyle may be higher. The Angina Plan is delivered over 12 weeks and comprises an initial in-depth consultation with a trained nurse and close facilitation by the nurse to encourage and discuss progress with agreed patient goals. The current online program was not facilitated in the same way; instead, the contact was initiated by the user via email. One might speculate that this would be a more cost-effective mode of delivery.

The proportion of intervention group participants completing the whole intervention was 40%, and 60% of participants progressed three-quarters of the way through the intervention (up to stage 3). This is comparable with Reid et al [8] who reported 43% of participants completed a Web-based physical activity intervention. Intervention completion rates in both the current study and Reid et al [8] are similar to the completion rate of traditional cardiac rehabilitation. In the United Kingdom between 2011 and 2012, an average of 52% of patients enrolling onto cardiac rehabilitation completed the intervention [4]. Overall, there were also regular website visits with an average of 3 log-ins per week. The mean number of website log-ins was 2 and 4 times per week in Southard et al [29] and Zutz et al [30], respectively.

Strengths and Limitations

This study evaluated the effects of an Internet-delivered self-managed cardiac rehabilitation program in an angina population with objectively measured physical activity as the primary outcome, a group seldom included within rehabilitation research or rehabilitation services despite current guidelines [31]. The researcher who collected the outcome measures also delivered the intervention. This allows for potential bias because participants with particularly high CHD risk could have unintentionally been encouraged more than other participants, which could have influenced the trial results. In future trials, researchers taking outcome measures should be blinded. In addition, it is necessary to consider measurement reactivity, in which measurement results in changes in the people being measured [32]. Although the study measured physical activity objectively, there still remains the possibility that participants may have adjusted their behavior while the activity monitor was worn.

The physical activity measurement period was 2 days. At the time, 2 days was the recommended monitoring period [33]. For future studies, we would propose wearing the monitor for a longer period, ideally 7 days. The study did not achieve the changes in physical activity that the power calculation was based on. In hindsight, this would appear to be an ambitious target because the power calculation was based on an intervention that was much more intense than the one described here. The data show that the intervention was effective in the short term, and the benefit was sustained in some outcomes at 6-month follow-up in the absence of access to the site or any ongoing support. In the future, we would wish to study the impact of continued access to the site for an extended period compared to best usual care. Due to limitations in funding, we were unable to collect any cost-effectiveness or health care utilization data, which would be desirable in future studies. Additionally, it would be valuable to assess if this intervention has an impact on smoking behavior. The current intervention does comprise a smoking cessation component, although the effect of this component was not examined in the current study because only 2 (4%) and 6 (13%) participants in the intervention and control group, respectively, were smokers at baseline. Future research should examine the intervention's impact on smoking cessation. The sample recruited in this study was primarily of a White British origin. Although this is not reflective of the general population, it is in-line with the ethnicity of patients currently receiving traditional cardiac rehabilitation as reported in a national audit. Challenges remain to find an acceptable intervention for ethnic minorities [4]. It would also be useful in future studies to compare the outcomes of an angina population using the Web-based rehabilitation program to an angina population receiving traditional rehabilitation.

In terms of the technological advances in health care, the program could also be developed into an application for use on a smartphone, and thereby enable the program to be available via mobile phone technology. Research examining the value of mobile phone-based interventions in increasing physical activity has been evaluated in a meta-analysis conducted by Fanning et al [34] and provides support for interventions using mobile technology to increase physical activity behavior.

Conclusions

The provision of support for those with angina is poor and these individuals are underrepresented in conventional cardiac

rehabilitation programs [4]. An Internet-based approach may offer an alternative self-management approach to either the Angina Plan or cardiac rehabilitation. The program is also likely to offer a lower-cost form of intervention and implementation of a Web-based alternative. This could widen the reach of rehabilitation and effectively increase service capacity. A large, pragmatic trial is required to examine the effectiveness and cost-effectiveness of this intervention when embedded into clinical practice. We would propose to offer access for a longer period of time. There may also be opportunities to explore the value of this intervention as an alternative to conventional rehabilitation for those who have suffered an acute cardiac event; this would help to increase the choice and scope of rehabilitation services.

Acknowledgments

We would like to thank the general practitioners in Coventry and Warwickshire that took part in this trial. SS was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM). Support was provided by the NIHR Leicester Respiratory Biomedical Research Unit. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Conflicts of Interest

The only study team member related to the intervention being evaluated is SS who participated in the design of the intervention evaluated in this manuscript. SS derives no financial benefit from the development of the intervention.

Multimedia Appendix 1

Program Registration.

[PDF File (Adobe PDF File), 90KB - jmir_v16i9e186_app1.pdf]

Multimedia Appendix 2

Program Registration.

[PDF File (Adobe PDF File), 127KB - jmir_v16i9e186_app2.pdf]

Multimedia Appendix 3

Feedback on user performance.

[PDF File (Adobe PDF File), 9KB - jmir_v16i9e186_app3.pdf]

Multimedia Appendix 4

Exercise diary.

[PDF File (Adobe PDF File), 89KB - jmir_v16i9e186_app4.pdf]

Multimedia Appendix 5

Stress management component.

[PDF File (Adobe PDF File), 51KB - jmir_v16i9e186_app5.pdf]

Multimedia Appendix 6

Communication with cardiac rehabilitation nurses via an online email link; ‘ask the expert’.

[PDF File (Adobe PDF File), 34KB - jmir_v16i9e186_app6.pdf]

Multimedia Appendix 7

CONSORT-EHEALTH checklist V1.6.2 [35].

[PDF File (Adobe PDF File), 999KB - jmir_v16i9e186_app7.pdf]

References

1. Buckley BS, Simpson CR, McLernon DJ, Murphy A, Hannaford P. Five year prognosis in patients with angina identified in primary care: incident cohort study. BMJ 2009 Aug 06;339(aug06 2):b3058-b3058. [doi: 10.1136/bmj.b3058]

2. Stewart S, Murphy NF, Murphy N, Walker A, McGuire A, McMurray JJ. The current cost of angina pectoris to the National Health Service in the UK. Heart 2003 Aug;89(8):848-853 [FREE Full text] [Medline: 12860855]
3. Cardiac rehabilitation service, Commissioning guide: Implementing NICE guidance. London: NICE; 2008 Mar. URL: http://www.nice.org.uk/media/49E/45/CardiacRehabilitationCommissioningGuide.pdf [accessed 2014-06-09] [WebCite Cache ID 6QCaJ8Nzr]
4. British Heart Foundation. The National Audit of Cardiac Rehabilitation: Annual Statistical Report 2013. 2013. URL: http://www.cardiacrehabilitation.org.uk/nacr/docs/2013.pdf [accessed 2014-06-09] [WebCite Cache ID 6QCaY6F84]
5. McGillion M, Arthur H, Victor JC, Watt-Watson J, Cosman T. Effectiveness of Psychoeducational Interventions for Improving Symptoms, Health-Related Quality of Life, and Psychological well Being in Patients with Stable Angina. Curr Cardiol Rev 2008 Feb;4(1):1-11 [FREE Full text] [doi: 10.2174/157340308783565393] [Medline: 19924272]
6. Lewin RJ, Furze G, Robinson J, Griffith K, Wiseman S, Pye M, et al. A randomised controlled trial of a self-management plan for patients with newly diagnosed angina. Br J Gen Pract 2002 Mar;52(476):194-6, 199 [FREE Full text] [Medline: 12030661]
7. Furze G, Cox H, Morton V, Chuang L, Lewin R, Nelson P, et al. Randomized controlled trial of a lay-facilitated angina management programme. J Adv Nurs 2012 Oct;68(10):2267-2279 [FREE Full text] [doi: 10.1111/j.1365-2648.2011.05920.x] [Medline: 22229483]
8. Reid RD, Morrin LI, Beaton LJ, Papadakis S, Kocourek J, McDonnell L, et al. Randomized trial of an internet-based computer-tailored expert system for physical activity in patients with heart disease. Eur J Prev Cardiol 2012 Dec;19(6):1357-1364. [doi: 10.1177/1741826711422988] [Medline: 21903744]
9. Antypas K, Wangberg SC. An Internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: short-term results of a randomized controlled trial. J Med Internet Res 2014;16(3):e77 [FREE Full text] [doi: 10.2196/jmir.3132] [Medline: 24618349]
10. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
11. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen AJ, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J Med Internet Res 2013;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
12. Richardson CR, Buis LR, Janney AW, Goodrich D, Sen A, Hess M, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]
13. Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, et al. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. Med Sci Sports Exerc 2004 May;36(5):897-904. [Medline: 15126727]
14. Dwyer TJ, Alison JA, McKeough ZJ, Elkins MR, Bye PT. Evaluation of the SenseWear activity monitor during exercise in cystic fibrosis and in health. Respir Med 2009 Oct;103(10):1511-1517 [FREE Full text] [doi: 10.1016/j.rmed.2009.04.013] [Medline: 19464863]
15. Hill K, Dolmage TE, Woon L, Goldstein R, Brooks D. Measurement properties of the SenseWear armband in adults with chronic obstructive pulmonary disease. Thorax 2010 Jun;65(6):486-491. [doi: 10.1136/thx.2009.128702] [Medline: 20522844]
16. Roe L, Strong C, Whiteside C, Neil A, Mant D. Dietary intervention in primary care: validity of the DINE method for diet assessment. Fam Pract 1994 Dec;11(4):375-381. [Medline: 7895964]
17. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
18. 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]
19. De Smedt D, Clays E, Doyle F, Kotseva K, Prugger C, Pająk A, EUROASPIRE Study Group. Validity and reliability of three commonly used quality of life measures in a large European population of coronary heart disease patients. Int J Cardiol 2013 Sep 1;167(5):2294-2299. [doi: 10.1016/j.ijcard.2012.06.025] [Medline: 22748284]
20. Schwarzer R, Jerusalem M. Causal and control beliefs. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Windsor, UK: NFER-NELSON; 1995:35-37.
21. Schwarzer R, Fuchs R. Self-efficacy and health behaviours. In: Predicting Health Behaviour: Research and Practice with Social Cognition Models. Buckingham, UK: Open University Press; 1996:163-196.
22. Höfer S, Lim L, Guyatt G, Oldridge N. The MacNew Heart Disease health-related quality of life instrument: a summary. Health Qual Life Outcomes 2004;2:3 [FREE Full text] [doi: 10.1186/1477-7525-2-3] [Medline: 14713315]
23. Spertus JA, Winder JA, Dewhurst TA, Deyo RA, Prodzinski J, McDonell M, et al. Development and evaluation of the Seattle Angina Questionnaire: a new functional status measure for coronary artery disease. J Am Coll Cardiol 1995 Feb;25(2):333-341 [FREE Full text] [Medline: 7829785]
24. Höfer S, Saleem A, Stone J, Thomas R, Tulloch H, Oldridge N. The MacNew Heart Disease Health-Related Quality of Life Questionnaire in patients with angina and patients with ischemic heart failure. Value Health 2012 Jan;15(1):143-150. [doi: 10.1016/j.jval.2011.07.003] [Medline: 22264982]

25. Activate Your Heart: Online Cardiac Rehabilitation Programme. URL: http://www.activateyourheart.org.uk/ [accessed 2014-06-09] [WebCite Cache ID 6QCmIne6H]
26. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
27. Tudor-Locke C, Bell RC, Myers AM, Harris SB, Ecclestone NA, Lauzon N, et al. Controlled outcome evaluation of the First Step Program: a daily physical activity intervention for individuals with type II diabetes. Int J Obes Relat Metab Disord 2004 Jan;28(1):113-119. [doi: 10.1038/sj.ijo.0802485] [Medline: 14569279]
28. Neubeck L, Ascanio R, Bauman A, Briffa T, Clark AM, Freedman B, et al. Planning locally relevant Internet programs for secondary prevention of cardiovascular disease. Eur J Cardiovasc Nurs 2011 Dec;10(4):213-220. [doi: 10.1016/j.ejcnurse.2010.07.003] [Medline: 20810323]
29. Southard BH, Southard DR, Nuckolls J. Clinical trial of an Internet-based case management system for secondary prevention of heart disease. J Cardiopulm Rehabil 2003;23(5):341-348. [Medline: 14512778]
30. Zutz A, Ignaszewski A, Bates J, Lear SA. Utilization of the internet to deliver cardiac rehabilitation at a distance: a pilot study. Telemed J E Health 2007 Jun;13(3):323-330. [doi: 10.1089/tmj.2006.0051] [Medline: 17603835]
31. National Institute for Health and Care Excellence. Cardiac rehabilitation services: Commissioning Guides 2013 URL: http://publications.nice.org.uk/cardiac-rehabilitation-services-cmg40 [accessed 2014-06-09] [WebCite Cache ID 6QCmerDCm]
32. French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br J Health Psychol 2010 Sep;15(Pt 3):453-468. [doi: 10.1348/135910710X492341] [Medline: 20205982]
33. Tudor-Locke CE, Myers AM. Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. Res Q Exerc Sport 2001 Mar;72(1):1-12. [Medline: 11253314]
34. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
35. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CHD: coronary heart disease
DMA: duration of moderate activity
DBP: diastolic blood pressure
DSA: duration of sedentary activity
EE: energy expenditure
GP: general practitioner
NIHR: National Institute for Health Research
QOL: quality of life
SAQ: Seattle Angina Questionnaire
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 23.02.14; peer-reviewed by L Neubeck, S Fraser; comments to author 11.04.14; revised version received 17.06.14; accepted 10.07.14; published 12.09.14.

Please cite as:
Devi R, Powell J, Singh S
A Web-Based Program Improves Physical Activity Outcomes in a Primary Care Angina Population: Randomized Controlled Trial
J Med Internet Res 2014;16(9):e186
URL: http://www.jmir.org/2014/9/e186/
doi:10.2196/jmir.3340
PMID:25217464

Review

Web-Based Intervention Programs for Depression: A Scoping Review and Evaluation

Tian Renton[1], BA (Hons); Herman Tang[1], BSc; Naomi Ennis[1], BA (Hons); Michael D Cusimano[1,2,3], MD, MHPE, FRCSC, DABNS, PhD, FACS; Shree Bhalerao[4], MD(Psych); Tom A Schweizer[3,5], PhD; Jane Topolovec-Vranic[1,2,6], PhD

[1]Trauma and Neurosurgery Program, St Michael's Hospital, Toronto, ON, Canada

[2]Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada

[3]Division of Neurosurgery, University of Toronto, Toronto, ON, Canada

[4]Medical Psychiatry Service, St Michael's Hospital, Toronto, ON, Canada

[5]Keenan Research Centre for Biomedical Science, St Michael's Hospital, Toronto, ON, Canada

[6]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jane Topolovec-Vranic, PhD
Trauma and Neurosurgery Program
St Michael's Hospital
Bond 3-012
30 Bond Street
Toronto, ON, M5B 1W8
Canada
Phone: 1 416 864 6060 ext 3421
Fax: 1 416 864 5629
Email: topolovec-vranicj@smh.ca

Abstract

Background: Although depression is known to affect millions of people worldwide, individuals seeking aid from qualified health care professionals are faced with a number of barriers to treatment including a lack of treatment resources, limited number of qualified service providers, stigma associated with diagnosis and treatment, prolonged wait times, cost, and barriers to accessibility such as transportation and clinic locations. The delivery of depression interventions through the Internet may provide a practical solution to addressing some of these barriers.

Objective: The purpose of this scoping review was to answer the following questions: (1) What Web-delivered programs are currently available that offer an interactive treatment component for depression?, (2) What are the contents, accessibility, and usability of each identified program?, and (3) What tools, supports, and research evidence are available for each identified program?

Methods: Using the popular search engines Google, Yahoo, and Bing (Canadian platforms), two reviewers independently searched for interactive Web-based interventions targeting the treatment of depression. The Beacon website, an information portal for online health applications, was also consulted. For each identified program, accessibility, usability, tools, support, and research evidence were evaluated and programs were categorized as evidence-based versus non-evidence-based if they had been the subject of at least one randomized controlled trial. Programs were scored using a 28-point rating system, and evidence- versus non-evidence-based programs were compared and contrasted. Although this review included all programs meeting exclusion and inclusion criteria found using the described search method, only English language Web-delivered depression programs were awarded an evaluation score.

Results: The review identified 32 programs meeting inclusion criteria. There was a great deal of variability among the programs captured in this evaluation. Many of the programs were developed for general adolescent or adult audiences, with few (n=2) focusing on special populations (eg, military personnel, older adults). Cognitive behavioral therapy was the most common therapeutic approach used in the programs described. Program interactive components included mood assessments and supplementary homework sheets such as activity planning and goal setting. Only 12 of the programs had published evidence in support of their efficacy and treatment of depressive symptoms.

Conclusions: There are a number of interactive depression interventions available through the Internet. Recommendations for future programs, or the adaptation of existing programs include offering a greater selection of alternative languages, removing registration restrictions, free trial periods for programs requiring user fees, and amending programs to meet the needs of special populations (eg, those with cognitive and/or visual impairments). Furthermore, discussion of specific and relevant topics to the target audience while also enhancing overall user control would contribute to a more accessible intervention tool.

(J Med Internet Res 2014;16(9):e209) doi:10.2196/jmir.3147

KEYWORDS

depression; Web-based interventions; interactive treatment; health care access; mental health; technology

Introduction

In October 2012, the World Health Organization estimated that 350 million people worldwide suffer from depression [1]. In Canada alone, 12.2% of the adult population has met depression diagnosis criteria within their lifetime [2]. It is expected that depression will be the world's largest medical burden on health by 2020 [1]. Beyond its burden on society, depression is associated with worse global outcomes for the affected individual, including reduced social functioning, lower quality of life in regards to health, inability to return to work, as well as suicide [3-5].

Not all individuals face the same risk. Certain populations are more susceptible to depression relative to others. For example, it is estimated that 14%-77% of traumatic brain injury survivors suffer from depression post injury, which is a much higher than the general population [6-11]. The prevalence of depression is also moderated by gender. Research has indicated that almost double the number of women (10%) relative to men (5%) suffer from depression within their lifetime [12]. Moderated by age, young people aged 15-24 years were found more likely to experience substance abuse disorders and mental health disorders than any other age group [13].

There are a number of depression treatment options available including medication, lifestyle, and psychological interventions. Unfortunately, many therapies are unavailable to patients due to a lack of qualified service providers. In Canada, it is estimated that only 33% of individuals seeking mental health services actually receive treatment [14]. Patients also experience barriers such as prolonged wait times, cost, and accessibility barriers such as transportation and clinic locations. Within Ontario, 60% of family physicians rated accessibility to psychiatrists as fair to poor [13].

Not only are individuals hindered by accessibility issues but also by the stigma associated with a mental illness diagnosis such as depression. Only 42% of Canadians said that they would socialize with a friend suffering from a significant mental health issue [13], and 46% of Canadians believe mental illness is used as a term to excuse poor behavior [13]. Associated stigma and treatment barriers stress the need for alternative health care options. These barriers may partially explain why a recent study found that almost 49% of individuals who believe they have suffered from depression or anxiety have never sought professional health care [15].

Recent advances in computer technology offer potential alternative treatment options. The percentage of the global population that had access to the Internet nearly doubled from 18% in 2006 to 35% in 2011 [16]. As the scope of applications supported by technology and their use have increased over time, educational resources and Web-based treatment interventions have emerged to target mental health issues. At present, treatment programs are available for eating disorders [17], smoking cessation [18], obesity [19], safe sexual practices [20], autism spectrum disorders [21], substance use disorders [22], physical activity for type 2 diabetics [23], lifestyle behavior improvement [24], cannabis use [25], reducing chronic obstructive pulmonary disease risk behaviors [26], mental health in tertiary students [27], and chronic illness management [28], as well as others.

Regarding Web-based intervention treatments for depression specifically, the Canadian Network for Mood and Anxiety Treatments suggested computer-assisted cognitive behavioral therapy (CBT) programs as a second line treatment in cases "where first line treatments are not indicated or cannot be used or when first line treatments have not worked" [29]. However, programs must provide at a minimum Level 3 evidence (non-randomized controlled prospective studies or case series or high quality retrospective studies) and additional clinical support [29]. Similarly, the Improving Access to Psychological Therapies program of the UK National Institute for Health and Clinical Excellence recommends computerized CBT as a low-intensity intervention for depression [30].

Reviews of existing Web-based programs for the treatment of depression are available on the Beacon website [31], as well as systematic reviews of the effectiveness of such programs [32-34]. On the Beacon website, health care experts review, categorize, and rate Web-based programs, mobile applications, and support groups. A brief summary is provided for each site including a description of the program, details of the site, target audience, access requirements, and the extent of research evidence in support of the program. A rating is provided for each site based on the degree to which research evidence is available. Currently, 45 websites, 5 mobile apps, and 8 support groups have been identified and reviewed on the Beacon site.

We conducted a scoping review of interactive Web-based treatment programs for depression to answer the following questions: (1) What Web-based programs are currently available that offer an interactive treatment component for depression?, (2) What are the contents, accessibility, and usability of each identified program?, and (3) What tools, supports, and research evidence are available for each identified program? Unlike the systematic reviews that have been published to date, our review included publicly available Web-based programs with and

without supporting research trials. In addition, we have compared and contrasted those with and without supporting evidence on predetermined evaluation criteria.

Methods

Search Strategy

A search for interactive programs targeting depression was conducted in April 2014 using Google, Yahoo, and Bing (Canadian versions)—popular and comprehensive search engines easily accessible to many individuals. A list of the search terms used is included in Textbox 1. A program was included in the review if it (1) targeted the treatment of depression, (2) was accessible via the Internet, (3) had an interactive treatment component (ie, was not purely educational) and required user participation (ie, homework, worksheets, mood assessment), and (4) was available in English (additional languages allowed). A program was excluded if it (1) solely provided information regarding depression (ie, psychoeducation), (2) solely targeted the treatment of mental health issues other than depression (eg, anxiety, bipolar disorder, post-traumatic stress disorder), (3) was not accessible to the public (private programs), (4) solely targeted health care professionals for training purposes, (5) offered only mood tracking applications, (6) required supplemental equipment that would not be publicly available, (7) was solely available for research purposes (ie, user must be enrolled in the study to access the program), (8) offered no treatment program, (9) could not be completed within the home or private setting (ie, must attend classes), and (10) offered online counseling only (ie, no program associated with the website).

For the purposes of this scoping review, interactive was defined as a program requiring user engagement and input (eg, mood assessments, user worksheets, and integrated program requirements for mandatory user feedback). If a program provided only reading materials on the symptoms of depression, treatments, and relapse prevention, it was classified as psychoeducational and did not meet the "interactive" inclusion criterion.

Textbox 1. Search terms used in the search engines Google, Bing, and Yahoo.

1. Online depression treatment
2. CBT depression online treatment
3. Depression and CBT online programs
4. Online methods of depression treatment
5. Computerized cognitive behavior therapy programs for treatment of depression
6. Online cognitive behavior therapy programs treating depression
7. Computer based depression treatment programs
8. Internet delivered depression treatment programs
9. E-therapy for depression

Each of the terms in Textbox 1 was entered into the three identified search engines. A search log outlines the number of hits included and excluded, in addition to duplicate entries (see Multimedia Appendix 1). A search was terminated for the term when five consecutive pages of search results failed to identify any new programs. The page on which the search was terminated is included in the search log.

The Beacon website [31] was additionally investigated for programs meeting our inclusion criteria. Nine additional programs were identified from Beacon that had not been identified using our search strategy. Two independent reviewers completed the search as indicated above, resulting in only five discrepancies among the programs identified. After careful evaluation, a consensus was reached regarding included programs, with the search providing 20 programs and Beacon an additional nine programs. The remaining three programs were previously known to reviewers: After Deployment, Students Against Depression, and Dealing With Depression.

Program Evaluation Criteria

In order to systematically evaluate each identified program, we used criteria adapted from previously published guidelines [35] to generate five categories of evaluation that were broken down into 14 subcategories. Table 1 provides a detailed description of the category of assessment, the investigation, and the evaluation focus.

Table 1. Categories of investigation used to evaluate each program.

Category	Investigation	Evaluation focus
Accessibility	Fee/Referral	Was there a fee or physician referral required to access the program?
	Language	Was the program available in alternative languages?
	Registration Requirement	Was personal information required to access the program? Were details provided regarding the registration process?
	Target Audience	Who was the program designed for?
Usability	Statistics	Registered users, completion rates, and attrition data (if available).
	Therapeutic Approach	What therapies/treatment approach(es) were offered?
	Mode of Delivery	How was the content delivered: size of text, audio or video offered, use of character examples, and case scenarios?
Tools	Additional Features	What additional features were available (eg, if users could monitor their progress/modules completion and mood over time)? Were email reminders and follow-up offered?
	Worksheets	Did the program offer worksheets (printable for offline use or integrated throughout program)? Indicated whether worksheets were mandatory or optional.
	Assessments	Were assessments offered within the program? Indicated whether assessments were mandatory for completion of program.
Support	Clinician Support	Did the program offer linking with a clinician (either user's own clinician or program specific clinician) and type of linkage (eg, telephone monitoring)?
	Peer Support	Did the program offer peer support (eg, forum, personal story sharing, or blogs)?
	Crisis Links	Did the program offer crisis or emergency contacts?
Evidence	Randomized Controlled Trial	Had the program been evaluated for efficacy with at least one randomized controlled trial?

A program evaluation scoring system was also created based on the above mentioned article [35] to provide an objective numerical score for each of the captured programs. Permissions were obtained from the publishing journal to adapt the scoring system: 28 evaluation points were adapted from the 12 facets and guidelines discussed in the article, and facets were adapted into yes or no, closed-ended inquiry questions. Programs meeting the inquiry were awarded 1 point. Programs not meeting the inquiry were given 0 points or a 0[a] in the case that the point could not be evaluated using the program of interest and/or attempts to contact program developers. Table 2 outlines the facets and guidelines of the adapted scoring system. Summative program scores (absolute and relative [%]) are listed in Multimedia Appendices 2 and 3. A detailed breakdown of the points awarded per program is provided in Multimedia Appendix 3.

Two reviewers independently examined each program and scored them using the evaluation criteria outlined in Table 2. The reviewers achieved initial agreement on 782 of the 812 (96.3%) evaluation points (28 possible points across 29 programs). The reviewers then met and resolved the discrepancies through further program investigation or consulting developer/reviewer email communications until 100% agreement was achieved.

An identified program was categorized as an evidence-based program (EBP) if it was the intervention in at least one published randomized controlled trial (RCT). All other programs were categorized as non-evidence-based programs (NBP).

Table 2. Facets and adapted evaluation criteria.

Facet # and description	Adapted evaluation description and #
1. Focus and target population	1. Were the primary focus/ goals/ objectives of the intervention stated?
	2. Was an initial assessment conducted for program/user suitability purposes?
	3. Was the target audience or age group defined?
2. Authorship details	4. Were the names and credentials of authors present?
	5. Was the ownership or developer name provided?
	6. Were links to the developer website provided?
	7. Was date of program/site update provided?
	8. Was country of origin stated?
3. Model of change	9. Was the model of change (ie, type of therapy utilized) defined/stated?
4. Type and dose of intervention	10. Were the number of modules or time to complete each module stated?
	11. Was the program structured/guided (ie, modules to be completed in a restricted and specific order =1) or unstructured/unguided (ie, modules could be accessed freely =0)?
	12. Was the intervention tailored to the user or was it generic for all users?
	13. Did users receive feedback?
	14. Could users track their progress throughout the program?
	15. Were the assessments validated/reliable?
5. Ethical issues	16. Were the risks of the program stated/benefits of program were stated?
	17. Were safe guards provided (ie, crisis links /telephone hotline numbers provided)?
	18. Was a unique user name or password provided to users?
	19. Was the site secure?
	20. Were the rights and use of user personal information provided?
6. Professional support	21. Was there a statement of professional support (ie, therapist integrated into the intervention)?
	22. For programs utilizing therapist support: were the credentials of the therapist provided?
7. Other support	23. Was support provided from additional sources (ie, peer discussion forums or blogs)?
	24. Was this type of support monitored by an overseeing authority?
8. Program interactivity	25. Was the interactivity of the program described and accurate (ie, how much time needing to be spent on module/homework assignments)?
9. Multimedia channel of delivery	26. Did the program offer a multimedia content delivery (ie, a combination use of text, video, graphics, and audio formats)?
10. Degree of synchronicity	This evaluation point was combined with point 13.
11. Audience reach	This evaluation point is available in Multimedia Appendix 2, under the heading “Target Audience”
12. Program evaluation	27. Was evidence for the program provided to the user (ie, attrition data/ success rate/ completion rate/ # of users in the program/ testimonials)?
	28. Was a randomized controlled trial completed for the program?

Results

Summary

We identified 27 websites collectively offering 32 programs in accordance with inclusion and exclusion criteria (see Multimedia Appendix 2 for identified program names). Two websites offered multiple programs targeting the treatment of depression: eCentre Clinic (The Mood Mechanic Course, The Wellbeing Course, The Wellbeing Plus Course, and The UniWellbeing Course), and This Way Up (Clinical Course, Worry and Sadness Self-Help Course, and Schools Course).

Three programs meeting inclusion criteria were not found using the searches conducted but were previously known to authors: After Deployment, Dealing With Depression, and Students Against Depression. Three programs (Interapy, Kleur Je Leven, and Internetpsykiatri) included in Tables 3-5 and Multimedia Appendix 2 were not included in the evaluation scoring log (Multimedia Appendix 3) due to language barrier restrictions. In sum, 32 programs met inclusion criteria and were included in this review for evaluation. Tables 3-5and Multimedia Appendices 2 and 3 are organized alphabetically and also by whether the program was evidence-based or not (EBP vs NBP). The first 12 programs included in Tables 3-5 and Multimedia

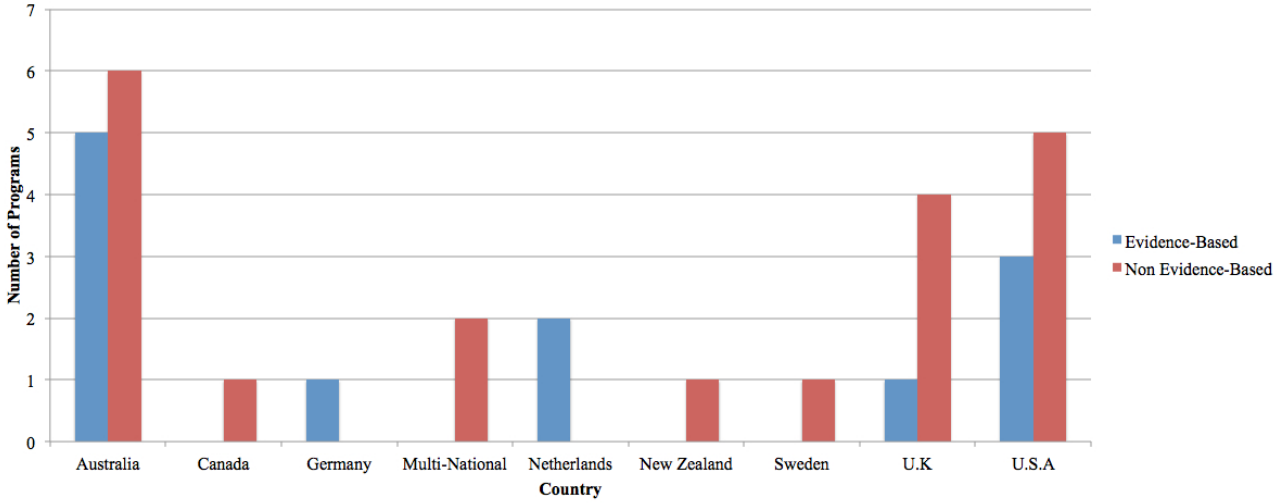
Appendices 2 and 3 were evidence-based (see Multimedia Appendix 4 for citations) with the remaining 20 programs having no identified evaluative RCTs. See Multimedia Appendix 5 for program screenshots.

Accessibility

Country of Origin

A summary of the countries where programs were developed is provided in Figure 1. The countries of origin for the EBP included Australia, Germany, Netherlands, the United Kingdom, and the United States. The countries of origin for the NBP were Australia, Canada, multinational collaborations, New Zealand, Sweden, the United Kingdom, and the United States.

Figure 1. Program development by country of origin.



Language

Four of the EBP were offered in alternative languages (Deprexis: Swedish and German; eCentre Clinic: Wellbeing Course: Arabic; MoodGYM: Simplified Chinese, Dutch, and Norwegian; and MoodHelper: Spanish). Two programs were available only in Dutch (Interapy and Kleur je Leven). Only one NBP was offered in another language and was not available in English (Internetpsykiatri: Swedish only). The remaining 19 programs were offered exclusively in English.

Fees or Referrals and Registration

The registration processes and fee structures varied by program for both the EBP and NBP and included those that were freely available (n=12), those with a fee required (n=8), those with an access code required (n=4), those for which an application to the course was required (n=4), those with multiple registration criteria (n=1), and those with unknown registration criteria (n=3). See Figures 2 and 3.

Eight programs had an associated fee ranging from AUD $55 (This Way Up, Clinic Course) for five/six lessons to US $400 for eight sessions (The National Stress Clinic). Most programs offered a free trial period, allowing users to interact with the program prior to paying the fee.

Most programs (n=21) required the input of personal information to access course material. Registration was restricted to certain countries for five programs (eCentre Clinic: The Wellbeing Course,The Mood Mechanic Course, The Wellbeing Plus Course, and The UniWellbeing Course, restricted to Australia; MoodHelper, restricted to the United States). Registration was not required for three programs, and registration requirements were unknown for three programs.

The programs that did not mandate registration allowed users to access program content without entering any personal information. However, users could not track their progress through these programs without registration. Programs with mandatory registration required users to input basic personal information before gaining access to program content, allowing for sessional data storage. For example, registration allowed users to create personal profiles through which mood assessments, worksheets, and module progression were recorded.

Figure 2. Fees and referrals required for program access.

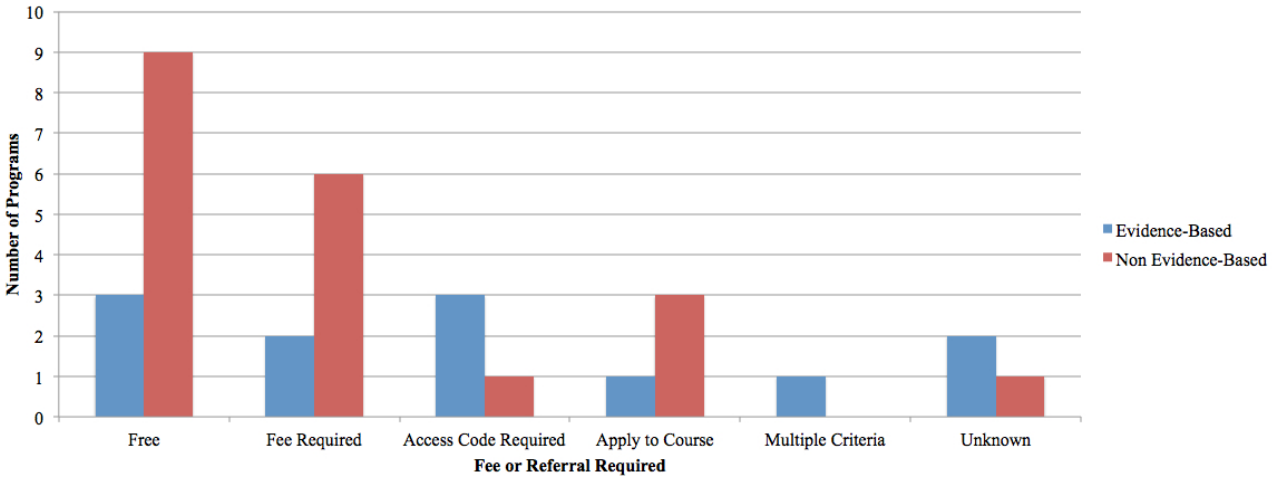
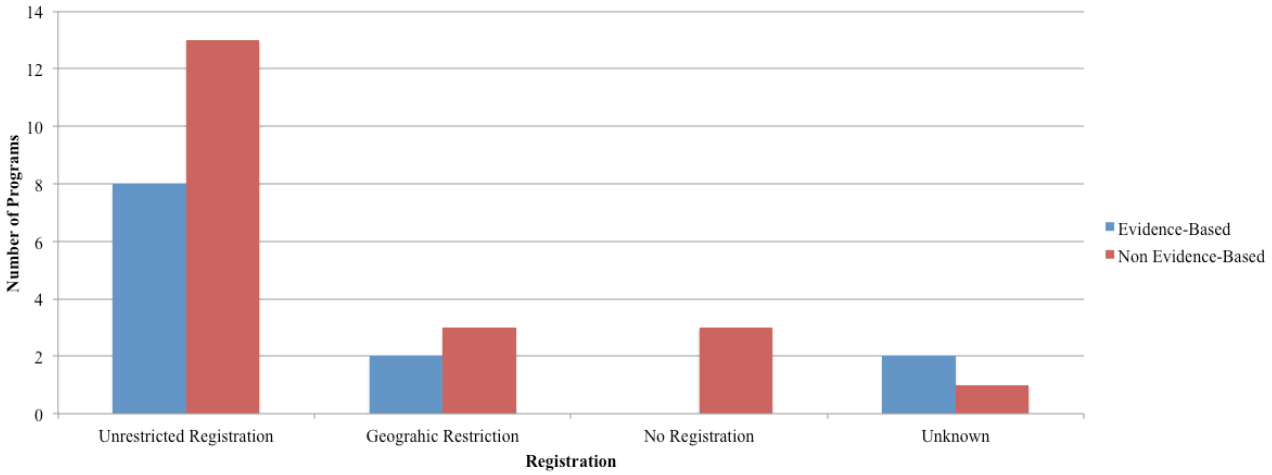


Figure 3. Registration required for program access.



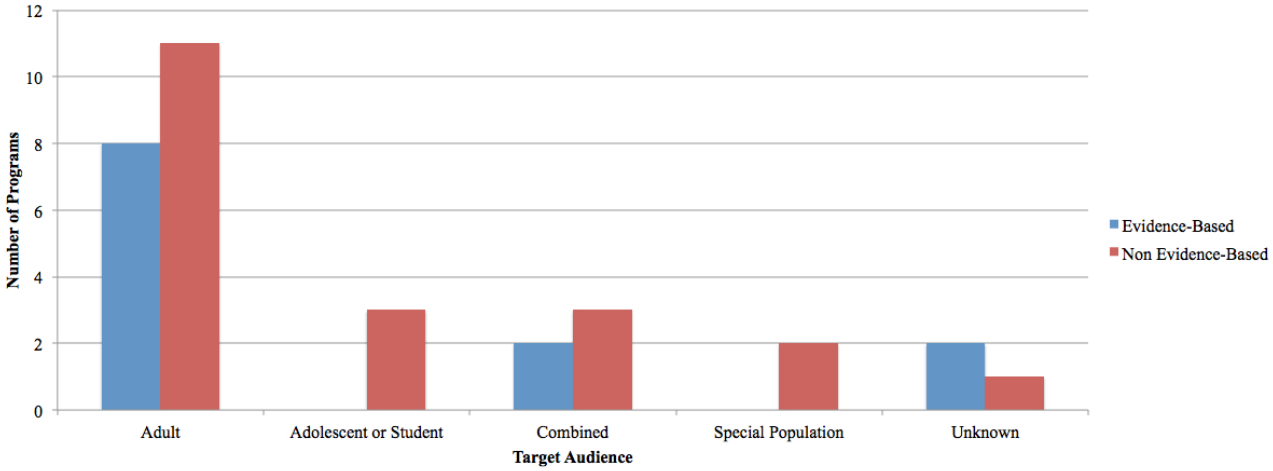
Target Audience

The identified programs predominantly targeted adult audiences (n=19), with a few programs developed for adolescents or students (n=3), combined populations (ie, adolescent and adult population; n=5), a specified special population (n=2), or unknown audiences (n=3) (see Figure 4).

The majority of programs contained content that was specific to the adult end-user. Sessions varied by program and offered an array of materials targeting adult concerns surrounding depression. Topics included but were not limited to problem solving, goal setting and planning, tackling financial issues, workplace stress, in addition to challenging negative thoughts. Programs targeting adolescents were mainly focused on academic life and the stressors associated with the learning environment. Topics included time management, stress associated with exams, relationships, confidence and self-esteem, as well as other social issues faced by youth.

Two NBP were developed for specific target populations: After Deployment (discharged military personnel) and eCentre Clinic Wellbeing Plus Course (older adults). Sessions covered many issues faced by these individuals including post-traumatic stress syndrome, insomnia, depression, and anxiety.

Figure 4. Target audience(s) of the programs.



Evaluation Score

Evaluation scores for EBP ranged from 68%-92%, with an average score of 80%. Evaluation scores for NBP ranged from 39%-92%, with an average score of 73%. A detailed breakdown of scores per program is provided in Multimedia Appendix 3.

Usability

Program Statistics and Attrition

A number of programs (n=17) provided user statistics and/or user reviews. Varying by program, statistics aimed to demonstrate overall program efficacy and effectiveness using program completion rates, number of registered users, number of users completing the course, and reductions in validated measure scores from pre- to post-intervention. See Table 3 for usability results.

Table 3. Usability of the programs evaluated.

Ref. #	Statistics available (# of users registered, number of unique visitors, attrition data, stated completion rate)	Therapeutic orientation/Intervention offered[a]	Content delivery[b]
1	89% of users rated the program as useful, 70% of users (mild to moderate depression) who completed the program required no further treatment	CBT	ANI/TXT++/VID
2	—	PE/CBT	AUD/TXT/VID/ANI
3	7000 patients currently enrolled in trials	CBT/PP	AUD/TXT++
4	Have offered free treatment to more than 4000 Australians	CBT/IPT	TXT++
5	—	CBT	ANI/TXT++
6	10,804 new registrations, 71,113 unique visitors, 9,199,943 Internet hits	CBT/IPT/PS/RX/PA	TXT++/VID
7	—	UNK	UNK
8	—	UNK	UNK
9	750,000 registered users	CBT	TXT++/ANI
10	Over 2500 users registered	CBT	TXT+/ANI
11	75% of users complete the course and require no further treatment, 6000 patients enrolled and 2400 clinicians as of Dec. 2012	CBT	ANI/TXT+
12	—	CBT	ANI/TXT+
13	—	UNK	AUD/TXT+/VID
14	54% average reduction in Patient Health Questionnaire (PHQ)-9 Score	MD/CBT	AUD/TXT+/VID
15	Reported decreases in Beck Depression Index scores and reportedly 60% of individuals who finish program will be “cured” of their depression/ testimonials	CBT	TXT++/ANI
16	—	CBT	VID/TXT++
17	—	UNK	AUD/TXT++/VID
18	Have offered free treatment to more than 4000 Australians	CBT/IPT	TXT++
19	Have offered free treatment to more than 4000 Australians	CBT/IPT	TXT++
20	Have offered free treatment to more than 4000 Australians	CBT/IPT	TXT++
21	—	UNK	UNK
22	222,078 registered users	CBT	ANI/TXT++
23	—	CBT/PP	VID/TXT++/AUD
24	600,000 users registered, 437,507 unique visitors, 79,607,184 Internet hits.	CBT/IPT	TXT++/ANI
25	—	CBT/IPT/PS/PP	TXT++
26	—	CBT	TXT+++
27	Testimonials and user reviews	CBT	UNK
28	User reviews	CBT	TXT++/ VID
29	—	CBT	AUD/TXT+++
30	User reviews	CBT	TXT++
31	—	UNK	VID (subtitles avail.)/TXT+++/AUD
32	—	CBT	TXT+/ANI

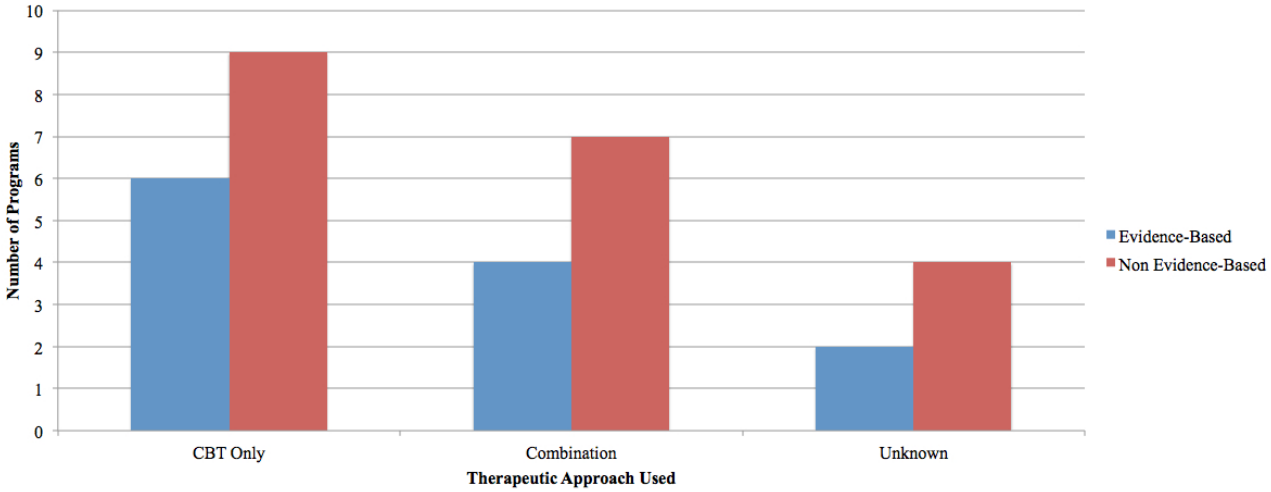
[a]MD:mindfulness, PA: physical activity, PE: psychoeducation, PP: positive psychology, PS: problem solving, RX: relaxation, UNK: unknown.

[b]ANI: animations/graphics, AUD: audio files, TXT: text based (+=small text blocks, ++=medium text blocks, +++=large text blocks), VID: video files.

Therapeutic Approach

Programs delivered their interventions through various therapeutic techniques. The majority of the EBP (n=6) delivered CBT-focused treatments, four offered integrated therapies (eg, CBT, Intrapersonal Therapy [IPT], psychoeducation, relaxation therapy, problem solving, and physical activity), and two did not define their therapeutic approach. Nine NBP provided CBT-based materials, seven offered combination therapy models, and four NBP were categorized as unknown as they did not define their therapeutic approach. See Figure 5.

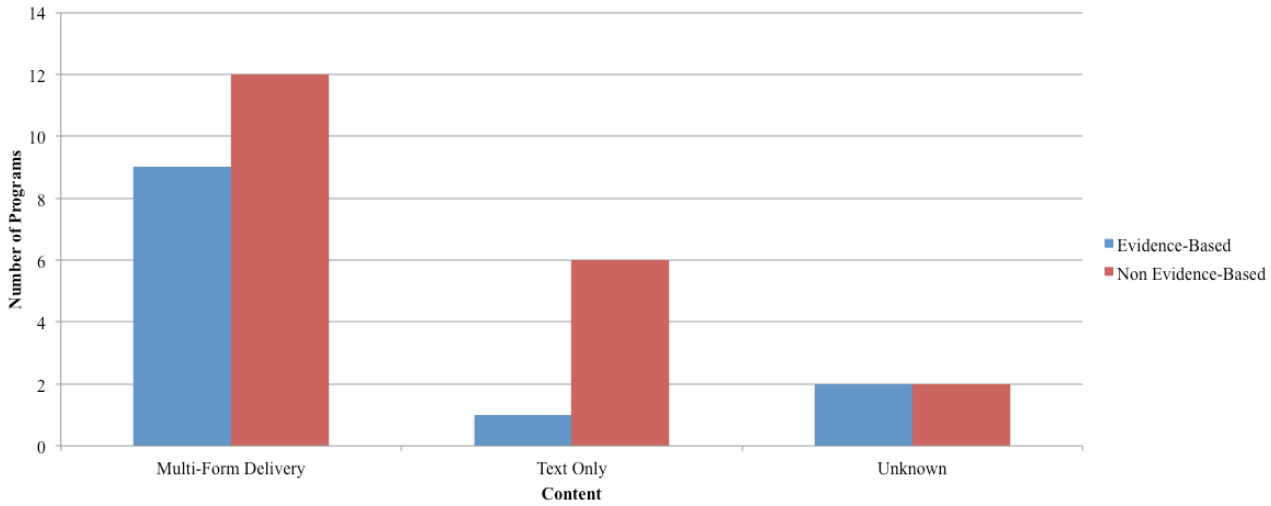
Figure 5. Therapeutic approach offered in programs evaluated.



Mode of Delivery

Treatment interventions were delivered in a number of ways for both EBP and NBP. Nine EBP contained a multimodal format (ie, combination of text, video, and audio) with only one delivered in a text-only format. Two EBP were unknown (could not be evaluated). The majority of the NBP (n=12) were offered in a multimodal format. Six NBP were offered in a text-only format, and two programs were categorized as unknown. See Figure 6.

Figure 6. Method of content delivery.



Quality Assessment

Table 4 provides information regarding the available tools within each program such as worksheets, assessments, and additional features (eg, email reminders, calendars, journal functions). Table 5 provides information regarding program supports from peers, from the clinician, and available crisis links.

Table 4. Tools associated with the programs evaluated.

Ref. #	Additional features[a]	Worksheets[b]	Assessments[c]
1	PT	W	UNK
2	PTP	W	A[V]
3	RE/MT/PTP/PT/F	W	A[V]
4	PT/F	W	A[V]
5	PT/F/MT	W	A
6	PT/MT/F	W	A[V]
7	UNK	UNK	UNK
8	UNK	UNK	UNK
9	PT/MT/F	W	A[V]
10	MT/PT/F	W	A
11	CA/PT/F	W	A[V]
12	CA/PT/F	W	A
13	PT/F/MT	W	A[V]
14	JOUR/PT/F/MT	W	A[V]
15	PTP/F/PT	UNK	A[V]
16	PT/F	W	A[V]
17	—	—	—
18	PT/F	W	A[V]
19	PT/F	W	A[V]
20	PT/F	W	A[V]
21	UNK	UNK	UNK
22	MT/PT/PTP/F	W	A[V]
23	PT/F	W	A
24	MT/PT/JOUR/PTP/F	W	A
25	PTP/F	W	A
26	PT/F/RE	W	A
27	PT/F/MT	W	A
28	MT/CA/PT/F/JOUR	—	A[V]
29	—	W	—
30	MT/PT/F	W	A[V]
31	PT/CA/F	W	A[V]
32	PT/F	W	A[V]

[a]CA: calendar application, JOUR: journal application, MT: mood tracking, F: feedback provided, PT: progress tracking, PTP: personalized treatment plan, RE: reminder emails.

[b]UNK, unknown, W: worksheets available.

[c]A: assessments available, A[V]: assessment validated.

Table 5. Support associated with the programs evaluated.

Ref. #	Peer support[a]	Clinician support[b]	Crisis links[c]
1	—	TC	UNK
2	—	—	UNK
3	—	TC	EC
4	—	TC	EC
5	—	—	EC
6	UNK	TC/—	UNK
7	UNK	UNK	UNK
8	UNK	UNK	UNK
9	—	—	EC
10	—	TC	EC
11	—	TC	EC
12	—	—	EC
13	PSF	—	EC
14	UNK	TC	—
15	—	—	—
16	—	—	EC
17	—	—	—
18	—	TC	EC
19	—	TC	EC
20	—	TC	EC
21	UNK	UNK	UNK
22	PSF	TC	EC
23	PSF	—	—
24	—	—	EC
25	PSF	—	—
26	PSF	TC	EC
27	PSF	TC	EC
28	—	—	EC
29	PSF	—	EC
30	PSF	—	—
31	—	TC	EC
32	—	—	EC

[a]PSF: peer support forum.

[b]TC: therapist contact via telephone/email and/or therapist linking to program, UNK: unknown.

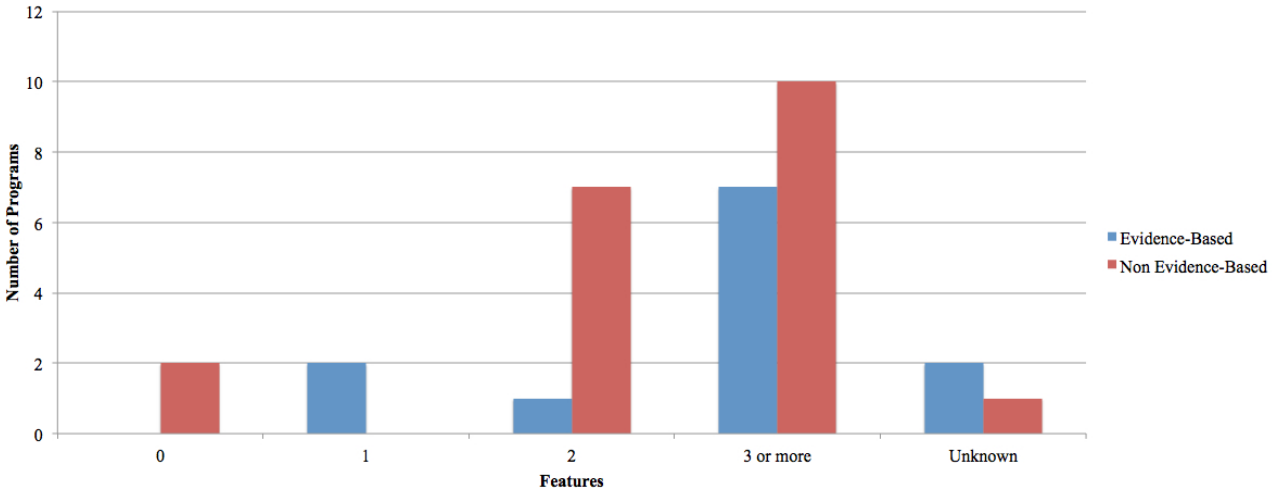
[c]EC: emergency contact information provided.

Tools

Additional Features

Many programs included additional features such as email reminders, calendar applications, journal space, progress tracking reports as well as mood tracking. EBP and NBP both offered a number of support tool options, with most programs offering multiple features (EBP=7, NBP=10). See Figure 7.

Figure 7. Additional features offered to program users.



Worksheets

Many programs provided lesson reinforcement activities and worksheets (EBP: n=10, NBP: n=16). Worksheets covered topics discussed in current or previous modules and encouraged user engagement. Many worksheets incorporated activity planning, goal setting, problem solving, and thought evaluation. Self-reflection activities further identified troublesome areas and encouraged corrective action.

Assessments

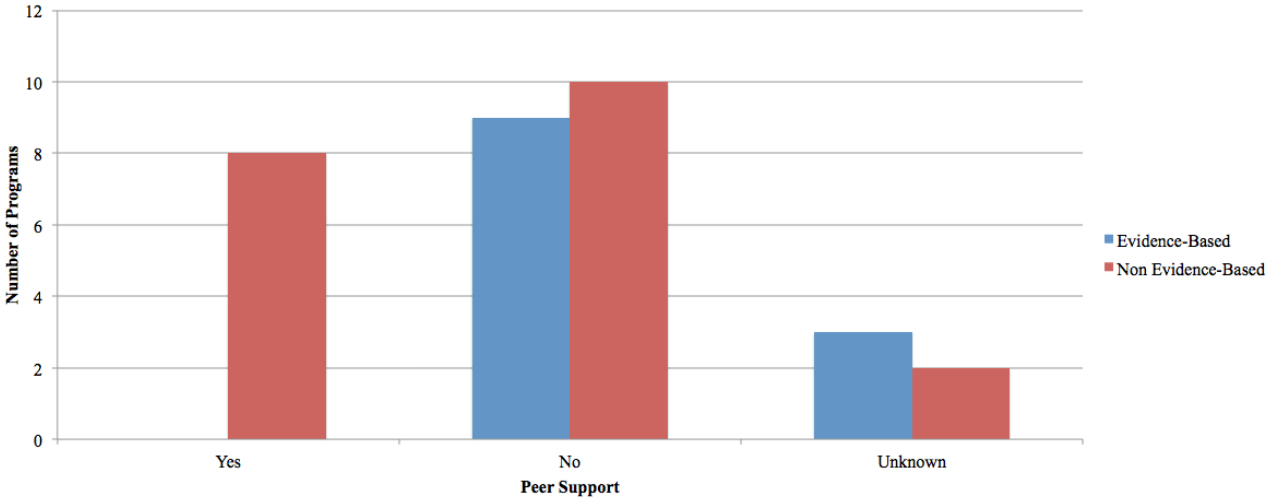
Most programs incorporated mood or depression assessment tools (EBP: n=8, NBP: n=17). Assessments were delivered prior to user registration, integrated into module content and/or independent of session programming. Programs administering assessments provided feedback and results immediately upon user completion. When evaluating some programs (EBP: n=3, NBP: n=1), it was unclear whether assessments were administered. In total, 17 programs used validated measures: Patient Health Questionnaire (PHQ)-9, Beck Depression Inventory (BDI), and Center for Epidemiologic Studies Depression Scale (CES-D).

Support

From Peers

None of the EBP offered a peer-support forum (three programs were categorized as unknown). The majority (n=11) of NBP did not include a peer-support forum (one program was categorized as unknown). Only eight NBP were found to have this service available to its users. See Figure 8.

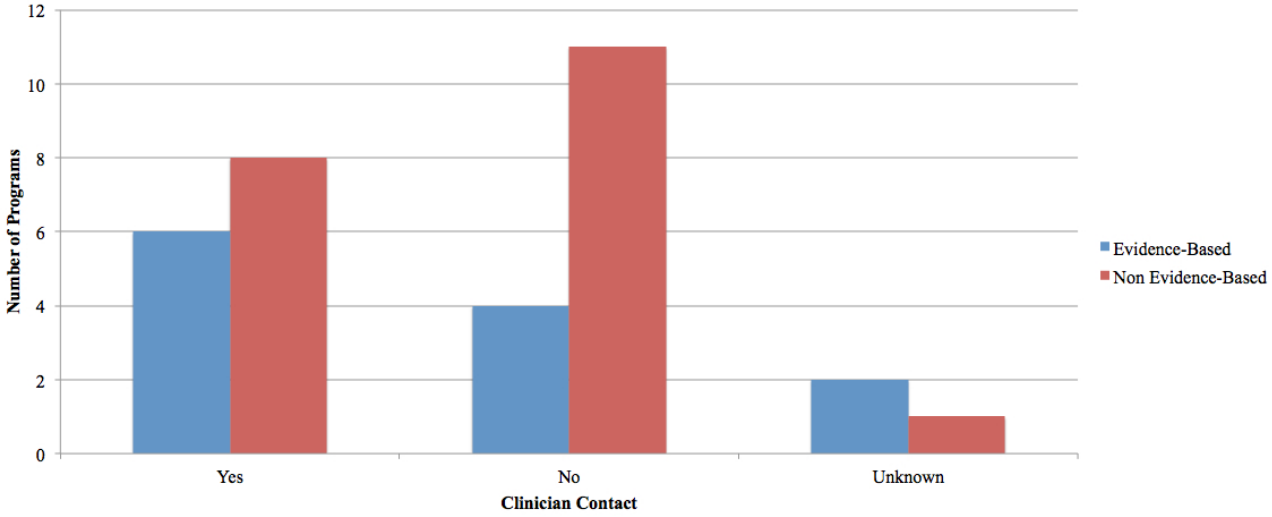
Figure 8. Support available from peers.



From the Clinician

Six EBP offered clinician support, four offered no clinician support, and two programs were unknown. Eight NBP offered clinician support, 11 programs did not offer clinician support, and one program was not evaluated. See Figure 9.

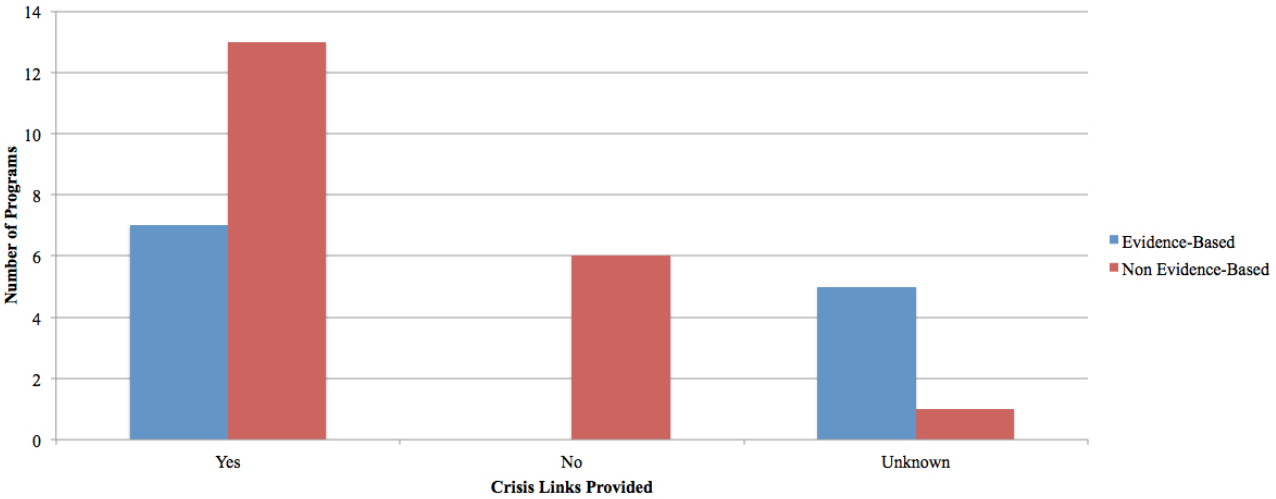
Figure 9. Support available from clinicians.



Crisis Links

Crisis links were defined as email addresses, phone numbers, and/or hotlines connected to distress centers providing counseling services to at-risk users. If present, telephone numbers and distress centers were from within the program's country of origin. In total, 20 programs (both EBP and NBP) had a crisis link with contact information and phone numbers. Most EBP provided this service (five programs were unknown); however, six NBP did not (one program unknown). See Figure 10.

Figure 10. Crisis link information provided.



Research and Publications

Only 12 of the 32 programs had at least one published RCT evaluating their efficacy. It was beyond the scope of the current review to summarize the results of these trials and the reader is directed to recently published systematic reviews and meta-analyses [32-34] (see Multimedia Appendix 4).

Discussion

Principal Findings

Use and wide spread dissemination of Web-based mental health care interventions is expanding and reflected by the number of currently available depression treatment programs captured in this scoping review. Web-based approaches may have several benefits beyond those of conventional psychotherapy [36-38]. Such approaches (1) are accessible at any time of day, unlike a traditional clinic setting with predetermined hours of operation, (2) are accessible from home therefore removing transportation barriers to treatment and health care provider shortages (particularly in rural communities), (3) are cost effective and reach a mass audience, (4) are self-guided by the user, and (5) allow the user to remain anonymous while in the setting of their choice (privacy). However, it is important to note that Web-based approaches may have several disadvantages including the potential perpetuation of secrecy associated with mental illness and depriving the individual of the opportunity for face-to-face interaction afforded by traditional therapeutic techniques.

We have identified 32 existing interactive Web-based programs and have found varying degrees of accessibility, quality, and evidence supporting their efficacy. Only 12 of the 32 programs had at least one peer-reviewed, published article describing the results of an efficacy study. In examining the programs, authors

noted a number of similarities and differences when comparing EBP to NBP by each evaluation point (see Multimedia Appendix 3). Using the adapted evaluation criteria [35], each program was rated using 28 categories. Totals for a few EBP and NBP were amended if a score could not be provided for a program in the category, as indicated by a superscript "a" ([a]) (ie, the information was not available on the program website and/or inquiries to program developers were not returned). EBP had a smaller range in scores (68%-92%) and a higher overall average score (80%) relative to the wider range for NBP (39%-92%) and a lower overall average score (73%). While generally comparable, there were some subtle differences between EBP and NBP programs on the evaluation criteria of accessibility, program structure, use of validated assessment tools, additional features, peer support forums, safeguards, and the provision of user-statistics as further discussed below.

While Web-based programs can generally improve access to mental health care, some aspects of existing programs may present users with alternative access barriers. For example, some programs were available only in one language (eg, Interapy and Kleur je Leven in Dutch, and Internetpsykiatri in Swedish), with the majority of NBP available only in English. The addition of alternative languages could promote open accessibility to any user seeking treatment. Furthermore, some programs had accessibility restrictions based on country of residence (see Multimedia Appendix 2), which may have been due to limited therapist availability for those programs providing clinician support. Given the global nature of the Internet, consideration could be made to avoid geographic restrictions on program availability.

While a higher proportion of the NBP (9/19[a]; 47%) were freely available (no fees or referrals required to access) as compared to the EBP (3/10[a]; 30%), a higher proportion of the NBP (6/19[a]; 32%) than the EBP (3/10[a]; 30%) had a fee for accessing the program. Requirements such as therapist referral, administrator acceptance for registration, and/or user fees may act as deterrents to use as they necessitate additional motivation and resources on the part of the user. The need to obtain a referral, enter personal information to register, and/or wait for access negates the benefits of anonymity and convenience afforded by Web-based tools. Conversely, registration requirements enable the user to track their progress and build on previously completed modules. In addition, the registration of personal information would allow the program deliverers to contact the user when in need, such as when increases in depression symptoms or suicide risk (which are more prevalent among depressed individuals as compared to other mental health disorders) are reported. Also, referral-based programs often allowed for integrated therapist contact. For fee-based programs, free demonstration/trial modules could be provided to allow users to assess the program prior to making a financial commitment.

In examining the programs captured in the review, CBT was the most commonly incorporated therapeutic approach (EBP=6, NBP=9; see Table 3). Other therapeutic techniques included psychoeducation, IPT, positive psychology, and narrative therapy. Recent evidence suggests that Web-delivered, self-guided IPT is as effective as Web-delivered CBT on symptoms of depression in general community samples [39]. Future investigations could continue to examine which techniques are most efficiently and effectively delivered through the Web and whether the mode of delivery (eg, text, audio, video) differentially affects outcome. In regards to program content and structure (evaluation point 11, ie, guided [modules must be completed in a set order] vs unguided [any module is accessible at any time]), 80% (8/10[a]) of EBP offered the guided approach relative to 63% (12/19[a]) of the NBP. Future studies should evaluate the relative effectiveness of guided versus unguided interventions.

Many of the programs captured in this review delivered treatment specifically for adolescents, adults, or both. A limited number of programs catered to special populations (eg, military personnel, older adults). Future programs could be geared toward the needs of special populations such as individuals with cognitive impairments or persons in a caregiving role. Accommodations for cognitively impaired individuals may include larger text sizing, multimodal delivery (audio and video files), in addition to programming specific to their impairments (eg, memory games, goal setting and problem solving). Caregivers of chronically ill patients also demonstrate increased psychological distress and burden [40,41]. Apart from depression, increased caregiver anxiety, guilt, rage, grief, substance abuse, and elevated risk of relapsing into a pre-existing mood disorder has been noted among caregivers [39,42,43]. Future programs could be tailored to meet the issues faced by caregivers in day-to-day life, including an emphasis on preserving and increasing available family resources in caregiving circumstances [44]. Program developers should use feedback from members of these populations to fully understand learning preferences as well as accessibility issues specific to these individuals.

Although a target audience was identified for each of the programs (ie, adults, students, special population), course content was often generic for all users within the targeted population. Personalized treatment plans (ie, generic programs vs individual treatment plans; evaluation point 12) were offered in only a few programs: 22% of EBP (2/9[a]) and NBP (4/18[a]). Personalized treatment plans may enhance user engagement by appealing to their specific treatment needs and offering relevant treatment information. Take for example a user suffering from only a mild form of depression; they may not have found additional anxiety information useful causing them to lose interest in continuing with the program despite its potential benefit. In programs offering a personalized treatment plan, program suggestions were based on an initial assessment. All but a few EBP and NBP offered assessments; however, not all of the assessment tools used were validated (evaluation point 15). Of EBP, 75% (6/8[a]) relative to 67% (12/18[a]) of NBP employed validated assessment tools (ie, BDI, PHQ-9, or CES-D). Programs should strive to offer validated assessment tools to provide users with accurate feedback in regards to their depressive and anxiety symptoms. During this emotionally sensitive time period, individuals could be heavily influenced

by program feedback and results, necessitating accurate and valid depictions of depression symptoms over time.

In addition to worksheets and assessments, some programs offered additional features that may help enhance usage and retention including emails offering encouragement, helpful quotes or testimonials, and reminders to complete modules; completion trackers for each session and/or the program overall; supplementary worksheets and mood assessments delivered during or after each session to assess and monitor progress; and automated feedback to the user. The majority of both EBP (7/10[a]; 70%) and NBP (10/20; 50%) offered three or more of these additional features (see Table 4). The effect of these features on user satisfaction and treatment efficacy should be further investigated.

Many of the programs included additional integrated therapist contact, peer support discussion forums, and crisis links. Programs offering therapist support (evaluation point 21) were delivered via telephone, video conference, or live chat (ie, instant messaging): 60% (6/10[a]) of EBP and 50% (8/19[a]) of NBP offered therapist support. Similarly, 60% (6/10[a]) of EBP and 50% (8/19[a]) of NBP provided a therapist name and their credentials (evaluation point 22). Providing users with this information may provide them peace of mind that they are being cared for and monitored by an accredited individual capable of intervening if required. A recent study using MoodGYM plus brief face-to-face therapist support indicated positive results in the reduction of symptoms of depression in a primary care setting [45]. Although additional therapist support has been shown to be effective [45,46], further investigation is needed to understand therapist user interaction impact on patient outreach, treatment experience, and concept reinforcement. Conversely, efficacy has also been demonstrated in programs that did not include clinician support [47,48]. When choosing an intervention program, this option is influenced by user learning needs and should be indicated by the user. Facilitating therapist-user communication through technology could aid in maintaining user anonymity and privacy. Potential negative implications of therapist contact include reduced availability, reduced user independence, and increased pressure to complete program requirements.

Other avenues of support offered within the evaluated programs included peer discussion forums, blogs, and shared user spaces (evaluation point 23). Unlike EBP, none of which provided peer-support forums, 44% (8/18[a]) of NBP offered this feature. However, only 24% (4/17[a]) of NBP offered forums that were monitored by an overseeing authority, facilitating safe user interaction and positive constructive topics of conversation (evaluation point 24). Peer support offers a level of familiarity not offered with clinician support. The need for relatedness to others enduring similar emotional issues can be both comforting and motivating; however, the effectiveness of peer support upon symptom resolution has yet to be evaluated in this context.

In addition to peer and clinician support, some programs offered crisis links via telephone hotlines, email, or chat functions. Hotlines provided support to users under distress when therapists or other social support options were unavailable. Safeguards (evaluation point 17) were available in all the evaluated EBP; however, only 68% (13/19[a]) of NBP offered this feature. Due to the sensitive nature of the treatment and topics discussed, all programs should offer or provide information for available crisis links.

To evaluate program usability, we contacted each of the program's administrators. Many were unable or did not wish to disclose user statistics in regards to registration, attrition, and program completion. Those that responded to inquiry emails or posted statistics on their program websites are listed in Table 3. Among the few programs providing data (EBP, 7/10[a] or 70%; NBP, 10/19[a] or 53%), statistics varied greatly. Examining the existing literature, completion rates in RCTs of Web-delivered treatments for depression have primarily ranged from 55%-67% [49-54] with rates as low as 20% [55]. Future research should examine which aspects of a program could promote retention and completion such as email reminders or motivators/incentives. End-user feedback may be useful in identifying less effective areas within programs and facilitate modification.

Although not included in this review (as they did not meet the inclusion and exclusion criteria), three novel and noteworthy Web-based treatment programs for depression were identified: Depression Quest [56], Moodbuster [57], and SPARX [58]. Depression Quest and SPARX are gaming programs that deliver innovative depression treatments to adolescents. Depression Quest invites the user to experience life from the perspective of a depressed individual. Users read scenarios, select one of the decisions provided, and navigate the path associated with their choice. Program developers aimed to (1) aid caregivers by providing insight into the depressed mindset, and (2) demonstrate to individuals affected by depression that they are not alone in their struggle. In SPARX, users navigate an avatar of their choice towards an end objective while fighting gloomy negative automatic thoughts along the way. Targeting depression, anxiety, and stress, clinical trials of the program have demonstrated positive results on symptoms of depression [59]. The third program, Moodbuster, provides an interactive treatment program similar to those included in the review; however, it also incorporates the use of biosensors worn on the body throughout the day. Biosensors are equipped to transmit information on emotionally influenced bodily responses like electrodermal activity, respiration, and electrocardiography changes. A monitoring system for medication intake is also provided to users prescribed pharmacological treatment. Sensors are set to monitor dose and intake information. With all sensors feeding back to the program, Moodbuster interprets the information and reasons which type of therapy is most likely to be effective. Although resource intensive, a program like Moodbuster may be effective for depression resistant to alternative forms of treatment; however, research is needed to further evaluate this treatment approach.

In summary, many interactive treatment programs for depression are available on the Web; however, the efficacy and validity of most of these programs (20/32, 63%) have not been evaluated using RCTs. When comparing those programs that are evidence-based to those that have not been empirically

evaluated, more of the EBP programs seemed to use a guided approach, employ validated assessment tools, offer additional features, incorporate safeguards, and provide user statistics. More of the NBP programs were available without fees or referrals (however, a higher proportion did request a user fee than the EBP) and offered peer-support forums. Based on our review, several programs emerged that are easily accessible, free to use, and have supporting evidence for their efficacy including E Couch, MoodGYM, and This Way Up (Self Help Course, Worry and Sadness; see Figure 11-13 for screenshots of these programs).

Although there is a strong and growing body of evidence in support of Web-based interventions, some perceive that the uptake and dissemination of such programs have not been commensurate with their potential to improve health-related outcomes. With respect to Web-based interventions for depression, potential barriers have been cited such as negative clinician and patient attitudes [60,61], legal and ethical regulations related to online clinician-patient interactions [62], a lack of practitioner willingness to refer patients to such interventions, and clinician fears of losing work [63]. In order to increase implementation and reach across a range of settings, Burnett and Glasgow have suggested using tailored messaging and social networking functionality, in particular leveraging newer technologies that offer novel ways for users to store, view, manipulate, share, and experience their personal data (eg, Web 2.0 design principles) [64].

Figure 11. Mood Gym introduction page.

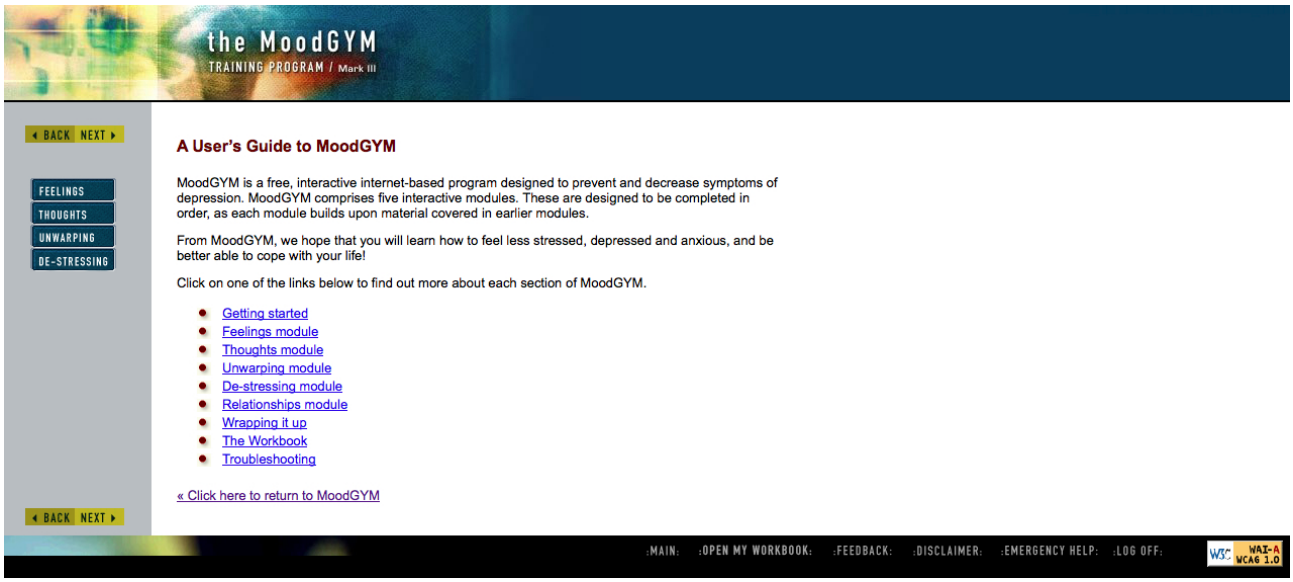


Figure 12. E Couch navigation page.

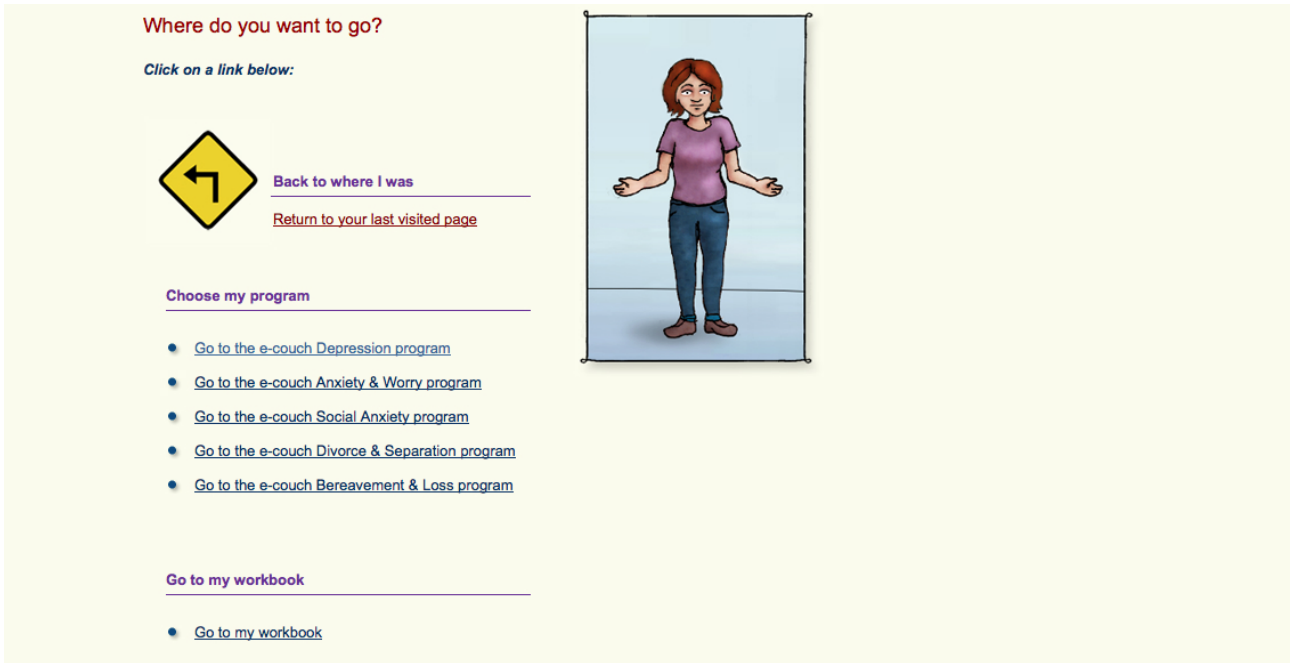


Figure 13. This Way Up welcome page.

Limitations

While we aimed to be comprehensive, systematic, and thorough in our review and evaluation, some information could not be ascertained due to referral and cost restrictions limiting program accessibility. We used the Canadian version of common search engines (Google, Yahoo, and Bing) to identify programs conforming to evaluation inclusion criteria. Searches were repeated until results became redundant; however, it is possible that despite our efforts to be inclusive, some programs were missed. It is also important to note that searches conducted in other countries (using the above mentioned search engines) may not yield the same search results. Provided that many search engines suggest local websites and may also receive funding from local advertisers, our search results may be highly specific to our proximity (Toronto, Canada). Moreover, the rapidly evolving nature of the Internet means that the same search conducted several months from now could yield a different set of results and conclusions. New programs may be developed, existing programs may be discontinued, adapted, or amalgamated, and new research trials could support or refute the efficacy of these programs.

It should also be noted that programs were evaluated from the perspective of a researcher and not of an end-user. A person with depression, for example, may have different opinions regarding the usability and quality of the identified programs. The intended end-user would likely provide developers with a more subjective opinion. Programs may also be biased to the type of user they capture, that is, individuals with less severe forms of depression. Individuals suffering from more severe forms of depression may face greater decreases in motivation and are less likely to access and participate in an intervention program. Consequently, users with severe forms of depression may be underrepresented in RCTs and in collected user feedback and data. Feedback is necessary to facilitate change as well as outline likes and dislikes for various features. Overall, end-user input should be sought as it is crucial to improving treatment delivery and program functionality.

Future Directions

At this time, there are few programs available for special populations (eg, caregivers, individuals with cognitive deficits, older adults). It is important that program functionality accommodate accessibility of varying populations to ensure adequate treatment is delivered. Studies investigating the needs of these special populations could inform the development of new programs and the adaptation of existing ones. Future programs could also aim for increased accessibility. This could entail multilanguage delivery, elimination of residency restrictions, elimination of registration fees or referrals, flexibility in module timing, and minimization of mandatory user response (eg, mandatory worksheets and mood assessments).

With respect to research, although the minority of the identified programs were evidence-based as defined by the presence of at least one evaluative RCT, the trials that have been completed to date have been generally of strong quality with adequate sample sizes [32-34]. In addition to testing the efficacy of these programs, future research should examine the barriers, facilitators, and modifiers to program uptake, compliance, and retention. The impact of aspects such as program structure and content, registration requirements, user fees, therapist support, and specific features (eg, reminders, feedback, peer-support groups) on program efficacy, compliance, retention, safety, and cost-effectiveness require further investigation. In addition, factors related to treatment response sustainability need to be examined.

Conclusions

In this review, we identified 27 websites offering 32 programs with interactive components aimed at reducing symptoms of depression among users. The programs varied widely in terms of content, accessibility, usability, method of delivery (eg, text,

audio, or video), and supplementary tools. A minority of the programs identified had empiric evidence to support their efficacy for the treatment of depressive symptoms.

In choosing to use or refer a Web-based treatment program for depression, the user may wish to consider the following factors: ease of use and accessibility, availability of additional features and support needs, and most critically, programs that have been validated with good quality research. Users are encouraged to critically evaluate their program choice and should investigate research supporting program claims.

Web-delivered interventions afford many potential advantages to individuals. Users can log on to their preferred program in the comfort of their own home 24 hours a day, 365 days a year. This may help to increase accessibility, reduce prolonged wait times, and address privacy concerns. Furthermore, it is potentially cost-efficient and convenient, allowing users to seek treatment when desired. Developers should continue to create such programs and tailor additional sites to the needs of specialty groups.

Acknowledgments

Authors wish to thank Karthik Natarajan for his additional support conducting searches and program evaluations. We also wish to thank Herman Tang for generously providing us a cover graphic reflective of the manuscript topic.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search log.

[XLSX File (Microsoft Excel File), 481KB - jmir_v16i9e209_app1.xlsx]

Multimedia Appendix 2

Accessibility of the programs evaluated.

[XLSX File (Microsoft Excel File), 491KB - jmir_v16i9e209_app2.xlsx]

Multimedia Appendix 3

Program evaluation log.

[XLSX File (Microsoft Excel File), 494KB - jmir_v16i9e209_app3.xlsx]

Multimedia Appendix 4

Randomized controlled trial citations for evidence-based programs.

[DOCX File, 486KB - jmir_v16i9e209_app4.docx]

Multimedia Appendix 5

Program screenshots.

[PDF File (Adobe PDF File), 13MB - jmir_v16i9e209_app5.pdf]

References

1. World Health Organization. Depression Fact Sheet. 2012. URL: http://www.who.int/mediacentre/factsheets/fs369/en/index.html [accessed 2014-08-22] [WebCite Cache ID 6S1VCE9GM]
2. Public Health Agency of Canada. The human face of mental health and mental illness in Canada. 2006. URL: http://www.phac-aspc.gc.ca/cd-mc/mi-mm/depression-eng.php [accessed 2014-08-22] [WebCite Cache ID 6S1VYWYhT]
3. Dikmen SS, Bombardier CH, Machamer JE, Fann JR, Temkin NR. Natural history of depression in traumatic brain injury. Arch Phys Med Rehabil 2004 Sep;85(9):1457-1464. [Medline: 15375816]
4. Holsinger T, Steffens DC, Phillips C, Helms MJ, Havlik RJ, Breitner JC, et al. Head injury in early adulthood and the lifetime risk of depression. Arch Gen Psychiatry 2002 Jan;59(1):17-22. [Medline: 11779276]
5. Hesdorffer DC, Rauch SL, Tamminga CA. Long-term psychiatric outcomes following traumatic brain injury: a review of the literature. J Head Trauma Rehabil 2009;24(6):452-459. [doi: 10.1097/HTR.0b013e3181c133fd] [Medline: 19940678]
6. Jorge RE, Robinson RG, Moser D, Tateno A, Crespo-Facorro B, Arndt S. Major depression following traumatic brain injury. Arch Gen Psychiatry 2004 Jan;61(1):42-50. [doi: 10.1001/archpsyc.61.1.42] [Medline: 14706943]

7. Hibbard MR, Ashman TA, Spielman LA, Chun D, Charatz HJ, Melvin S. Relationship between depression and psychosocial functioning after traumatic brain injury. Arch Phys Med Rehabil 2004 Apr;85(4 Suppl 2):S43-S53. [Medline: 15083421]
8. Rapoport MJ, McCullagh S, Streiner D, Feinstein A. Age and major depression after mild traumatic brain injury. Am J Geriatr Psychiatry 2003;11(3):365-369. [Medline: 12724117]
9. Seel RT, Kreutzer JS. Depression assessment after traumatic brain injury: an empirically based classification method. Arch Phys Med Rehabil 2003 Nov;84(11):1621-1628. [Medline: 14639561]
10. Seel RT, Kreutzer JS, Rosenthal M, Hammond FM, Corrigan JD, Black K. Depression after traumatic brain injury: a National Institute on Disability and Rehabilitation Research Model Systems multicenter investigation. Arch Phys Med Rehabil 2003 Feb;84(2):177-184. [doi: 10.1053/apmr.2003.50106] [Medline: 12601647]
11. Bombardier CH, Fann JR, Temkin NR, Esselman PC, Barber J, Dikmen SS. Rates of major depressive disorder and clinical outcomes following traumatic brain injury. JAMA 2010 May 19;303(19):1938-1945 [FREE Full text] [doi: 10.1001/jama.2010.599] [Medline: 20483970]
12. Dobson K. Canadian Psychological Association. 2009. "Psychology Works" fact sheet: depression URL: http://www.cpa.ca/docs/file/publications/factsheets/psychologyworksfactsheet_depression.pdf [accessed 2014-08-26] [WebCite Cache ID 6S7yzqrU6]
13. Centre for Addiction and Mental Health. Statistics on mental health and addiction. 2012. URL: http://www.camh.ca/en/hospital/about_camh/newsroom/for_reporters/pages/addictionmentalhealthstatistics.aspx [accessed 2014-08-26] [WebCite Cache ID 6S7zbdLOX]
14. Statistics Canada. Canadian community health survey: mental health and well-being. 2003. URL: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SurvId=20251&SurvVer=1&InstaId=20892&InstaVer=1&SDDS=5015&lang=en&db=imdb&adm=8&dis=2
15. Canadian Mental Health Association. The effects of depression and anxiety on Canadian society. 2002. URL: http://www.cmha.ca/public_policy/the-effects-of-depression-and-anxiety-on-canadian-society/ [accessed 2013-08-06] [WebCite Cache ID 6IfkV8MnQ]
16. International Telecommunication Union. ITU World Telecommunication/ICT Indicators database. 2011. The world in 2011: ICT facts and figures URL: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2011.pdf [accessed 2013-08-06] [WebCite Cache ID 6IfkWcYsR]
17. Dölemeyer R, Tietjen A, Kersting A, Wagner B. Internet-based interventions for eating disorders in adults: a systematic review. BMC Psychiatry 2013 Aug;13:207 [FREE Full text] [doi: 10.1186/1471-244X-13-207] [Medline: 23919625]
18. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
19. Coons MJ, Demott A, Buscemi J, Duncan JM, Pellegrini CA, Steglitz J, et al. Technology interventions to curb obesity: a systematic review of the current literature. Curr Cardiovasc Risk Rep 2012 Apr;6(2):120-134 [FREE Full text] [doi: 10.1007/s12170-012-0222-8] [Medline: 23082235]
20. Miranda J, Côté J, Godin G, Blais M, Otis J, Guéhéneuc YG, et al. An internet-based intervention (condom-him) to increase condom use among HIV-positive men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2013 Oct;2(2):e39 [FREE Full text] [doi: 10.2196/resprot.2723] [Medline: 24132072]
21. Grynszpan O, Weiss PL, Perez-Diaz F, Gal E. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. Autism 2014 May;18(4):346-361. [doi: 10.1177/1362361313476767] [Medline: 24092843]
22. Marsch LA, Carroll KM, Kiluk BD. Technology-based interventions for the treatment and recovery management of substance use disorders: a JSAT special issue. J Subst Abuse Treat 2014 Jan;46(1):1-4. [doi: 10.1016/j.jsat.2013.08.010] [Medline: 24041749]
23. Connelly J, Kirk A, Masthoff J, MacRury S. The use of technology to promote physical activity in Type 2 diabetes management: a systematic review. Diabet Med 2013 Dec;30(12):1420-1432. [doi: 10.1111/dme.12289] [Medline: 23870009]
24. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
25. Tait RJ, Spijkerman R, Riper H. Internet and computer based interventions for cannabis use: a meta-analysis. Drug Alcohol Depend 2013 Dec 1;133(2):295-304. [doi: 10.1016/j.drugalcdep.2013.05.012] [Medline: 23747236]
26. Voncken-Brewster V, Tange H, de Vries H, Nagykaldi Z, Winkens B, van der Weijden T. A randomised controlled trial testing a web-based, computer-tailored self-management intervention for people with or at risk for chronic obstructive pulmonary disease: a study protocol. BMC Public Health 2013 Jun;13:557 [FREE Full text] [doi: 10.1186/1471-2458-13-557] [Medline: 23742208]
27. 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]
28. McDermott MS, While AE. Maximizing the healthcare environment: a systematic review exploring the potential of computer technology to promote self-management of chronic illness in healthcare settings. Patient Educ Couns 2013 Jul;92(1):13-22. [doi: 10.1016/j.pec.2013.02.014] [Medline: 23566427]

29. Parikh SV, Segal ZV, Grigoriadis S, Ravindran AV, Kennedy SH, Lam RW, Canadian Network for MoodAnxiety Treatments (CANMAT). Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. II. Psychotherapy alone or in combination with antidepressant medication. J Affect Disord 2009 Oct;117 Suppl 1:S15-S25. [doi: 10.1016/j.jad.2009.06.042] [Medline: 19682749]
30. Clark, DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. International Review of Psychiatry 2011 Aug;23:375-384 [FREE Full text] [Medline: 22026487]
31. Centre for Mental Health Research at the Australian National University. Beacon 2.0. Depression services URL: https://beacon.anu.edu.au/ [accessed 2014-05-20] [WebCite Cache ID 6Pi8xvRnw]
32. 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]
33. 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]
34. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. Med J Aust 2010 Jun 7;192(11 Suppl):S4-11. [Medline: 20528707]
35. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting internet intervention research. Cogn Behav Ther 2011 Jun;40(2):82-97. [doi: 10.1080/16506073.2011.573807] [Medline: 25155812]
36. Mooney G, Speed J. The association between mild traumatic brain injury and psychiatric conditions. Brain Inj 2001 Oct;15(10):865-877. [doi: 10.1080/02699050110065286] [Medline: 11595083]
37. Mooney G, Speed J, Sheppard S. Factors related to recovery after mild traumatic brain injury. Brain Inj 2005 Nov;19(12):975-987. [doi: 10.1080/02699050500110264] [Medline: 16263640]
38. McAllister TW, Arciniegas D. Evaluation and treatment of postconcussive symptoms. NeuroRehabilitation 2002;17(4):265-283. [Medline: 12547976]
39. George LK, Gwyther LP. Caregiver well-being: a multidimensional examination of family caregivers of demented adults. Gerontologist 1986 Jun;26(3):253-259. [Medline: 3721232]
40. Thomsen IV. Late outcome of very severe blunt head trauma: a 10-15 year second follow-up. J Neurol Neurosurg Psychiatry 1984 Mar;47(3):260-268 [FREE Full text] [Medline: 6707671]
41. Dura JR, Stukenberg KW, Kiecolt-Glaser JK. Anxiety and depressive disorders in adult children caring for demented parents. Psychol Aging 1991 Sep;6(3):467-473. [Medline: 1930763]
42. Whitlatch CJ, Zarit SH, von Eye A. Efficacy of interventions with caregivers: a reanalysis. Gerontologist 1991 Feb;31(1):9-14. [Medline: 2007480]
43. Schulz R, O'Brien AT, Bookwala J, Fleissner K. Psychiatric and physical morbidity effects of dementia caregiving: prevalence, correlates, and causes. Gerontologist 1995 Dec;35(6):771-791. [Medline: 8557205]
44. Anderson MI, Parmenter TR, Mok M. The relationship between neurobehavioural problems of severe traumatic brain injury (TBI), family functioning and the psychological well-being of the spouse/caregiver: path model analysis. Brain Inj 2002 Sep;16(9):743-757. [doi: 10.1080/02699050210128906] [Medline: 12217201]
45. Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, et al. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. J Med Internet Res 2013;15(8):e153 [FREE Full text] [doi: 10.2196/jmir.2714] [Medline: 23916965]
46. Donker T, Bennett K, Bennett A, Mackinnon A, van Straten A, Cuijpers P, et al. Internet-delivered interpersonal psychotherapy versus internet-delivered cognitive behavioral therapy for adults with depressive symptoms: randomized controlled noninferiority trial. J Med Internet Res 2013;15(5):e82 [FREE Full text] [doi: 10.2196/jmir.2307] [Medline: 23669884]
47. 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]
48. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
49. 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]
50. 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 Nov;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094]
51. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005 Jun;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]

52. 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]
53. 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]
54. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009 Jun;43(6):571-578. [doi: 10.1080/00048670902873722] [Medline: 19440890]
55. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
56. Quinn Z. Depression Quest. 2013. URL: http://www.depressionquest.com/ [accessed 2014-08-26] [WebCite Cache ID 6S84YipWi]
57. VU University Amsterdam. Moodbuster. The Netherlands ICT4 Depression URL: http://www.ict4depression.eu/ [accessed 2014-08-26] [WebCite Cache ID 6S857mmHC]
58. The University of Auckland. SPARX. 2012. URL: http://linkedwellness.com/play-sparx- [WebCite Cache ID 6pxlv65e9]
59. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012;344:e2598 [FREE Full text] [Medline: 22517917]
60. Stallard P, Richardson T, Velleman S. Clinicians' attitudes towards the use of computerized cognitive behaviour therapy (cCBT) with children and adolescents. Behav Cogn Psychother 2010 Oct;38(5):545-560. [doi: 10.1017/S1352465810000421] [Medline: 20615273]
61. Mohr DC, Siddique J, Ho J, Duffecy J, Jin L, Fokuo JK. Interest in behavioral and psychological treatments delivered face-to-face, by telephone, and by internet. Ann Behav Med 2010 Aug;40(1):89-98 [FREE Full text] [doi: 10.1007/s12160-010-9203-7] [Medline: 20652466]
62. Dever Fitzgerald T, Hunter PV, Hadjistavropoulos T, Koocher GP. Ethical and legal considerations for internet-based psychotherapy. Cogn Behav Ther 2010;39(3):173-187. [doi: 10.1080/16506071003636046] [Medline: 20485997]
63. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
64. 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]

Abbreviations

CBT: cognitive behavioral therapy
EBP: evidence-based program
IPT: interpersonal therapy
NBP: non-evidence-based program
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 04.12.13; peer-reviewed by T Donker, S Patten; comments to author 14.03.14; revised version received 22.05.14; accepted 10.07.14; published 23.09.14.

Please cite as:
Renton T, Tang H, Ennis N, Cusimano MD, Bhalerao S, Schweizer TA, Topolovec-Vranic J
Web-Based Intervention Programs for Depression: A Scoping Review and Evaluation
J Med Internet Res 2014;16(9):e209
URL: http://www.jmir.org/2014/9/e209/
doi:10.2196/jmir.3147
PMID:25249003

Original Paper

Effects of a Guided Web-Based Smoking Cessation Program With Telephone Counseling: A Cluster Randomized Controlled Trial

Michael Mehring[1], MD; Max Haag[1]; Klaus Linde[1], MD; Stefan Wagenpfeil[2], PhD; Antonius Schneider[1], MD

[1]Institute of General Practice, Klinikum rechts der Isar, Technische Universität München, Munich, Germany

[2]Institute for Medical Biometry, Epidemiology und Medical Informatics (IMBEI), Universitätsklinikum des Saarlandes, Homburg/Saar, Homburg, Germany

Corresponding Author:
Michael Mehring, MD
Institute of General Practice
Klinikum rechts der Isar
Technische Universität München
Orleansstr. 47
Munich, 81667
Germany
Phone: 49 614658914
Fax: 49 614658915
Email: michael.mehring@tum.de

Abstract

Background: Preliminary findings suggest that Web-based interventions may be effective in achieving significant smoking cessation. To date, very few findings are available for primary care patients, and especially for the involvement of general practitioners.

Objective: Our goal was to examine the short-term effectiveness of a fully automated Web-based coaching program in combination with accompanied telephone counseling in smoking cessation in a primary care setting.

Methods: The study was an unblinded cluster-randomized trial with an observation period of 12 weeks. Individuals recruited by general practitioners randomized to the intervention group participated in a Web-based coaching program based on education, motivation, exercise guidance, daily short message service (SMS) reminding, weekly feedback through Internet, and active monitoring by general practitioners. All components of the program are fully automated. Participants in the control group received usual care and advice from their practitioner without the Web-based coaching program. The main outcome was the biochemically confirmed smoking status after 12 weeks.

Results: We recruited 168 participants (86 intervention group, 82 control group) into the study. For 51 participants from the intervention group and 70 participants from the control group, follow-up data were available both at baseline and 12 weeks. Very few patients (9.8%, 5/51) from the intervention group and from the control group (8.6%, 6/70) successfully managed smoking cessation (OR 0.86, 95% CI 0.25-3.0; *P*=.816). Similar results were found within the intent-to-treat analysis: 5.8% (5/86) of the intervention group and 7.3% (6/82) of the control group (OR 1.28, 95% CI 0.38-4.36; *P*=.694). The number of smoked cigarettes per day decreased on average by 9.3 in the intervention group and by 6.6 in the control group (2.7 mean difference; 95% CI -5.33 to -0.58; *P*=.045). After adjustment for the baseline value, age, gender, and height, this significance decreases (mean difference 2.2; 95% CI -4.7 to 0.3; *P*=.080).

Conclusions: This trial did not show that the tested Web-based intervention was effective for achieving smoking cessation compared to usual care. The limited statistical power and the high drop-out rate may have reduced the study's ability to detect significant differences between the groups. Further randomized controlled trials are needed in larger populations and to investigate the long-term outcome.

Trial Registration: German Register for Clinical Trials, registration number DRKS00003067; http://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ ID=DRKS00003067 (Archived by WebCite at http://www.webcitation.org/6Sff1YZpx).

(J Med Internet Res 2014;16(9):e218) doi:10.2196/jmir.3536

KEYWORDS

smoking cessation; Web-based; randomized controlled trial; primary care

Introduction

Tobacco smoking is a major preventable cause of death worldwide. The use of tobacco is estimated to kill 5.4 million people a year. By 2030, tobacco will contribute to the deaths of more than 8 million people a year. The overwhelming majority of those deaths are predicted to occur in the developing world [1]. Tobacco smokers are more prone to develop cancer and are also at substantially increased risk of developing heart disease, stroke, emphysema, and other severe diseases [2]. Prevention and cessation are the two principal strategies against tobacco smoking. Surveys are already indicating that almost 70% of smokers would like to stop smoking completely [3,4]. There is evidence that cessation advice given by a doctor is an efficient way to support smokers to quit and that more intensive interventions in general practice increases the abstinence rate [5]. Especially for younger smokers, the Internet may provide a vehicle to support this approach [6]. There were 2.4 billion Internet users worldwide with an increasing trend in 2012 [7]. The use of Web-based smoking cessation material provides low costs per user and results in high cost-effectiveness [8]. Web-based programs are convenient for users because the content can be accessed easily anytime and anywhere. For some people, the greater level of anonymity online than in in-person counseling may be appealing.

The evidence of a variety of computer and other electronic aids were summarized in a meta-analysis by Chen et al [9]. They concluded that computer and other electronic aids increase—to a small extent—the likelihood of prolonged smoking cessation compared with no intervention. A recently published Cochrane review [10] came to the conclusion that some Web-based interventions can assist smoking cessation, particularly those that are interactive and tailored to individuals. However, there were no consistent effects detected from trials that compared Internet interventions with usual care. Likewise, most of the included trials relied only on self-reported smoking status. Biochemical validation of self-reported cessation was only attempted in six of 28 trials.

Therefore, the aim of the present study was to compare the short-term effectiveness of a Web-based coaching program in combination with telephone counseling to usual care (see Multimedia Appendix 1). The Web-based program we tested, HausMed, combines an individually tailored strategy for smoking cessation with automated advice and feedback elements, in addition to monitoring via Internet and telephone counseling in general practice. Such a tool would facilitate the management of patients as they receive support from their general practitioner (GP) during the primary care process guided by the Web-based program. To date, very few findings within primary care patients and involving general practitioners are available. Additionally, the biochemical validation of the self-reported cessation status by a cotinine urine test was implemented in the present investigation to confirm the documented main outcome.

Methods

Design

The study was designed as a two-armed, unblinded cluster-randomized controlled trial. At the beginning of the study, around 2000 Bavarian general practitioners (GPs) received a fax by the Bavarian Association of General Practitioners with information about the research project. All interested GPs were sequentially registered for randomization. After giving written consent, the participating practices were randomized either to the intervention or the control arm. The sequence of randomization (allocation 1:1) was provided by a methodologist, who did not participate in the execution of the study, via the program Research Randomizer [11]. Randomization was concealed by using sequentially numbered, opaque, sealed envelopes held by the study coordinator. Before starting the recruitment of patients, physicians and practice nurses received detailed instructions by the research team on the study process (both intervention and control group) and on the coaching program (only intervention group).

Physicians assigned to the control arm were asked to change nothing in their usual way of counseling and to treat participants in the same manner as if they would have been non-participants. There was no structured documentation of the care provided. The patients recruited by intervention practices received free access to the Web-based coaching program. The patients of practices participating in the control arm were advised by the GPs in their individual way of usual care to quit smoking. The study was approved by the Medical Ethics Committee of the Technische Universität München (April 19, 2011) and was in accordance with ethical standards for human experimentation established by the Declaration of Helsinki. All participants gave written informed consent. A data and safety monitoring board was established before the beginning of the study.

Participants and Procedures

Participating physicians were general practitioners in Bavaria, Germany. The GPs were requested to recruit individuals with the desire for smoking cessation (see Multimedia Appendix 2). Individuals at least 18 years old and with Internet access were potentially eligible. Exclusion criteria were aged younger than 18 years, insufficient German language skills, and lack of Internet access. Further exclusion criteria were psychiatric disorders and posttraumatic stress disorder.

After the GP decided that patient should participate, an information form was given and discussed with the patients, and a participation form had to be signed. At the same time, the baseline data acquisition took place. All participants were asked to fill in a standardized questionnaire together with the GP. The standardized questionnaire comprised the following information: age, sex, height, weight, physical activity, years of tobacco use, number of cigarettes smoked per day, number of previous quit attempts, use of current nicotine replacement therapy (NRT), and reasons for smoking. Participants of the intervention group

received a free Web-code. The physician filled in a form together with the patient with information about the potential existence and grade of a chronic obstructive pulmonary disease. This form with the Internet code was used by the patient for specification during the registration process of the Internet program. Participants of both groups were requested after 12 weeks to document the follow-up evaluation together with their physician. The follow-up comprised again information about smoking status, weight, physical activity, number of cigarettes smoked per day, and use of current NRT. At follow-up, a biochemical validation of the self-reported cessation status was also implemented through a cotinine urine test. Cotinine is detected in the urine for 2-4 days after the use of tobacco.

Physicians in the intervention group received €50 per participant for time and effort. Physicians in the control group received €25. Participants in the intervention group received free access to the smoking cessation program, which usually costs €79. Participants in the control group received €10 as an incentive to come into practice for follow-up investigation after 12 weeks. No methodical changes were made during the entire study period.

Intervention

Anamnestic and health data were documented in a structured registration form including information about the potential existence and grade of a chronic obstructive pulmonary disease from the GP. The patient received a copy of this form in order to use the health data for subscribing via Internet into the HausMed coaching program [12] at home. A specific program was installed to allow the participants to log in without charge. After completion of a pre-assessment, the program generated individual coaching based on the given information of the physicians (registration form), the physical characteristics, and the everyday behavior of the participants.

The coaching program is based on the generally accepted principles of cognitive behavioral therapy and combined psychoeducation and motivational techniques with behavioral-therapeutic elements [13]; for example, education, realistic goal-setting, and individual resources, and in particular, the behavioral change theory targeted to smoking cessation by using inexpensive Internet and mobile technologies in combination with existing health care resources of GPs. The content of the coaching program aimed at achieving a lasting change of behavioral patterns with the help of individualized education, motivation, exercise guidance, daily SMS text messaging (short message service, SMS) reminders, self-monitoring via Internet and, finally, through active monitoring and approximately three telephone calls during the 12 weeks by the GPs or their staff. The framework of the program is based on the idea by Buchkremer and Batra, of the Department of Psychiatry and Psychotherapy, University of Tübingen, Germany [14,15]. The development and implementation of the Web-based smoking cessation program was carried out by WeCARE GmbH, Göttingen, Germany. The coaching program was subdivided into 12 different constitutive modules. Each module was performed for one week and contains particular tasks (Textbox 1), which were supported by corresponding daily SMS reminders. The participant had to perform a specific task each day and received a daily SMS regarding that specific task.

The reminder contained adapted information to maintain motivation, impart daily tips, and encourage daily performance of the respective task. The specific daily tasks were offered on the first day of each module. The coaching program also offered a variety of printed material (emergency plan, relaxation exercises, questionnaires, information, self-agreements, etc), which was connected to the respective task and included interactive buttons, video clips, and learning progress quizzes to examine learning success (Figures 1 and 2).

Textbox 1. Goals of the Web-based coaching program exercises.

- Your path to a life without tobacco
- Get down to brass tacks
- This end is a beginning
- A look at the substitute bench
- To understand the withdrawal
- Reward pays twice
- Alternatively, for patients with COPD: Chronic bronchitis
- Food as an alternative?
- Relax without nicotine
- Race against addiction
- People are creatures of habit
- Alternative for light smokers: one is none
- With knowingness against the levity
- Alternative for stress-smoking: stress does not dissolve in smoke
- The journey is the destination

Figure 1. Specific daily tasks including interactive buttons, video clips, and learning progress quizzes.

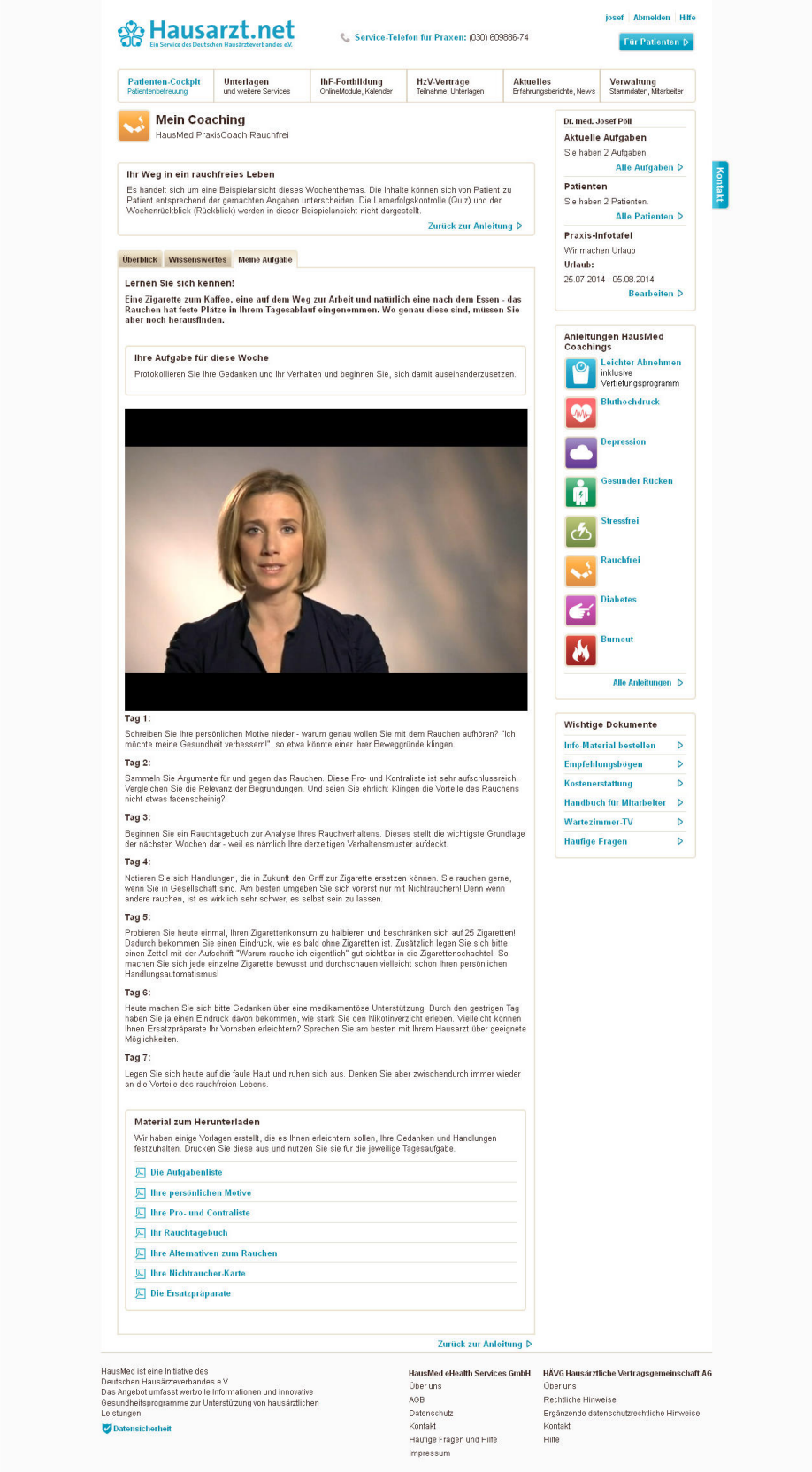
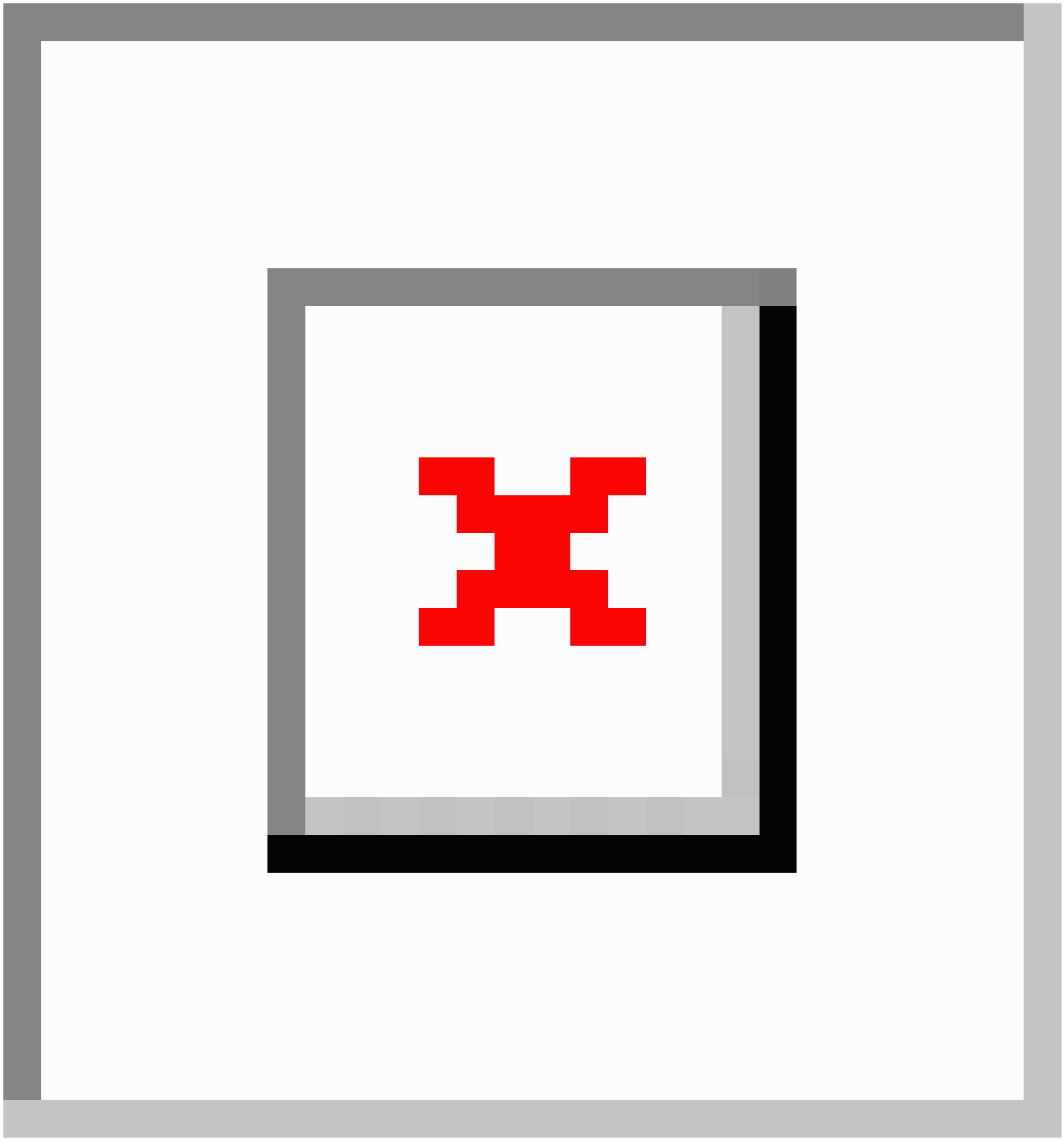


Figure 2. Sample printed materials: emergency plan, relaxation exercises, questionnaires, information, and self-agreements.



At the end of each week, participants were asked to give feedback via the Internet concerning their condition and level of motivation and whether or not they did their weekly tasks. Participants could also communicate among themselves on a forum or asked a HausMed team member in case they had any questions. The active monitoring (or rather supervising) of the entire 12-week coaching course was carried out by the GP through a separate login account in a secured physician area (motivation, condition, and status of the module exercise). In addition to that, three specified telephone calls from the GP or a qualified practice nurse (Weeks 2, 4, and 12) were implemented to motivate and support the participants. If either a participant's motivation or condition declined notably at any point during the coaching period or if the module exercise was not completed, additional counseling from the GP or practice nurse was given over the telephone. There was no limitation to the frequency of website use, but participants were given a goal of using the website at least once a week and GPs were advised to log in into the program twice a week. No changes were made to the coaching program within the entire study period.

Outcome Measures

The primary outcome measure was biochemically confirmed smoking status 12 weeks after inclusion into the study by use of a cotinine urine test. Secondary outcome measures were self-reported smoking status, number of NRTs, weight in kilograms, number of smoked cigarettes per day, physical activity (range from 0-4), and breathing difficulties (range from

0-4). A higher number on the scale refers to a more frequent physical activity and to more breathing difficulties. No changes were made on the trial outcome measures within the entire study period.

Statistical Analysis

Sample size calculation was performed with G*Power 3 correcting for the cluster design (intracluster correlation coefficient=.05, average cluster size=3) for two-sided testing (alpha is 5% and a power of 80%). For the expected effect, abstinence rates of 30% versus 10% were assumed. Using these assumptions, the calculated total sample size for primary outcome smoking cessation was 152 participants. Taking expected attrition into account, we aimed at recruiting a total of 180 participants in about 80 general practices.

Baseline data are presented descriptively. Group differences were calculated for all participants whose smoking status was available at baseline and follow-up (completer collective). Sensitivity analysis was performed by an intent-to-treat analysis assuming that participants with missing values had no smoking status change at all. The strongly variable cluster size caused major numerical problems in the linear mixed model analysis. As it was not possible to adjust for intracluster correlations properly and because of the high variability of patients in practices, it was decided to perform the main analysis using Fisher's exact test without accounting for the clusters. For the main outcome smoking status, we also performed secondary analyses based on logistic regression analyses with adjustments for age, gender, height, and number of cigarettes smoked per day. We further conducted generalized estimating equations as sensitivity analysis to account for practices as patient clusters. All analyses were performed using SPSS version 19.0.

Results

Originally 92 practices were interested in participating and were randomized. However, 16 practices withdrew early after randomization (7 GPs from the intervention and 9 GPs from the control group), and 34 practices (19 GPs from the intervention and 15 GPs from the control group) did not recruit any participants for the study (Figure 3). Altogether, 168 patients were recruited from 42 practices (86 patients in 20 intervention practices; 82 patients in 22 control practices) between May 19, 2011, and April 1, 2013. More than half of participants (54.5%, 66/121) were female, and the average age was 45.5 years. At 12 weeks, 35 participants in the intervention group did not show up for the measurement. In the usual care group, 12 participants had missing values at 12 weeks, and 1 of them had a missing baseline value. For 121 participants (51 from the intervention and 70 from the control group), information on smoking status was available both at baseline and after 12 weeks (complete-case). The proportion of non-completers (35/86; 12/82) was significantly higher (chi-square test, $P<.001$) in the intervention than in the usual care group.

The intervention and control group were similar in gender, age, weight, number of cigarettes smoked per day, and number of years with nicotine consumption. However, participants in the control group were significantly taller. A total of 7 participants used NRTs, and one participant in each group used varenicline at enrollment. There was no significant group difference found for the use of NRT or for the intake of varenicline (Table 1).

The self-reported cessation rate among the intervention group participants was 17.6% (9/51) and among the control group participants was 14.3% (10/70) without a significant group difference ($P=.623$). A logistic regression without adjustment (OR 0.78, 95% CI 0.29-2.08) and after adjustment for age, gender, and number of cigarettes smoked per day at baseline (OR 0.62, 95% CI 0.22-1.78) revealed similar results. Within the intent-to-treat analysis, self-reported cessation rate among the intervention group was 10.5% (9/86). The self-reported cessation rate of the control group was 11.3% (10/82). Results from the logistic regression without adjustment (OR 1.19, 95% CI 0.46-3.09) and after adjustment for age, gender, height, and number of cigarettes smoked per day at baseline (OR 1.04, 95% CI 0.39-2.80) were similar.

The results of the biochemical validation of the self-reported cessation status revealed that only a few (3/8) of the documented cotinine tests of the interventions group were positive, although they self-reported smoking cessation. None of these 3 participants were on nicotine replacement therapies, which would have explained the positive cotinine tests. The validation of the self-reported cessation status from the control group showed no disconfirmation. All six conducted cotinine tests were negative. No cotinine tests were administered for 1 participant of the intervention group and 4 participants of the control group who reported smoking cessation. The cessation rate by use of the biochemical validation was 9.8% for the intervention group (5/51) and 8.6% for the control group (6/70). Results were similar from the logistic regression without adjustment (OR 0.86, 95% CI 0.25-3.0) and after adjustment for age, gender, and number of cigarettes smoked per day at baseline (OR 0.63, 95% CI 0.17-2.40). The secondary analysis using a generalized estimating equation showed also a non-significant result ($P=.74$). Within the intent-to-treat analysis, the confirmed cessation rate among the intervention group was 5.8% (5/86) and for the control group 7.3% (6/82). Results from the logistic regression without adjustment (OR 1.28, 95% CI 0.38-4.36) and after adjustment for age, gender, and number of cigarettes smoked per day at baseline (OR 1.01, 95% CI 0.28-3.62) were similar (Table 2).

Figure 3. Participant flow of the study (GP=general practitioner; ITT=intent-to-treat; CC-Analysis=complete-case).

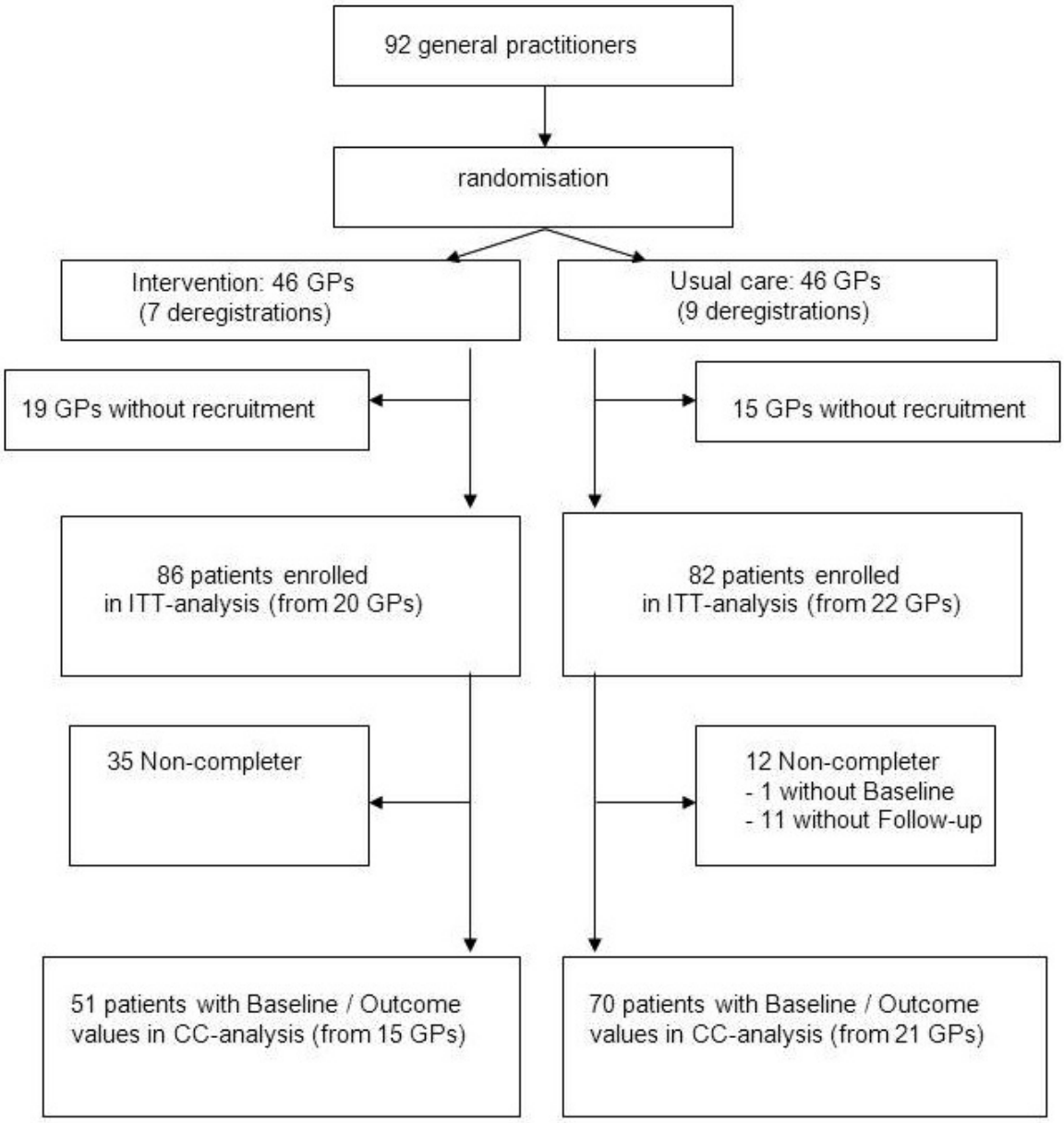


Table 1. Baseline characteristics at enrollment.

		Intervention			Usual care				
		n	mean	SD	n	mean	SD	Δ mean	*P*
Age, years		86	42.2	12.6	81	45.8	12.8	2.6	.072[a]
Weight, kg		85	73.3	15.7	81	78.0	17.2	4.3	.068[a]
Height, cm		85	169.8	8.7	81	173.5	9.3	4.2	.009[a]
Cigarettes smoked per day, n		86	18.2	7.4	81	16.6	9.4	2.8	.218[a]
Tobacco use, years		86	22.7	12.0	81	24.3	12.1	2.4	.410[a]
Use of NRT		4			3				1.0[b]
	Nicotine patch	3			2				1.0[b]
	Nicotine gum	1			1				1.0[b]
Intake of varenicline		1			1				1.0[b]
Gender, n (%)									.274[b]
	Females	51 (60.0)			40 (49.4)				
	Males	34 (40.0)			41 (50.6)				

[a]*P* values from Student *t* test.

[b]*P* values from Fisher's exact.

Table 2. Results of self-reported and biochemically confirmed cessation status for intent-to-treat and complete-case analyses (*P* values are from Fisher's exact test).

	Intent-to-treat					Complete-case				
	Intervention group		Usual care			Intervention group		Usual care		
	n	%	n	%	*P*	n	%	n	%	*P*
Self-reported cessation	9	10.5	10	12.2	.810	9	17.6	10	14.3	.623
Not biochemically confirmed										
Relapse	77	89.5	72	87.8		42	82.4	60	85.7	
Disconfirmed	3	3.5	0	0		3	5.9	0	0	
Missing test	1	1.2	4	4.9		1	2	4	5.7	
Biochemically confirmed	5	5.8	6	7.3	.762	5	9.8	6	8.5	1.0
Total N	86	100	82	100		51	100	70	100	

The result from the secondary outcome, number of cigarettes smoked per day, revealed within the complete-case analysis that the mean difference after 12 weeks of the intervention group was 2.7 cigarettes less than the mean difference of the control group (95% CI -5.33 to -0.58; *P*=.045). After adjustment for number of cigarettes smoked per day at baseline, age, gender, and height, the difference between groups was no longer significant (95% CI 0.27-4.72; *P*=.080). There were no statistically significant group differences for other secondary outcomes like weight, physical activity, use of NRT, intake of varenicline, and breathing difficulties (Table 3). After 12 weeks, 8 participants used NRT and 4 participants used varenicline. One participant of the usual care group with a documented use of NRT did not further specify the used NRT. No participants using nicotine replacement had either a self-reported or biochemically confirmed cessation status.

Adverse events from 26 participants were documented. In the intervention group, 4 participants reported weight gain, 2 participants had increased perceived stress, 1 participant had a sleep disorder, and 1 participant had increased irritability. In the usual care group, 6 participants had increased perceived stress, 5 participants had cardiovascular problems, 4 participants reported fatigue, 4 participants reported weight gain, 2 participants had sweating, 1 participant had a sleep disorder, and 1 participant specified increased irritability. Additionally, one serious adverse event occurred that was not directly related to the intervention: a participant in the usual care group died due to a spontaneous rupture of an aortic aneurysm. This event was considered to be unrelated to the intervention and the study procedures.

Table 3. Results of NRT, breathing difficulties, physical activity, weight, and number of smoked cigarettes per day (complete-case analysis).

		Intervention			Usual care				
		n	mean	SD	n	mean	SD	Δ mean	*P*
Use of NRT									
Baseline		4			3				.462[a]
	Nicotine patch	3			2				.651[a]
	Nicotine gum	1			1				1.0[a]
After 12 weeks		4			4				.723[a]
	Nicotine patch	0			2				.511[a]
	Nicotine gum	4			1				.167[a]
Intake of varenicline									
Baseline		1			1				1.0[a]
After 12 weeks		1			3				.640[a]
Breathing difficulties: (scale from 0 to 4)									
Baseline		51	1.0	1.2	70	1.0	1.0	0	.980[b]
After 12 weeks		51	0.69	0.99	69	0.75	1.0	0.06	.704[b]
Physical activity: (scale from 0 to 4)									
Baseline		51	1.4	1.0	70	1.4	1.0	0	.513[b]
After 12 weeks		51	1.25	1.25	70	1.44	1.1	0.19	.231[b]
Weight, kg									
Baseline		51	74.4	16.3	70	78.7	17.3	4.3	.169[b]
After 12 weeks		48	74.8	17.2	67	80.1	17.2	5.3	.107[b]
Number of smoked cigarettes per day									
Baseline		51	19.5	7.0	70	16.7	8.9	2.8	.063[b]
After 12 weeks		51	10.2	7.7	69	10.2	8.3	-0.02	.989[b]
Difference		51	-9.3	7.1	69	-6.6	7.3	2.7	.045[b]

[a]*P* values from Fisher's exact test.

[b]*P* value from Student *t* test.

Discussion

Principal Findings

We found that the Web-based coaching program in combination with telephone counseling and monitoring in general practice was not effective for achieving smoking cessation compared to usual care.

Some trials have already shown that smoking cessation programs can be delivered effectively via the Internet [8,16-18], although Web-based interventions are often accompanied by a high attrition rate [19,20]. The evidence has been summarized in a recent meta-analysis by Chen et al [9]. They concluded that computer and other electronic aids increase to a small extent the likelihood of prolonged smoking cessation compared to no intervention. Another meta-analysis of Web- and computer-based smoking cessation programs indicates that there is currently sufficient evidence to support their use [21]. They stated within a meta-analysis of 22 randomized controlled trials (RCTs) that Web- or computer-based smoking cessation programs led to a 1.5 times higher abstinence rate compared to control groups. The effect of Web- or computer-based programs was therefore similar to that of counseling interventions. Their pooled cessation rate for Web- or computer-based programs over the long term (12 months) was 9.9% (95% CI 8.9-10.9). The advantage of a Web-based intervention compared to a control group cannot be confirmed by the present results, although our abstinence rates from both groups in the short term were comparable to theirs. In a Cochrane review, Civljak et al [10] also detected no consistent effects from trials that compared Internet interventions with usual care.

For example, Muñoz et al [22] suggested that quit rates obtained by using Internet interventions for smoking cessation are comparable with quit rates reported from smoking cessation

therapies or smoking cessation groups. Our present result supports this fact as there was no significant difference between the intervention and the usual care in a primary care setting. Previous findings with a biochemical confirmation of the self-reported cessation support our results. Patten et al [23] revealed in adolescent smokers no significant treatment differences between a brief office intervention and a Web-based intervention. Their smoking abstinence rates after 24 weeks were 12% versus 6% for the brief office intervention and Internet-based intervention, respectively. The present discrepancy of the self-reported smoking status (17.6%) and the biochemical validation (9.8%) was noticeable even if this conspicuousness is based on only three individuals in the treatment group with a disconfirmed smoking status. It has been reported that biochemical validation does not modify the conclusions of low-intensity interventions trials [24]. Glasgow et al observed a disconfirmation rate between 4-5% over all groups. Compared to the present disconfirmation, we identified a similar rate only in the intervention group of 3.5-5.9%. In the control group, all conducted cotinine tests were negative so there was no disconfirmation to be documented. The self-reported cessation rate from 17.6% (complete-case) of the intervention group compared to the confirmed rate from 9.8%, beside the missing validation of 2% (n=1), leads to the suggestion of a certain overestimation from the present self-reported cessation rate. Nevertheless, this discrepancy could have still occurred by chance due to the small number of disconfirmed results.

The present study underlines how difficult it is to achieve smoking cessation in primary care and how difficult it is to motivate patients to quit smoking. Based on a parallel trial conducted with comparable conditions, we identified that a Web-based intervention with telephone counseling led to a significantly greater weight reduction [25] compared to usual care. This advantage was not shown regarding smoking cessation. Many reasons have been already identified to explain the low effectiveness of smoking cessation interventions in general practice. Patients' low motivation to quit [26,27], patients' low compliance [26,28], patients' resistance to speak about smoking [29,30], lack of time [31,32], lack of economical reimbursement [28,33], lack of skills and low self-efficacy [32,34], consideration of smokers' other problems [30,35], and an uneasy feeling when talking about smoking cessation [31,32] have been already identified as barriers. Notwithstanding, even interventions with a small effect are capable of substantially decreasing diseases that are associated with nicotine use [36].

Strengths and Limitations

Strengths of the present study were the embedding of the study in a realistic primary care setting and the use of a biochemical validation of the self-reported cessation status. However, some important methodological aspects for the interpretation of the study results need to be considered. First, the randomization of the present study was conducted at the practice level before individual participants were included. Thus, physicians knew whether they recruited patients for the intervention or the control group, which could lead to bias. Second, due to the highly variable cluster sizes the statistical analysis of our data was not straightforward. Classical linear mixed models taking the cluster design into account could not be used because of numerical problems. Therefore, we used simple Fishers' exact test (which ignores intracluster correlation) and an additional multilevel analysis (which runs into problems when cluster sizes differ) as sensitivity analysis. Third, according to our power calculations the target number of participants was not completely reached due to a slow recruitment of participants, and at a certain point the study had to be stopped, which may have reduced the study's ability to detect significant differences between the groups. Fourth, the proportion of participants without follow-up values was undoubtedly higher in the intervention than in the usual care group. This could be partly due to the fact that participants in the control group received a small financial incentive while those in the intervention group did not. Participants in the intervention group might also have been less willing to have an additional practice visit after completing the program than those in the control group who had little practice contact otherwise. Therefore, our complete-case analysis with the self-reported cessation might overestimate the rates to some extent. Within the intent-to-treat analysis, where the missing post values were replaced without a change of smoking status (baseline carried forward), the cessation rates were clearly smaller. Fifth, the content of the usual care was not further evaluated. The practitioners of the control group were asked to change nothing in their usual way of counseling and to treat their participants in the same manner as usual. There was no additional documentation of their counseling provided.

Conclusions

Our findings suggest that the tested Web-based coaching program in combination with telephone counseling and monitoring in general practice was not effective for achieving smoking cessation compared to usual care. The effect was similar to usual primary care and comparable to other Web-based interventions. We identified a discrepancy of self-reported smoking cessation and the biochemical validation, which should be reconsidered for further studies. The limited statistical power and the high drop-out rate may have reduced the study's ability to detect significant differences between the groups. Further RCTs are needed in order to investigate long-term outcomes and interventions in larger populations, as well as in the contents of usual primary care.

Authors' Contributions

AS, KL, and MM designed the study. MM wrote the initial protocol with supervision from AS and KL. MM coordinated the study with MH. MM, KL, AS, MH, and SW did the analysis. MM drafted the manuscript with contributions from AS, KL, SW, and MH. All authors read and approved the final manuscript. MM is the guarantor.

Conflicts of Interest

This study was completely funded by HausMed eHealth Services GmbH (Berlin, Germany). The sponsor did not have access to study data and did not influence the development of this manuscript. AS, KL, and MM are employed at University Hospital Klinikum rechts der Isar, Technische Universität München. MH is a medical student. SW is employed at the Institute for Medical Biometry, Epidemiology und Medical Informatics (IMBEI), Universitätsklinikum des Saarlandes, Homburg/Saar.

Multimedia Appendix 1

HausMed - Presentation Video.

[MP4 File (MP4 Video), 20MB - jmir_v16i9e218_app1.mp4]

Multimedia Appendix 2

Waiting room advertisement.

[PPT File (Microsoft PowerPoint Presentation), 1MB - jmir_v16i9e218_app2.ppt]

Multimedia Appendix 3

CONSORT-EHEALTH checklist V1 6.2 [37].

[PDF File (Adobe PDF File), 995KB - jmir_v16i9e218_app3.pdf]

References

1. World Health Organization. Fact sheet N° 339. 2013. URL: http://www.who.int/mediacentre/factsheets/fs339/en/ [accessed 2014-05-12] [WebCite Cache ID 6PVulbQr1]
2. World Health Organization. Building blocks for Tobacco Control - A Handbook. 2004. URL: http://whqlibdoc.who.int/publications/2004/9241546581_eng.pdf?ua=1 [accessed 2014-05-12] [WebCite Cache ID 6PVzOl3sT]
3. Lader D. Smoking-related Behaviour and Attitudes. 2008. Opinions Survey Report No 40 URL: http://www.ons.gov.uk/ons/rel/lifestyles/smoking-related-behaviour-and-attitudes/2008-09/index.html [accessed 2014-05-12] [WebCite Cache ID 6PW0bgVHN]
4. Centers for Disease Control. MMWR. Morbidity and Mortality Weekley Report 2011; 60; 2011. Quitting smoking among adults - United States - 2001-2010 URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6044a2.htm [accessed 2014-09-17] [WebCite Cache ID 6Sf7PAMv5]
5. Silagy C, Stead LF. Physician advice for smoking cessation. Cochrane Database Syst Rev 2001(2):CD000165. [doi: 10.1002/14651858.CD000165] [Medline: 11405953]
6. Brown J, Michie S, Raupach T, West R. Prevalence and characteristics of smokers interested in internet-based smoking cessation interventions: cross-sectional findings from a national household survey. J Med Internet Res 2013;15(3):e50 [FREE Full text] [doi: 10.2196/jmir.2342] [Medline: 23506944]
7. Royal Pingdom. 2012. Internet 2012 in numbers URL: http://royal.pingdom.com/2013/01/16/internet-2012-in-numbers/ [accessed 2014-05-12] [WebCite Cache ID 6PVyo35Rz]
8. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
9. Chen YF, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. Health Technol Assess 2012;16(38):1-205, iii [FREE Full text] [doi: 10.3310/hta16380] [Medline: 23046909]
10. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
11. Research Randomizer Program. URL: http://www.randomizer.org/ [accessed 2014-07-18] [WebCite Cache ID 6RA14U999]
12. HausMed eHealth Services GmbH. Berlin, Germany URL: http://www.hausmed.de/ [accessed 2014-07-18] [WebCite Cache ID 6RA1xITdL]
13. Miller WR, Rollnick S. Motivational Interviewing. In: Helping People Change (Applications of Motivational Interviewing). New York: The Guilford Press; 2002.
14. Batra A, Buchkremer G. Nichtrauchen!. Stuttgart: Kohlhammer Verlag; 2013.
15. Batra A. Treatment of tobacco dependence. Dtsch Arztebl Int 2011 Aug;108(33):555-564 [FREE Full text] [doi: 10.3238/arztebl.2011.0555] [Medline: 21912578]
16. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]

17. Prochaska JO, Butterworth S, Redding CA, Burden V, Perrin N, Leo M, et al. Initial efficacy of MI, TTM tailoring and HRI's with multiple behaviors for employee health promotion. Prev Med 2008 Mar;46(3):226-231 [FREE Full text] [doi: 10.1016/j.ypmed.2007.11.007] [Medline: 18155287]
18. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8 Suppl 1:S59-S67. [Medline: 17491172]
19. Mañanes G, Vallejo MA. Usage and effectiveness of a fully automated, open-access, Spanish Web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2014;16(4):e111 [FREE Full text] [doi: 10.2196/jmir.3091] [Medline: 24760951]
20. 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]
21. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
22. Muñoz RF, Aguilera A, Schueller SM, Leykin Y, Pérez-Stable EJ. From online randomized controlled trials to participant preference studies: morphing the San Francisco Stop Smoking site into a worldwide smoking cessation resource. J Med Internet Res 2012;14(3):e64 [FREE Full text] [doi: 10.2196/jmir.1852] [Medline: 22739225]
23. Patten CA, Croghan IT, Meis TM, Decker PA, Pingree S, Colligan RC, et al. Randomized clinical trial of an Internet-based versus brief office intervention for adolescent smoking cessation. Patient Educ Couns 2006 Dec;64(1-3):249-258. [doi: 10.1016/j.pec.2006.03.001] [Medline: 16616449]
24. Glasgow RE, Mullooly JP, Vogt TM, Stevens VJ, Lichtenstein E, Hollis JF, et al. Biochemical validation of smoking status: pros, cons, and data from four low-intensity intervention trials. Addict Behav 1993;18(5):511-527. [Medline: 8310871]
25. Mehring M, Haag M, Linde K, Wagenpfeil S, Frensch F, Blome J, et al. Effects of a general practice guided web-based weight reduction program--results of a cluster-randomized controlled trial. BMC Fam Pract 2013;14:76 [FREE Full text] [doi: 10.1186/1471-2296-14-76] [Medline: 23981507]
26. Young JM, Ward JE. Implementing guidelines for smoking cessation advice in Australian general practice: opinions, current practices, readiness to change and perceived barriers. Fam Pract 2001 Feb;18(1):14-20 [FREE Full text] [Medline: 11145622]
27. Cornuz J, Ghali WA, Di Carlantonio D, Pecoud A, Paccaud F. Physicians' attitudes towards prevention: importance of intervention-specific barriers and physicians' health habits. Fam Pract 2000 Dec;17(6):535-540 [FREE Full text] [Medline: 11120727]
28. O'Loughlin J, Makni H, Tremblay M, Lacroix C, Gervais A, Déry V, et al. Smoking cessation counseling practices of general practitioners in Montreal. Prev Med 2001 Dec;33(6):627-638. [doi: 10.1006/pmed.2001.0937] [Medline: 11716660]
29. McLeod D, Somasundaram R, Howden-Chapman P, Dowell AC. Promotion of smoking cessation by New Zealand general practitioners: a description of current practice. N Z Med J 2000 Nov 24;113(1122):480-485. [Medline: 11198538]
30. Coleman T, Murphy E, Cheater F. Factors influencing discussion of smoking between general practitioners and patients who smoke: a qualitative study. Br J Gen Pract 2000 Mar;50(452):207-210 [FREE Full text] [Medline: 10750230]
31. Gallefoss F, Drangsholt K. [Smoking cessation intervention and barriers against it among general practitioners in Vest-Agder county]. Tidsskr Nor Laegeforen 2002 Nov 10;122(27):2608-2611 [FREE Full text] [Medline: 12523189]
32. Helgason AR, Lund KE. General practitioners' perceived barriers to smoking cessation-results from four Nordic countries. Scand J Public Health 2002;30(2):141-147. [doi: 10.1080/14034940210133799] [Medline: 12028863]
33. Saywell RM, Jay SJ, Lukas PJ, Casebeer LL, Mybeck KC, Parchman ML, et al. Indiana family physician attitudes and practices concerning smoking cessation. Indiana Med 1996;89(2):149-156. [Medline: 8867414]
34. Kottke T, Willms D, Solberg L, Brekke M. Physician-delivered smoking cessation advice: issues identified during ethnographic interviews. Tobacco Control 1994 Mar 01;3(1):46-49. [doi: 10.1136/tc.3.1.46]
35. Bremberg S, Nilstun T, Kovac V, Zwitter M. GPs facing reluctant and demanding patients: analysing ethical justifications. Fam Pract 2003 Jun;20(3):254-261 [FREE Full text] [Medline: 12738693]
36. Ockene JK. Physician-delivered interventions for smoking cessation: strategies for increasing effectiveness. Prev Med 1987 Sep;16(5):723-737. [Medline: 3317393]
37. 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

COPD: chronic obstructive pulmonary disease
GP: general practitioner
NRT: nicotine replacement therapy
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 13.05.14; peer-reviewed by H Brendryen, S Hall; comments to author 27.05.14; revised version received 18.07.14; accepted 23.08.14; published 24.09.14.

Please cite as:
Mehring M, Haag M, Linde K, Wagenpfeil S, Schneider A
Effects of a Guided Web-Based Smoking Cessation Program With Telephone Counseling: A Cluster Randomized Controlled Trial
J Med Internet Res 2014;16(9):e218
URL: http://www.jmir.org/2014/9/e218/
doi:10.2196/jmir.3536
PMID:25253539

Original Paper

Short-Term Effectiveness of Web-Based Guided Self-Help for Phobic Outpatients: Randomized Controlled Trial

Robin N Kok[1,2], MSc; Annemieke van Straten[1], PhD; Aartjan T F Beekman[3], PhD; Pim Cuijpers[1,4], PhD

[1]Department of Clinical Psychology and the EMGO institute for Health and Care Research, VU University Amsterdam, Amsterdam, Netherlands
[2]National Institute for Mental Health Research, The Australian National University, Canberra, Australia
[3]Department of Psychiatry and the EMGO institute for Health and Care Research, VU University Medical Centre, Amsterdam, Netherlands
[4]Innovation Incubator, Division Health-Training.Online, Leuphana University, Lüneburg, Germany

Corresponding Author:
Robin N Kok, MSc
Department of Clinical Psychology and the EMGO institute for Health and Care Research
VU University Amsterdam
van der Boechorststraat 1
Amsterdam, 1081 BT
Netherlands
Phone: 31 205983833
Fax: 31 205988756
Email: r.n.kok@vu.nl

Abstract

Background: Internet-based guided self-help has been successfully used in the general population, but it is unknown whether this method can be effectively used in outpatient clinics for patients waiting for face-to-face psychotherapy for phobias.

Objective: The aim was to assess the clinical effectiveness of Phobias Under Control, an Internet-based intervention based on exposure therapy with weekly guidance.

Methods: We conducted a randomized controlled trial, recruiting 212 outpatients scheduled to receive face-to-face psychotherapy for any type of phobia at an outpatient clinic. Participants suffering from at least 1 DSM-IV or ICD-10 classified phobia (social phobia, agoraphobia with or without panic disorder, and/or specific phobia as ascertained by a telephone interview at baseline) were randomly allocated to either a 5-week Internet-based guided self-help program based on exposure therapy with weekly student support followed by face-to-face psychotherapy (n=105) or a wait-list control group followed by face-to-face psychotherapy (n=107). Primary outcome was the Fear Questionnaire (FQ). Secondary outcomes were the Beck Anxiety Inventory (BAI) and Center of Epidemiological Studies-Depression scale (CES-D). Assessments took place by telephone at baseline (T0) and on the Internet at posttest (T1, self-assessment at 5 weeks after baseline). Missing data at T1 were imputed.

Results: At posttest, analysis of covariance on the intention-to-treat sample showed significant but small effect sizes between intervention and control groups on the FQ (d=0.35, *P*=.02), CES-D (d=0.34, *P*=.03), and a nonsignificant effect size on the BAI (d=0.28. *P*=.05). Although initial acceptance was good, high nonresponse was observed, with 86 of 212 participants (40.5%) lost to follow-up at T1 and only 14 of 105 (13.3%) intervention participants finishing all 5 weeks.

Conclusions: Phobias Under Control is modestly effective in lowering phobic and depressive symptoms in a relatively short period and may be clinically beneficial when implemented in routine outpatient practice.

Trial Registration: Netherlands Trial Register NTR2233; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2233 (Archived by WebCite at http://www.webcitation.org/6O2ioOQSs).

(J Med Internet Res 2014;16(9):e226) doi:10.2196/jmir.3429

KEYWORDS

phobias; phobic disorders; anxiety disorders; Web-based intervention; Internet therapy; randomized controlled trial; outpatients

Introduction

Phobias are among the most common mental disorders and the most common type of anxiety disorders [1]. Specific phobias are the most common form of anxiety disorders for both genders, with a total 12-month prevalence of 7.1%, followed by social phobia (4.8%) and agoraphobia without panic disorder (1.2%). All phobias have a negative impact on quality of life and psychosocial functioning [2], and the societal burden of phobias is considerable [3,4]. Despite detrimental effects on quality of life, research has shown a substantial delay of more than 10 years between onset of symptoms and first therapy attendance [5]. In a recent study, social phobia was found not only to have the earliest onset age, but also an even longer delay—on average 28 years—in seeking treatment [6]. Notwithstanding the impact of a phobia on a patient's quality of life [2,7], phobias are often not the primary reason for seeking treatment from an outpatient clinic [8,9] and it has been argued that commonly occurring comorbid disorders, such as depression, mask underlying social phobia leading to underdiagnosis in primary care [10]. This suggests widespread undertreatment [11] for these disorders, even though there is robust evidence of efficacious psychological treatments for agoraphobia [12], social phobia [13], and specific phobias [14], most notably exposure therapy and cognitive behavioral therapy (CBT).

Internet-based interventions are increasingly popular adaptations of evidence-based psychotherapies as a replacement of, or adjunct to, traditional face-to-face therapies. Starting with computer-based, offline interventions (eg, [15,16]), existing therapies such as CBT, exposure therapy, systematic desensitization, and relaxation were found to be efficacious [17] and were rewritten to suit delivery on the Internet [18,19]. In past years, Internet interventions have been found efficacious for a number of anxiety disorders [17,18,20,21] and phobias, including agoraphobia [22], specific phobias [14,23,24], and social anxiety disorder [25-27]. Thus, Internet-delivered psychological treatments for anxiety and phobias are feasible, acceptable, and effective.

Typically, outpatients exhibit higher levels of anxiety and a greater number of comorbid and more complex diagnoses, as well as greater psychosocial impairment when compared with general and primary care populations [5]. Previous research has primarily focused on self-referred participants from primary care settings or from the general population [28], and although some evidence exists on the effectiveness of routine psychological interventions in outpatients [29], only a limited number of trials have specifically evaluated Internet-based treatments in outpatient clinics and secondary care for common mental disorders [30-33]. To the best of our knowledge, there appear to be no large-scale high-quality trials evaluating the efficacy of Internet-based exposure therapy in phobic outpatients.

Because waiting lists are commonplace in outpatient clinics, time spent waiting for face-to-face treatment could be spent effectively by offering a (guided) self-help intervention to patients. Delegating the routine, basic elements of exposure treatment to a guided Internet-based situation could shorten face-to-face therapy and limit therapist involvement, making the treatment more cost-effective [19,34]. Previous research has indicated that Internet-based therapy for social phobia might be cost-effective relative to face-to-face therapy [35,36]. Furthermore, because pretreatment dropout is common in outpatient clinics [37], a second postulated benefit may be that continually engaging the patients in their treatments throughout the wait-list period will result in lower pretreatment attrition or "no shows."

The objective of the current trial was to assess the short-term clinical effectiveness of offering Internet-based guided self-help to outpatients compared to a wait-list control. To our knowledge, this is the first large-scale randomized controlled trial of Internet-based treatment for phobias in outpatients. As such, it will also provide valuable information on the acceptability and feasibility of such an intervention in outpatient clinics. This paper describes the principal short-term outcomes of this multifaceted trial.

Methods

Trial Design

A full trial protocol is available elsewhere [38]. This trial was approved by the Medical Ethics Committee of the VU Medical Centre, Amsterdam (registration number 2010/77) and registered with the Dutch Trial Registry (NTR2233). A total of 481 participants who recently applied for psychological treatment at an outpatient clinic consented to be contacted by our research group and were referred to the researchers from August 2010 to December 2013. After briefing the participants about the aims of the study, screening, and obtaining informed consent in writing, eligible participants (n=212) were administered a telephone baseline questionnaire and participants were randomized to either the intervention group (n=105) or treatment as usual (n=107). Patients who were ineligible (n=111), declined participation, or could not be contacted (n=153) remained on the waiting list for face-to-face treatment. The research did not interfere with the outpatient clinics' wait-list duration or start of treatment and participants could start face-to-face psychotherapy after the intervention or control group period.

Participants

Recruitment Procedure

A total of 8 specialized anxiety disorder outpatient clinics in medium-to-large cities in the west of the Netherlands participated. Clinics were selected for a high monthly volume of patients for practical reasons. Participants were referred to the outpatient clinics by their general practitioners (GPs), briefly screened, and placed on a waiting list. Recruitment commenced in August 2010 and was stopped in December 2013 to allow for sufficient follow-up time. Waiting lists for outpatient psychotherapy are common in the Netherlands, and time spent on a waiting list is usually at least 6 weeks from first referral to first treatment session. At the start of the wait-list period, participants presenting with a phobia as a primary or secondary disorder were referred to the researchers and screened by telephone using the Composite International Diagnostic Interview (CIDI) [39] for presence of any phobia by master's

level students. Consequently, exclusion criteria were checked and baseline measures were administered. During this wait-list period, a nontherapeutic meeting with a health care professional from the outpatient clinic took place to ascertain treatment needs and to determine optimal face-to-face treatment for all participants. Additional details on recruitment are available elsewhere [38].

Eligibility Criteria for Participants

All computer-literate patients with a possible phobia (social phobia, agoraphobia with or without panic disorder, specific phobia) were referred to the researchers by the outpatient clinic even if a phobia was not the primary reason for seeking treatment at an outpatient clinic. Participants had to (1) be 18 years or older, (2) be currently enrolled to receive face-to-face psychotherapy at 1 of the participating outpatient clinics, and (3) have a *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision; *DSM-IV-TR*) or *International Classification of Diseases, Tenth Revision* (*ICD-10*) diagnosis of any phobia as established by the CIDI. Psychotropic medication use was allowed if stable for at least the duration of the intervention or control group period. Patients presenting with psychotic disorders or at elevated risk for suicide were excluded from the trial, but remained on the waiting list for face-to-face psychotherapy at their outpatient clinics.

Interventions

Internet-Based Guided Self-Help

The Internet-based intervention is an adaptation of an existing self-help book on phobias [40]. The intervention is offered at no cost to the participant, takes 5 weeks to complete, and is based on psychoeducation and exposure therapy. The broad and nonspecific focus of the intervention is on identifying and correcting avoidance behavior by using exposure, a common and evidence-based therapeutic component of most phobia therapies [41]. This broad focus facilitates using the intervention for the entire range of phobias. The intervention was presented to the prospective participants as a free-of-charge voluntary course to start reducing their phobic symptoms during the wait-list time. They were told that the intervention was based on evidence-based principles and that the elements they would encounter during the intervention would essentially be the same as in their upcoming face-to-face psychotherapy, allowing for a head start in their treatment. Participants were informed that face-to-face treatment would commence at the scheduled time, regardless of whether they enrolled in the study, and that their decision to participate or not would neither postpone nor advance their face-to-face treatment.

During the intervention, participants build a hierarchy (see Figure 1) of fear-inducing situations or stimuli and expose themselves to these situations or stimuli gradually. The participant completes exposure exercises as homework assignments and reports on his or her accomplishments to the coach each week. In the first weeks, the participant makes an inventory of his or her avoidance and safety behaviors and defines a focal point for exposure situations and a desired behavioral goal. The participant then plans a number of gradual exposure exercises to be executed for the upcoming week, with exposure exercises becoming gradually more challenging each week (see Figure 2). The coach monitors the fear hierarchy and planning and replies with a supportive message once a week for 5 weeks, relevant to the participant's homework experiences through the secure online platform. All coaching was supervised by an experienced psychotherapist. The intervention is tunneled (ie, no new material is available to the participant until the participant has reported on that week's achievements and the coach has provided feedback on these achievements). If applicable, the coach sends a standardized reminder message through the secure online platform if the participant did not use the website that week. All actions on the platform (eg, new feedback received, new exercise available) prompted an immediate automated email to the participant. Material from previous weeks remains accessible to the participant. Online coaching messages were delivered through a secured message system on the intervention website by trained and supervised master's level students of clinical psychology. The participant completes exposure exercises alone and reports on completed exercises weekly. Throughout the intervention period, the participants were kept on the waiting list for face-to-face psychotherapy.

The website platform was migrated to an updated version during the recruitment period. This migration was performed to ensure continuing safety of participant data in accordance with Dutch law and to resolve or mitigate critical bugs and shortcomings in website functionality. Website content, however, remained unaltered throughout recruitment. No substantial website downtime was observed during recruitment.

Figure 1. Screenshot of Phobias Under Control: fear hierarchy.

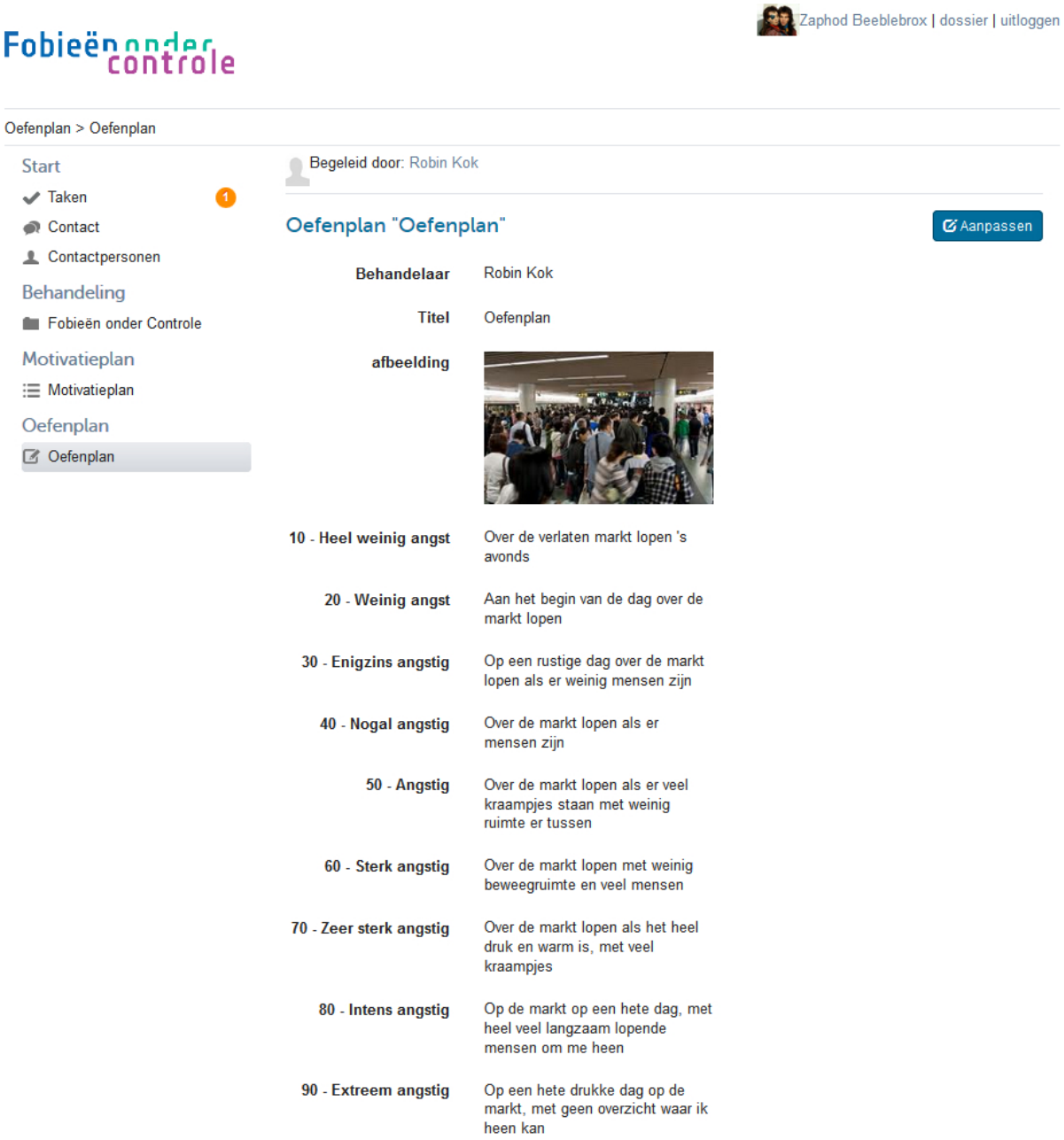
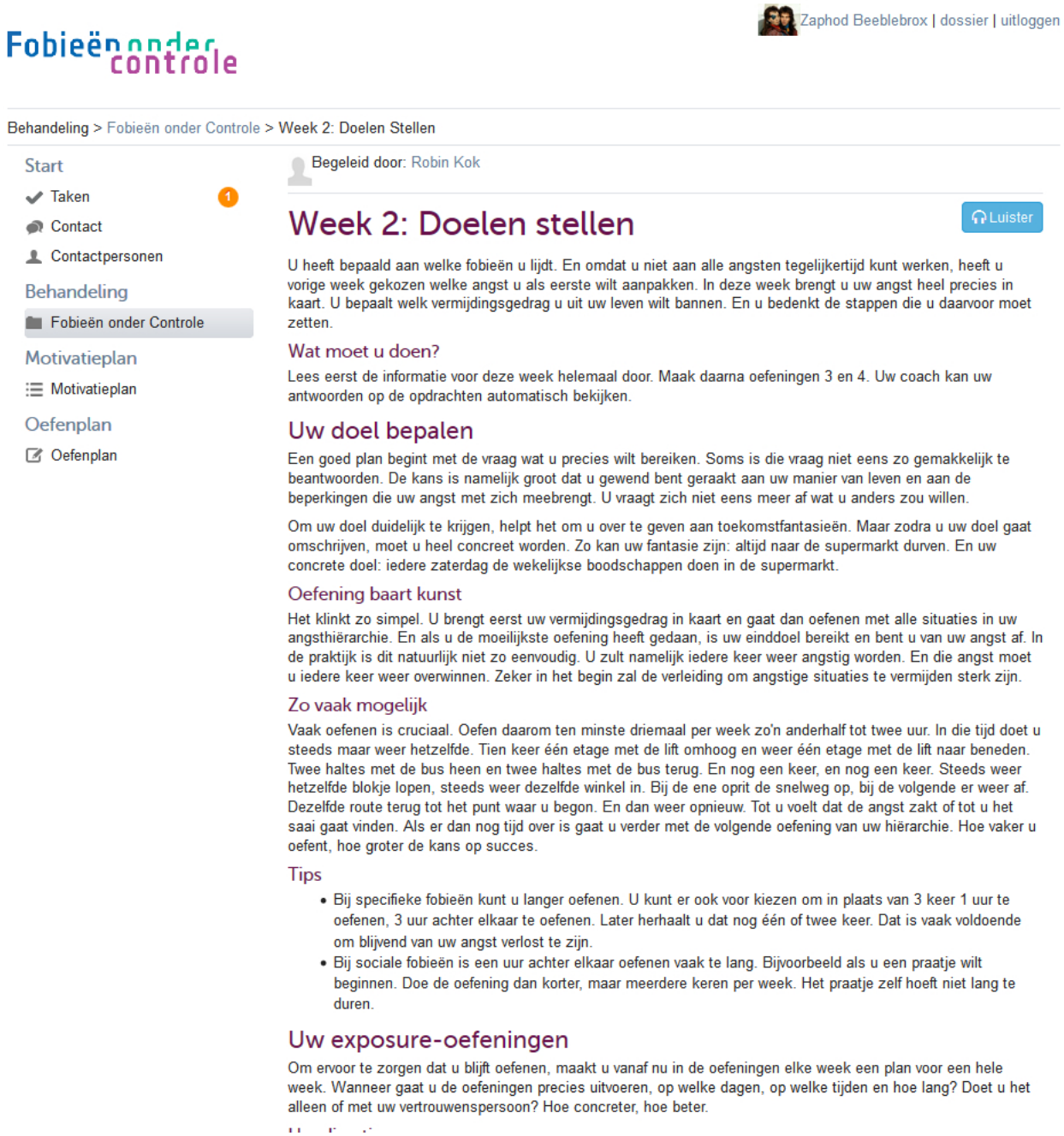


Figure 2. Screenshot of Phobias Under Control: intervention content.



Waiting List Condition

Participants in the waiting list group remained on the waiting list for face-to-face psychotherapy. Additionally, to comply with ethics committee regulations and to provide an incentive for enrolling in the trial, control group participants received a self-help book [40] based on exposure therapy, the de facto standard treatment in phobias. This book was sent to the control group participants free of charge with no instructions or support.

Assessments

All outcome measures were administered by phone at baseline (T0) and as self-assessment on the Internet at posttest (5 weeks from randomization, T1), using Web-based questionnaire software visibly associated to VU University Amsterdam. This relatively short period was selected to minimize the posttest assessments taking placing during face-to-face psychotherapy if face-to-face psychotherapy should incidentally take place earlier than 6 weeks after inclusion into the trial. To reduce study dropout, intensive reminder emails and telephone reminder calls were used for T1 assessments. Despite several email and telephone reminders, there was considerable variability in follow-up time (mean 50, SD 15.3 days), yet there was no significant difference in follow-up time between the intervention and wait-list control groups. All trial data were stored on a secured network complying with Dutch safety and privacy standards at the time of inclusion and accessible only to research staff. Data were anonymized as soon as possible.

Outcomes

Outcome measures are described in more detail elsewhere [38].

Primary Outcome Measures

Fear Questionnaire

The primary outcome measure was the Fear Questionnaire (FQ) [42]. This instrument measures severity of fear and avoidance of phobic stimuli. The psychometric validity of the FQ has been established for the Dutch version [43]. Internal consistency was good, with Cronbach alpha ranging from.78 (blood-injection-injury subscale) to 0.84 (total score).

Secondary Outcome Measures

Anxiety

The Beck Anxiety Inventory (BAI) [44] is a 21-item self-report questionnaire that focuses primarily on physiological manifestations of anxiety. The BAI has been validated for patients with agoraphobia [45] and other anxiety disorders [46]. Internal consistency was excellent (Cronbach α=.92).

Depressive Symptoms

The Center for Epidemiological Studies-Depression Scale (CES-D) [47] was administered as a self-rated questionnaire on the Internet. A Dutch version of the CES-D has been validated in an Internet-administrated form [48]. Internal consistency in this sample was good (Cronbach α=.70).

Process Outcome Measures

Adherence

Following a recent definition of intended usage [49], we defined *intervention adherence* as "the extent to which individuals should experience the content (of the intervention) to derive maximum benefit." Because some exercises were deemed to have a larger impact on lowering symptom severity (eg, reporting on performing exposure exercises is more beneficial than filling in a readiness to change questionnaire), different weights were assigned to different exercises accordingly, to a total of 20% for each of the 5 weeks. The intended usage was defined as 100% (ie, finishing the 8 exercises the participants were supposed to finish in 5 weeks). The main use metric was having finished an exercise as verified by the coach.

Treatment Satisfaction

Satisfaction with the Internet intervention or the self-help book was evaluated using the Client Satisfaction Questionnaire (CSQ-8) [50] which has been validated for use in a Dutch population [51]. Internal consistency was good (Cronbach α=.85). A few free-text items on participant satisfaction and experiences specific to the Web-based intervention were added.

Sample Size

To obtain 90% statistical power with a 2-sided alpha equal to .05 and assuming a mean standardized effect size (Cohen's *d*) of 0.7 in the intervention group and 0.2 in the control group, we calculated that 170 participants were needed to establish a clinical effect of the Internet intervention compared to wait-list controls. Assuming a dropout rate of 30% at 1-year follow-up, 244 participants should be included.

Randomization

A computer-generated randomization table was prepared by a researcher not involved in the data collection (AvS). Randomization was stratified at clinic level and performed at a 1:1 ratio. To ensure approximately equal randomization ratios per clinic, blocks of 8 were used. An external researcher not involved in the project supervised a list of sequentially numbered allocations and assigned participants to the conditions. All project members involved in data collection were unaware of allocation status until randomization was definitive. Participants were enrolled by a master's level research assistant.

Blinding

Due to the nature of this trial, neither participants nor researchers could be blinded to treatment allocation. All outcome measures are self-report questionnaires, which makes blinding unnecessary.

Statistical Methods

Data were analyzed with SPSS for Windows, version 20 (IBM Corp, Armonk, NY, USA) according to the intention-to-treat (ITT) principle. Using the multiple imputation function implemented in SPSS 20, we imputed missing data at posttest yielding 50 imputed datasets with 50 iterations each using the multiple imputation option with predictive mean matching. Predictors for the imputing procedure were pretest and (nonmissing) posttest scores, as well as age, clinic, education level, gender, randomization status, and quality of life at pretest. Because SPSS does not automatically calculate pooled statistics for imputed datasets when using ANCOVA, we calculated these statistics by pooling the saved residuals from each imputed dataset and reported mean values and 95% confidence intervals. Between-group effects on the primary outcome measure (FQ) at posttest were calculated with an ANCOVA, with baseline scores of the FQ, BAI, and CES-D entered as a covariate. Within- and between-group effect sizes were reported as Cohen's *d*. Effect sizes of *d*=0.2 are interpreted as small, effect sizes between 0.2 and 0.5 are interpreted as moderate, and effect sizes of 0.8 and upwards are interpreted as large. Due to the large amount of missing data at posttest, data are also presented separately for participants with full follow-up information. No interim analyses were performed. No stopping guidelines were postulated.

Changes to Protocol

There were no changes from the published study protocol [38].

Results

Sample

See Figure 3 for a flowchart and overview of participants in this trial. Of 481 participants assessed for eligibility, 212 were randomized to either intervention (n=105) or control (n=107). Baseline data are presented in Table 1. Apart from psychotropic medication use, there were no significant differences between intervention and control groups at baseline. Participants were mainly Dutch, female, and highly educated.

Table 1. Baseline characteristics of participants.

Characteristics	Total sample (N=212)	Intervention (n=105)	Control (n=107)	*P* [a]
Demographics				
Age, mean (SD)	34.6 (11.7)	35.7 (11.7)	33.4 (11.6)	.09
Female, n (%)	130 (61.0)	58 (55)	72 (67)	.10
Higher education,[b] n (%)	120 (57.1)	58 (56)	62 (58)	.69
Disposable income (€),[c] mean (SD)	1524 (761)	1545 (727)	1503 (796)	.70
Dutch parents, n (%)				
Both parents Dutch	144 (68.6)	72 (69)	72 (689)	.98
One parent Dutch	23 (11)	11 (11)	12 (11)	
Neither parent Dutch	43 (21)	21 (20)	22 (21)	
Psychotropic medication, n (%)	43 (20)	14 (13%)	29 (27)	.01
Baseline scores, mean (SD)				
FQ	40.28 (22.71)	42.43 (23.41)	38.19 (21.93)	.24
BAI	44.81 (13.41)	45.15 (13.76)	44.48 (13.13)	.81
CES-D	24.82 (8.47)	24.96 (8.61)	24.69 (8.37)	.84
CIDI phobia diagnoses [d]				
Specific phobia, n (%)				
Animal-type	19 (9)	9 (9)	10 (9)	.84
Nature-type	35 (17)	21 (20)	14 (13)	.18
Blood-injection-injury	51 (24)	25 (24)	26 (24)	.93
Situational-type	72 (34)	38 (36)	34 (32)	.50
Agoraphobia without panic disorder	36 (17)	18 (17)	18 (17)	.95
Agoraphobia with panic disorder	87 (41)	48 (45)	40 (37)	.28
Social phobia	113 (53.3)	51 (49)	62 (58)	.17
Number of phobias, mean (SD)	1.95 (1.08)	1.99 (1.11)	1.91 (1.06)	.57
Number of phobias, n (%)				
1 phobia	91 (43)	43 (41)	48 (45)	
2 phobias	61 (29)	32 (31)	29 (27)	
≥3 phobias	60 (28)	30 (29)	30 (28)	

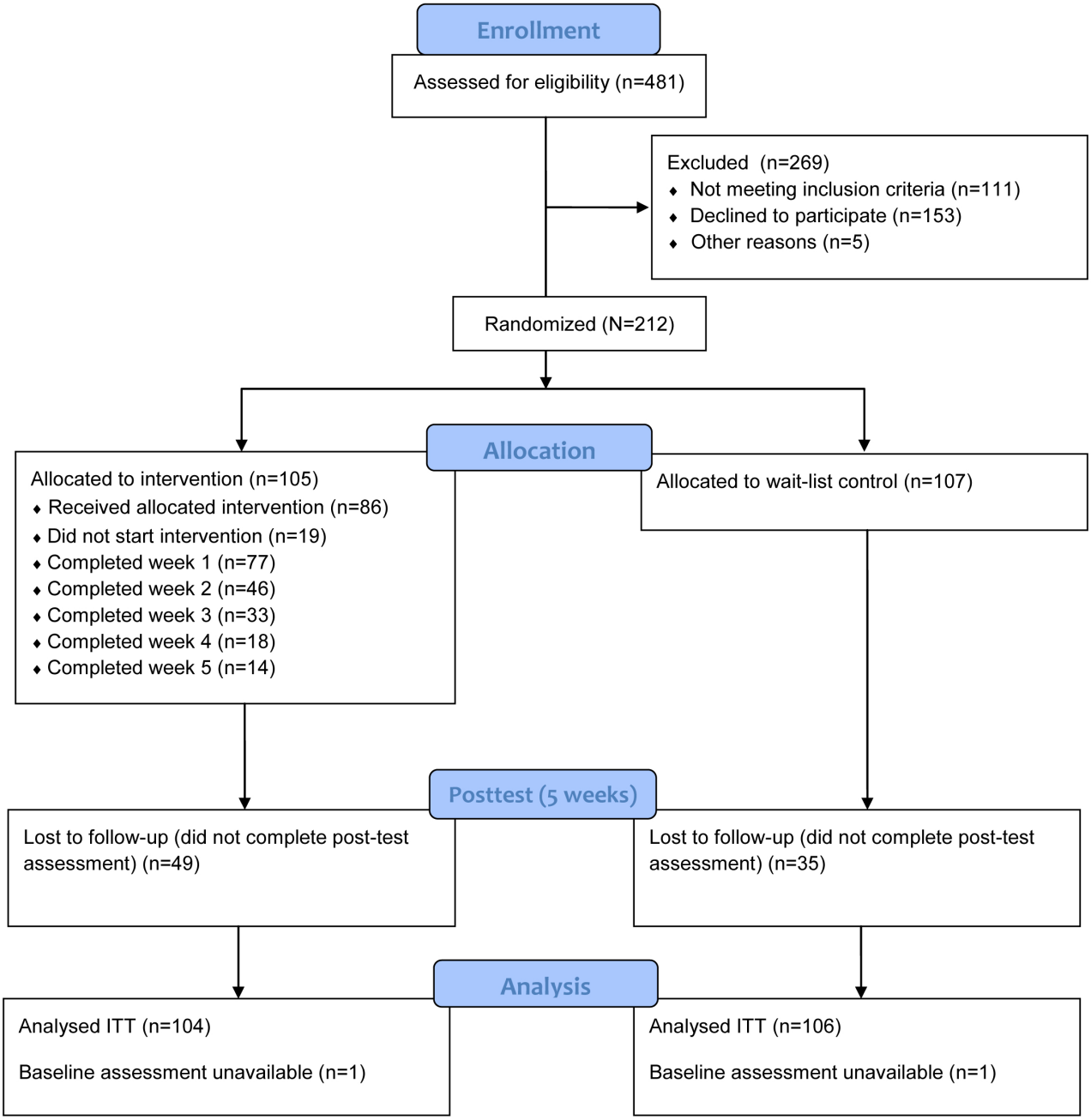
[a] Tested with *t* test or chi-square test as appropriate.

[b] Equivalent to a Bachelor's degree or higher.

[c] N=180.

[d] Percentages add up to over 100% due to multiple possible diagnoses per participant.

Figure 3. Participant flowchart.



Study Dropout

Posttreatment assessments were completed by 126 of 212 (59.2%) participants. Due to database corruption, 2 pretreatment assessments were unavailable. We tested whether there were significant differences in all baseline characteristics described in Table 1 between those who completed the follow-up assessments and those who did not, with dropouts scoring significantly higher on the BAI (t_{210}=2.275, P=.02) and the CES-D (t_{210}=2.489, P=.01), and dropouts being younger in age (t_{210}=2.022, P=.04), less often highly educated (χ^2_{210}=8.1, P=.004), and taking psychotropic medication less often at the time of assessment (χ^2_{210}=35.4, P<.001). Intervention group participants were also more likely to be nonresponders (67%, 84/126 of completers were control group participants; χ^2_{210}=4.3, P=.04). Table 2 shows the statistically significant differences between nonresponders and study completers. The differences of other characteristics described in Table 1 were not statistically significant.

Table 2. Differences between dropouts and study completers.

Characteristics	Completer (n=126)	Dropout (n=84)	*P* [a]
Baseline CES-D score, mean (SD)	23.70 (8.34)	26.63 (8.39)	.01
Baseline BAI score, mean (SD)	43.08 (13.54)	47.35 (12.96)	.02
Age, mean (SD)	35.83 (11.82)	32.52 (11.35)	.04
Higher education, n (%)	82 (65%)	38 (45%)	.004
On psychotropic medication, n (%)	43 (34%)	0 (0%)	.<001

[a] Tested with *t* test or chi-square test as appropriate.

Intervention Adherence and Satisfaction

Of the 105 participants, 78 (76.4%) started using the intervention, 14 (13.3%) finished week 5, and 9 (8.8%) met the intended usage of 100% (all 8 exercises) in 5 weeks. Average adherence, as expressed as percentage of intended usage, was 37.5% (SD 30.7%); median number of exercises completed (out of a possible 8) was 3 (IQR 4.0). As found in a previous meta-analysis [52], higher education in this sample was associated with higher adherence ($F_{2,103}$=8.132, P=.005). Intervention participants were moderately positive about their coach (average grade 6.8 of 10, SD 1.06) and indicated that the quality of the feedback messages was satisfactory (10/43, 23%), good (18/43, 42%), or very good to excellent (13/43, 30%). The number of messages received was also evaluated as being balanced (not too many, not too few) by most participants (32/43, 74%). Mean scores for all 8 CSQ-8 items were acceptable (mean 2.78, SD 0.58-0.81; possible item range 1-4).

Completers and Intention-to-Treat Analyses

After imputing missing values at posttest and correcting for baseline scores of the FQ, BAI, and CES-D, ANCOVA showed a significant difference in FQ scores between intervention and control groups at posttest ($F_{2,208}$=6.327, 95% CI 5.977-6.686; P=.02, 95% CI .01-.02; partial η^2=.030, 95% CI .03-.03). ANCOVAs also showed a significant difference in CES-D scores between intervention and control groups at posttest ($F_{2,208}$=6.121, 95% CI 5.550-6.669; P=.03, 95% CI .02-.03; partial η^2=.029, 95% CI .03-.03), but no significant difference in BAI scores between intervention and control groups at posttest ($F_{2,208}$=4.097, 95% CI 3.818-4.376; P=.05, 95% CI .04-.06; partial η^2=.020, 95% CI .02-.02). Changes in scores are presented in Table 3 and presented graphically in Figure 4.

Table 3. Main results, imputed intention-to-treat sample (N=210).

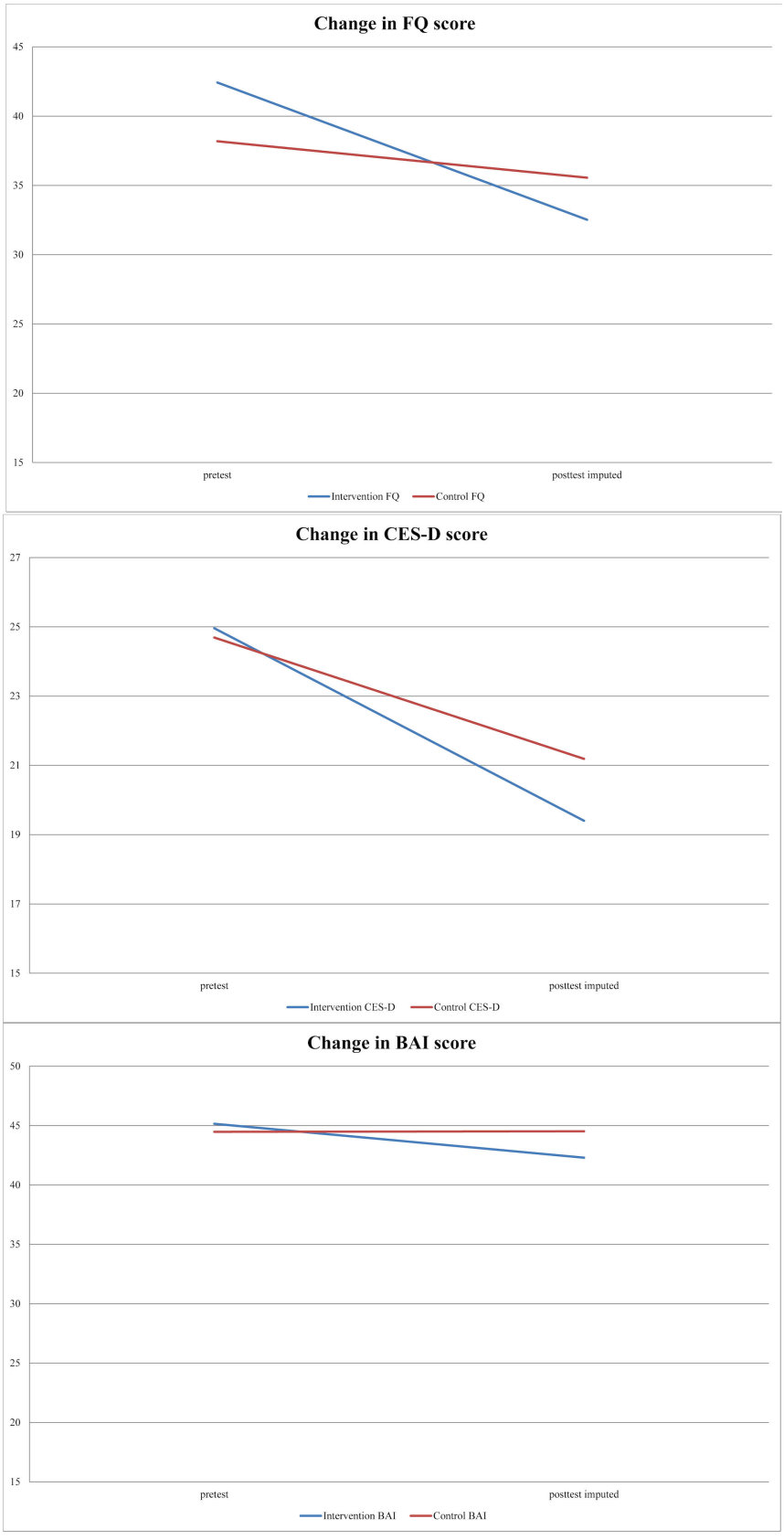
Results	Pretest, mean (SD)	Posttest ITT, mean (SD)	ANCOVA[a]		Effect size, *d* (95% CI)		NNT[b]
			$F_{2,208}$	*P*	Within group	Between group	
Intervention							
FQ	42.02 (23.39)	32.52 (18.48)	6.33	.16	0.42 (0.31, 0.52)	0.35 (0.07, 0.62)	5.10
CES-D	24.99 (8.58)	19.40 (5.97)	6.12	.26	0.75 (0.51, 0.97)	0.34 (0.07, 0.61)	5.26
BAI	45.01 (13.78)	42.30 (10.68)	4.10	.53	0.22 (0.03, 0.40)	0.28 (0.01, 0.55)	6.41
Wait-list control							
FQ	38.34 (21.97)	35.56 (18.41)	—	—	0.13 (0.03, 0.23)	—	—
CES-D	24.75 (8.39)	21.19 (7.06)	—	—	0.46 (0.27, 0.65)	—	—
BAI	44.57 (13.16)	44.52 (10.79)	—	—	0.00 (-0.14, 0.15)	—	—

[a] Controlled for baseline scores.

[b] Number needed to treat.

After correcting for baseline scores and age (full follow-ups only), we found no significant differences in posttest scores between intervention and control groups for FQ ($F_{2,208}$=.137, P=.71, partial η^2=.001), CES-D ($F_{2,208}$=2.086, P=.15, partial η^2=0.017), or BAI ($F_{2,208}$=0.333, P=.57, partial η^2=0.003).

Figure 4. Change in scores between baseline and imputed posttest scores.



Adverse Events

No adverse events were reported or observed during the trial. In a recent trial, a few participants reported temporary adverse effects in an Internet-based intervention for social anxiety disorder [53]; in particular, exacerbation of anxiety symptoms and negative well-being. One control group participant (1/107, 0.5%) reported a worsening of complaints after having started the intervention. However, it is unknown whether this was due

to the intervention or to other circumstances not related to this study.

Discussion

Principal Findings

Results show that an Internet-based guided intervention for phobic outpatients can be effective, with modest but significant effect sizes (d=0.28-0.35). The between-group effect sizes were small but significant in the imputed sample. Interestingly, effect sizes were similar for the FQ and CES-D. Because spontaneous recovery is common in depression [54] but not in phobia [5,6], it is unlikely that the decrease in phobic complaints can be attributed to spontaneous recovery rather than to the intervention. The intervention was not targeted specifically at depression; therefore, the significant decrease in depressive symptoms as measured by the CES-D might be attributable to spontaneous recovery rather than the intervention, although due to randomization this spontaneous recovery should probably have occurred in the control condition as well. Another possible explanation is the decrease in phobic complaints in the intervention group led to a commensurate decrease in depressive symptoms. Because the control group participants received a self-help book as compensation for their time invested in completing the baseline and follow-up assessments, there is a possibility that the between-group effect sizes are a more conservative estimate of the real effect size. Additionally, effect sizes for the BAI were lowest of all 3 primary outcome measures. This may be a result of the BAI being concerned mostly with those physical sensations often associated with panic attacks rather than phobias.

Comparison to Earlier Literature

Compared with earlier research, recruitment of patients in the outpatient clinics seemed to be reasonably successful. Of the 481 patients assessed for eligibility, of 370 eligible patients we randomized 212 of the 244 participants intended. In all, 153 patients (31.8%) declined to participate in the study (before being screened) and 111 (23.1%) did not meet inclusion criteria. In total, we were able to include 44.1% of all referrals with approximately 2 patients needed to be screened to include a single participant and approximately 3 hours required for each included participant. This does, however, not include the amount of time needed at the outpatient clinic for administrative tasks. In contrast, a recent study of computerized CBT for anxiety in secondary care [55] managed to randomize only 8% (88 of 1141) of referrals. The recruitment percentage also compares rather favorably to trial recruitment in primary care, which can be problematic in terms of duration and numbers of participants recruited as compared to the planned recruitment time and numbers of participants [56]. For example, 1 study focusing on Internet-delivered CBT for depression managed to recruit only 7 patients from 11 general practices in 8 months [57], a marked contrast to our trial. This may be a result of keeping close contact with outpatient staff, although it is difficult to draw comparisons because similar studies with tight integration in routine outpatient clinics are scarce and health care systems differ on national levels. When comparing effect sizes of the current trial with earlier trials of Internet-based anxiety and phobia interventions, we found that the effect sizes of the current trial are low overall [19,58]. In face-to-face psychotherapy, effect sizes of psychological treatment for participants meeting diagnostic criteria were found to be lower than those not meeting the criteria (eg, [13]), but another study found large and sustained effects of Internet-based CBT for panic disorder with and without agoraphobia in routine psychiatric care [33]. Although these previous trials used CBT compared to exposure therapy, it should be noted that the intervention of the current trial included a number of CBT components and that there is no unambiguous evidence that CBT should outperform exposure therapy for phobias per se, with studies finding either no difference or a negligible advantage for either CBT of exposure therapy in phobias (eg, [14,19,59,60]). There is an exception for social anxiety disorder, where cognitive therapies appear to outperform exposure therapy [13,61,62]. It should be noted that just over half of the participants in the current trial were diagnosed with social anxiety disorder, indicating that perhaps greater attention should have been paid to the cognitive therapy elements in the intervention. However, comorbidity of different phobias was large in the current trial, and participants suffering from multiple phobias were encouraged to focus on a single, well-circumscribed area of phobic avoidance at the start of the intervention. Many participants explicitly chose to work on specific phobias or agoraphobia rather than social anxiety, indicating that exposure therapy was indeed the right choice for these participants.

Overall, few (n=9) participants completed all 8 exercises, and 14 completed all 5 weeks, indicating that some exercises were skipped by the participants. Ending a treatment early may not necessarily be a negative finding [63]. Although not many participants stated a reason for not finishing the intervention, the primary reasons given were a lack of time (8 of 27) and the intervention not being suited to the needs of the participant (5 of 27), both commonly cited barriers to the uptake of online interventions [64]. Despite the high acceptability, as expressed by the low percentage of participants refusing to enter the trial outright, low adherence to this intervention remains a cause for concern because it seems to be lower than generally found in other Internet interventions for anxiety disorders [64]. Possible causes for the lower adherence in this intervention may be that the intervention was too broad, targeting all types of phobias within 1 intervention. Additionally, participants were aware that they would receive face-to-face psychotherapy regardless of whether they finished the online intervention, which may have lowered their motivation to persist. Furthermore, the intervention itself may not have been persuasive enough to encourage repeated use [49]. Earlier efforts to identify predictors of treatment dropout in social anxiety disorder, for example, have not yielded consistent results [65], and future research into the causes of premature treatment termination is needed.

A particular strength of this study is the high acceptance rate among outpatients as compared with other studies in outpatients and specialized health care centers, which indicates that this sample is clinically relevant and that the results may generalize well across other outpatient samples using similar recruitment strategies. Implementation in routine practice would perhaps facilitate better uptake due to dropping the constraints

surrounding research-oriented RCT setting (eg, randomization, filling in extra questionnaires).

Limitations

A number of limitations should be taken into account when interpreting the results from this trial. Firstly, the number of participants fell slightly short from the target number of participants (212 randomized versus 244 targeted). Far-reaching cutbacks in Dutch mental health care during the recruitment into this trial resulted in a dwindling number of patients seeking help with outpatient clinics, effectively shrinking the overall participant pool. Secondly, recruitment through outpatient clinics depends on outpatient clinic staff and may be liable to selection bias. Although the included sample of participants seems relatively representative of a clinical sample, selection bias may have occurred during the outpatient clinics' own selection procedures over which we had no influence. Thirdly, although we corrected for missing values at follow-up by using multiple imputation, the results should be interpreted with caution due to the large amount of missing data. Imputation is an approximation based on a combination of chosen predictors, techniques, and imputing algorithms, which may yield varying results in different datasets [66], making extensive sensitivity analyses time consuming and inconclusive at best. Although some argue that using covariates yields similar results to imputing [67], this may depend heavily on the dataset, and multiple imputation remains the solution of choice for missing data [68]. Regardless of the method for accounting for missing data, the large amount of missing data in this trial is a limitation and means that results should be interpreted with caution. Thirdly, there was considerable variability in the time between baseline assessment and follow-up (median 48 days, range 29-138 days). However, there were no significant correlations between baseline and posttest scores and the time in days between baseline assessment and posttest assessment, which indicates that the posttest scores as used are a representative assessment. Finally, offering the control group participants a self-help book may have influenced the between-group results because wait-list participants using the self-help book may have improved. However, because an improvement in control group participants would mean a smaller contrast between groups, this would lead to a more conservative estimate of treatment effects. Furthermore, it has been proposed that the use of a wait-list group design may not be a representative control group in that it is functionally different from a no-treatment group. It has also been put forward as actually being a "nocebo" group in that waitlisted participants actually do worse than no-treatment participants [69], which would theoretically lead to an inflation of between-group effect sizes. In the light of the current trial, however, both intervention and control groups were scheduled to receive face-to-face psychotherapy; as such, the arguments pertaining to possibly higher (or lower) effect sizes when using a wait-list control group do not necessarily apply to the current trial.

Implications and Future Research

In summary, adding an online guided intervention to routine face-to-face treatment may prove beneficial for outpatients, regardless of type of phobia diagnosis. However, effect sizes were markedly lower (d=0.28-0.35) than those found in research on psychological treatment for phobias in the general population (d=0.70-1.84) [12-14] and for anxiety in primary care (d=0.57) [70], and were only found for the imputed ITT sample. Because there was systematic attrition in this trial, the significant differences between completers and dropouts may provide valuable information to identify focal points for targeted attrition reduction, although previous efforts to identify nontechnological factors influencing attrition in online interventions have yielded inconclusive results [71-75]. Given the combination of high acceptability and low adherence, future research should focus on optimizing the usability and persuasive design of this intervention to improve retention and adherence [49] to maximize potential benefits of an intervention that efficiently uses the time spent waiting for face-to-face psychotherapy. Independent replication of the current results in different outpatient settings and countries is needed to verify the findings before robust inferences can be made, but the current results are promising.

Acknowledgments

We thank Manja de Neef, MA for coauthoring the intervention and providing professional supervision and additional input and feedback during the trial. We are very grateful for the cooperation of the trial participants and the participating outpatient clinics: PsyQ, GGZ inGeest, Prezens, GGZ Noord-Holland-Noord, and Altrecht. Furthermore, we thank Marleen Bakker, MSc; Julia van Oenen, MSc; and Kim Setkowski, BSc for their invaluable support in participant coaching and assessments, data collection, and administration. This study is funded by ZonMw (project number 80-82310-97-10015). The funding body had no influence in the design, execution, analysis, or interpretation of the results of this study.

Authors' Contributions

PC coauthored the intervention. PC and AvS drafted the study protocol. RK authored this paper. AvS, PC, and AB provided feedback and suggestions for this manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

AB has received grants for research from Eli Lilly, AstraZeneca, Janssen, and Shire and as a speaker from Lundbeck and Eli Lilly. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [76].

[PDF File (Adobe PDF File), 982KB - jmir_v16i9e226_app1.pdf]

References

1. Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 1998 Dec;33(12):587-595. [Medline: 9857791]
2. Mendlowicz MV, Stein MB. Quality of life in individuals with anxiety disorders. Am J Psychiatry 2000 May;157(5):669-682. [Medline: 10784456]
3. Konnopka A, Leichsenring F, Leibing E, König H. Cost-of-illness studies and cost-effectiveness analyses in anxiety disorders: a systematic review. J Affect Disord 2009 Apr;114(1-3):14-31. [doi: 10.1016/j.jad.2008.07.014] [Medline: 18768222]
4. 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]
5. Dingemans AE, van Vliet IM, Couvée J, Westenberg HG. Characteristics of patients with social phobia and their treatment in specialized clinics for anxiety disorders in the Netherlands. J Affect Disord 2001 Jul;65(2):123-129. [Medline: 11356235]
6. Green AC, Hunt C, Stain HJ. The delay between symptom onset and seeking professional treatment for anxiety and depressive disorders in a rural Australian sample. Soc Psychiatry Psychiatr Epidemiol 2012 Sep;47(9):1475-1487. [doi: 10.1007/s00127-011-0453-x] [Medline: 22116199]
7. Wittchen HU, Beloch E. The impact of social phobia on quality of life. Int Clin Psychopharmacol 1996 Jun;11 Suppl 3:15-23. [Medline: 8923105]
8. Zimmerman M, Mattia JI. Principal and additional DSM-IV disorders for which outpatients seek treatment. Psychiatr Serv 2000 Oct;51(10):1299-1304. [Medline: 11013331]
9. Dalrymple KL, Zimmerman M. Treatment-seeking for social anxiety disorder in a general outpatient psychiatry setting. Psychiatry Res 2011 May 30;187(3):375-381. [doi: 10.1016/j.psychres.2011.01.004] [Medline: 21310497]
10. Weiller E, Bisserbe JC, Boyer P, Lepine JP, Lecrubier Y. Social phobia in general health care: an unrecognised undertreated disabling disorder. Br J Psychiatry 1996 Feb;168(2):169-174. [Medline: 8837906]
11. Wittchen H. Social phobia: still neglected, poorly recognized and rarely treated? Acta Psychiatr Scand 2003 Sep;108(s417):3-3. [doi: 10.1034/j.1600-0447.108.s417.9.x]
12. Sánchez-Meca J, Rosa-Alcázar A, Marín-Martínez F, Gómez-Conesa A. Psychological treatment of panic disorder with or without agoraphobia: a meta-analysis. Clin Psychol Rev 2010 Feb;30(1):37-50. [doi: 10.1016/j.cpr.2009.08.011] [Medline: 19775792]
13. Acarturk C, Cuijpers P, van Straten A, de Graaf R. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol Med 2009 Feb;39(2):241-254. [doi: 10.1017/S0033291708003590] [Medline: 18507874]
14. Kok RN, Donker T, Batelaan N, Beekman ATF, van Straten A, Cuijpers P. Psychological treatment of specific phobias: a meta-analysis. Clinical Psychology Review 2015:1-1 (forthcoming)(forthcoming).
15. Marks IM, Kenwright M, McDonough M, Whittaker M, Mataix-Cols D. Saving clinicians' time by delegating routine aspects of therapy to a computer: a randomized controlled trial in phobia/panic disorder. Psychol Med 2004 Jan;34(1):9-17. [Medline: 14971623]
16. Carr AC, Ghosh A, Marks IM. Computer-supervised exposure treatment for phobias. Can J Psychiatry 1988 Mar;33(2):112-117. [Medline: 3365636]
17. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. Cogn Behav Ther 2009;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
18. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
19. Schneider AJ, Mataix-Cols D, Marks IM, Bachofen M. Internet-guided self-help with or without exposure therapy for phobic and panic disorders. Psychother Psychosom 2005;74(3):154-164. [doi: 10.1159/000084000] [Medline: 15832066]
20. Andersson G. Guided internet treatment for anxiety disorders. As effective as face-to-face therapies? Stud Health Technol Inform 2012;181:3-7. [Medline: 22954817]
21. 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]
22. Kiropoulos LA, Klein B, Austin DW, Gilson K, Pier C, Mitchell J, et al. Is internet-based CBT for panic disorder and agoraphobia as effective as face-to-face CBT? J Anxiety Disord 2008 Dec;22(8):1273-1284. [doi: 10.1016/j.janxdis.2008.01.008] [Medline: 18289829]

23. Andersson G, Waara J, Jonsson U, Malmaeus F, Carlbring P, Ost LG. Internet-based self-help versus one-session exposure in the treatment of spider phobia: a randomized controlled trial. Cogn Behav Ther 2009;38(2):114-120. [doi: 10.1080/16506070902931326] [Medline: 20183690]
24. Müller B, Kull S, Wilhelm FH, Michael T. One-session computer-based exposure treatment for spider-fearful individuals--efficacy of a minimal self-help intervention in a randomised controlled trial. J Behav Ther Exp Psychiatry 2011 Jun;42(2):179-184. [doi: 10.1016/j.jbtep.2010.12.001] [Medline: 21315879]
25. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther 2011 Mar;49(3):158-169. [doi: 10.1016/j.brat.2010.12.007] [Medline: 21255767]
26. Carlbring P, Nordgren LB, Furmark T, Andersson G. Long-term outcome of Internet-delivered cognitive-behavioural therapy for social phobia: a 30-month follow-up. Behav Res Ther 2009 Oct;47(10):848-850. [doi: 10.1016/j.brat.2009.06.012] [Medline: 19631312]
27. Hedman E, El Alaoui S, Lindefors N, Andersson E, Rück C, Ghaderi A, et al. Clinical effectiveness and cost-effectiveness of Internet- vs. group-based cognitive behavior therapy for social anxiety disorder: 4-year follow-up of a randomized trial. Behav Res Ther 2014 Aug;59:20-29. [doi: 10.1016/j.brat.2014.05.010] [Medline: 24949908]
28. Proudfoot J, Ryden C, Everitt B, Shapiro DA, Goldberg D, Mann A, et al. Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. Br J Psychiatry 2004 Jul;185:46-54 [FREE Full text] [Medline: 15231555]
29. Hans E, Hiller W. A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. Clin Psychol Rev 2013 Dec;33(8):954-964. [doi: 10.1016/j.cpr.2013.07.003] [Medline: 23988455]
30. Kenter RM, van Straten A, Hobbel SH, Smit F, Bosmans J, Beekman A, et al. Effectiveness and cost effectiveness of guided online treatment for patients with major depressive disorder on a waiting list for psychotherapy: study protocol of a randomized controlled trial. Trials 2013;14:412 [FREE Full text] [doi: 10.1186/1745-6215-14-412] [Medline: 24289099]
31. Bell CJ, Colhoun HC, Carter FA, Frampton CM. Effectiveness of computerised cognitive behaviour therapy for anxiety disorders in secondary care. Aust N Z J Psychiatry 2012 Jul;46(7):630-640. [doi: 10.1177/0004867412437345] [Medline: 22327097]
32. Moessner M, Bauer S, Ozer F, Wolf M, Zimmer B, Kordy H. Cost-effectiveness of an internet-based aftercare intervention after inpatient treatment in a psychosomatic hospital. Psychother Res 2014 Jul;24(4):496-503. [doi: 10.1080/10503307.2013.845919] [Medline: 24188127]
33. Hedman E, Ljótsson B, Rück C, Bergström J, Andersson G, Kaldo V, et al. Effectiveness of internet-based cognitive behaviour therapy for panic disorder in routine psychiatric care. Acta Psychiatr Scand 2013 Dec;128(6):457-467. [doi: 10.1111/acps.12079] [Medline: 23406572]
34. Marks IM, Kenwright M, McDonough M, Whittaker M, Mataix-Cols D. Saving clinicians' time by delegating routine aspects of therapy to a computer: a randomized controlled trial in phobia/panic disorder. Psychol Med 2004 Jan;34(1):9-17. [Medline: 14971623]
35. Hedman E, Andersson E, Ljótsson B, Andersson G, Rück C, Lindefors N. Cost-effectiveness of Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: results from a randomized controlled trial. Behav Res Ther 2011 Nov;49(11):729-736. [doi: 10.1016/j.brat.2011.07.009] [Medline: 21851929]
36. Titov N, Andrews G, Johnston L, Schwencke G, Choi I. Shyness programme: longer term benefits, cost-effectiveness, and acceptability. Aust N Z J Psychiatry 2009 Jan;43(1):36-44. [doi: 10.1080/00048670802534424] [Medline: 19085526]
37. Issakidis C, Andrews G. Pretreatment attrition and dropout in an outpatient clinic for anxiety disorders. Acta Psychiatr Scand 2004 Jun;109(6):426-433. [doi: 10.1111/j.1600-0047.2004.00264.x] [Medline: 15117287]
38. Kok RN, van Straten A, Beekman A, Bosmans J, de Neef M, Cuijpers P. Effectiveness and cost-effectiveness of web-based treatment for phobic outpatients on a waiting list for psychotherapy: protocol of a randomised controlled trial. BMC Psychiatry 2012;12:131 [FREE Full text] [doi: 10.1186/1471-244X-12-131] [Medline: 22937959]
39. World Health Organization. Composite International Diagnostic Interview (CIDI). Geneva: World Health Organization; 1990.
40. De Neef M, Cuijpers P. Fobieën. Amsterdam: Boom Uitgeverij; 2007.
41. Deacon BJ, Abramowitz JS. Cognitive and behavioral treatments for anxiety disorders: a review of meta-analytic findings. J Clin Psychol 2004 Apr;60(4):429-441. [doi: 10.1002/jclp.10255] [Medline: 15022272]
42. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. Behav Res Ther 1979;17(3):263-267. [Medline: 526242]
43. Van Zuuren FJ. The fear questionnaire. Some data on validity, reliability and layout. Br J Psychiatry 1988 Nov;153:659-662. [Medline: 3255455]
44. Beck AT, Steer RA. Beck Anxiety Inventory Manual. San Antonio, TX: Psychological Corperation; 1993.
45. de Beurs E, Wilson KA, Chambless DL, Goldstein AJ, Feske U. Convergent and divergent validity of the Beck Anxiety Inventory for patients with panic disorder and agoraphobia. Depress Anxiety 1997;6(4):140-146. [Medline: 9559283]

46. Leyfer OT, Ruberg JL, Woodruff-Borden J. Examination of the utility of the Beck Anxiety Inventory and its factors as a screener for anxiety disorders. J Anxiety Disord 2006;20(4):444-458. [doi: 10.1016/j.janxdis.2005.05.004] [Medline: 16005177]
47. 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]
48. Donker T, van Straten A, Marks I, Cuijpers P. Brief self-rated screening for depression on the Internet. J Affect Disord 2010 May;122(3):253-259. [doi: 10.1016/j.jad.2009.07.013] [Medline: 19679358]
49. 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]
50. Nguyen TD, Attkisson CC, Stegner BL. Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire. Eval Program Plann 1983;6(3-4):299-313. [Medline: 10267258]
51. de Brey H. A cross-national validation of the client satisfaction questionnaire: the Dutch experience. Eval Program Plann 1983;6(3-4):395-400. [Medline: 10267266]
52. Swift JK, Greenberg RP. Premature discontinuation in adult psychotherapy: a meta-analysis. J Consult Clin Psychol 2012 Aug;80(4):547-559. [doi: 10.1037/a0028226] [Medline: 22506792]
53. Boettcher J, Rozental A, Andersson G, Carlbring P. Side effects in Internet-based interventions for Social Anxiety Disorder. Internet Interventions 2014 Mar;1(1):3-11. [doi: 10.1016/j.invent.2014.02.002]
54. Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 2002 Sep;181:208-213 [FREE Full text] [Medline: 12204924]
55. Carter FA, Bell CJ, Colhoun HC. Suitability and acceptability of computerised cognitive behaviour therapy for anxiety disorders in secondary care. Aust N Z J Psychiatry 2013 Feb;47(2):142-152. [doi: 10.1177/0004867412461384] [Medline: 23047956]
56. Bower P, Wilson S, Mathers N. Short report: how often do UK primary care trials face recruitment delays? Fam Pract 2007 Dec;24(6):601-603 [FREE Full text] [doi: 10.1093/fampra/cmm051] [Medline: 17872907]
57. Woodford J, Farrand P, Bessant M, Williams C. Recruitment into a guided internet based CBT (iCBT) intervention for depression: lesson learnt from the failure of a prevalence recruitment strategy. Contemp Clin Trials 2011 Sep;32(5):641-648. [doi: 10.1016/j.cct.2011.04.013] [Medline: 21570485]
58. Reger MA, Gahm GA. A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety. J Clin Psychol 2009 Jan;65(1):53-75. [doi: 10.1002/jclp.20536] [Medline: 19051274]
59. Ost LG, Alm T, Brandberg M, Breitholtz E. One vs five sessions of exposure and five sessions of cognitive therapy in the treatment of claustrophobia. Behav Res Ther 2001 Feb;39(2):167-183. [Medline: 11153971]
60. Bornas X, Tortella-Feliu M, Llabrés J. Do all treatments work for flight phobia? Computer-assisted exposure versus a brief multicomponent nonexposure treatment. Psychotherapy Research 2006 Jan;16(1):41-50. [doi: 10.1080/10503300500091058]
61. Clark DM, Ehlers A, Hackmann A, McManus F, Fennell M, Grey N, et al. Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. J Consult Clin Psychol 2006 Jun;74(3):568-578. [doi: 10.1037/0022-006X.74.3.568] [Medline: 16822113]
62. Ougrin D. Efficacy of exposure versus cognitive therapy in anxiety disorders: systematic review and meta-analysis. BMC Psychiatry 2011;11:200 [FREE Full text] [doi: 10.1186/1471-244X-11-200] [Medline: 22185596]
63. Hilvert-Bruce Z, Rossouw PJ, Wong N, Sunderland M, Andrews G. Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. Behav Res Ther 2012 Aug;50(7-8):463-468. [doi: 10.1016/j.brat.2012.04.001] [Medline: 22659155]
64. 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]
65. Eskildsen A, Hougaard E, Rosenberg NK. Pre-treatment patient variables as predictors of drop-out and treatment outcome in cognitive behavioural therapy for social phobia: A systematic review. Nord J Psychiatry 2010 Apr;64(2):94-105. [doi: 10.3109/08039480903426929] [Medline: 20055730]
66. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
67. Groenwold RH, Donders AR, Roes KC, Harrell FE, Moons KG. Dealing with missing outcome data in randomized trials and observational studies. Am J Epidemiol 2012 Feb 1;175(3):210-217 [FREE Full text] [doi: 10.1093/aje/kwr302] [Medline: 22262640]
68. Liublinska V, Rubin DB. Re: "dealing with missing outcome data in randomized trials and observational studies". Am J Epidemiol 2012 Aug 15;176(4):357-8; author reply 358 [FREE Full text] [doi: 10.1093/aje/kws215] [Medline: 22850795]
69. Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, et al. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatr Scand 2014 Sep;130(3):181-192. [doi: 10.1111/acps.12275] [Medline: 24697518]

70. Seekles W, Cuijpers P, Kok R, Beekman A, van Marwijk H, van Straten A. Psychological treatment of anxiety in primary care: a meta-analysis. Psychol Med 2013 Feb;43(2):351-361. [doi: 10.1017/S0033291712000670] [Medline: 22717105]
71. Blankers M, Koeter MW, Schippers GM. Baseline predictors of treatment outcome in Internet-based alcohol interventions: a recursive partitioning analysis alongside a randomized trial. BMC Public Health 2013;13:455 [FREE Full text] [doi: 10.1186/1471-2458-13-455] [Medline: 23651767]
72. Wojtowicz M, Day V, McGrath PJ. Predictors of participant retention in a guided online self-help program for university students: prospective cohort study. J Med Internet Res 2013;15(5):e96 [FREE Full text] [doi: 10.2196/jmir.2323] [Medline: 23697614]
73. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
74. Karyotaki E, Smit F, Kleiboer A, Mira Pastor A, Berger T, Cristina B. Predictors of treatment dropout in self-guided web-based interventions for depression: An individual patient data meta-analysis. Psychol Med 2015:1-1 (forthcoming)(forthcoming).
75. Al-Asadi AM, Klein B, Meyer D. Pretreatment attrition and formal withdrawal during treatment and their predictors: an exploratory study of the anxiety online data. J Med Internet Res 2014;16(6):e152 [FREE Full text] [doi: 10.2196/jmir.2989] [Medline: 24938311]
76. 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

BAI: Beck Anxiety Inventory
CBT: cognitive behavioral therapy
CIDI: Composite International Diagnostic Interview
FQ: Fear Questionnaire
GP: general practitioner
IQR: Interquartile range
ITT: intention-to-treat
NNT: number needed to treat
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 26.03.14; peer-reviewed by T Berger, C Lewis; comments to author 24.06.14; revised version received 02.07.14; accepted 17.07.14; published 29.09.14.

Please cite as:
Kok RN, van Straten A, Beekman ATF, Cuijpers P
Short-Term Effectiveness of Web-Based Guided Self-Help for Phobic Outpatients: Randomized Controlled Trial
J Med Internet Res 2014;16(9):e226
URL: http://www.jmir.org/2014/9/e226/
doi:10.2196/jmir.3429
PMID:25266929

Viewpoint

mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use

Thomas Lorchan Lewis[1*], BSc (Hons), MB ChB; Jeremy C Wyatt[2*], BA, MBBS, FRCP, DM, FACMI

[1]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[2]Leeds Institute of Health Sciences, Faculty of Medicine, Health & Psychology, University of Leeds, Leeds, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Thomas Lorchan Lewis, BSc (Hons), MB ChB
Warwick Medical School
University of Warwick
Medical Teaching Centre
Gibbet Hill Road
Coventry, CV4 7AL
United Kingdom
Phone: 44 7876453511
Fax: 44 2476528375
Email: tlewis@doctors.org.uk

Abstract

The use of mobile medical apps by clinicians and others has grown considerably since the introduction of mobile phones. Medical apps offer clinicians the ability to access medical knowledge and patient data at the point of care, but several studies have highlighted apps that could compromise patient safety and are potentially dangerous. This article identifies a range of different kinds of risks that medical apps can contribute to and important contextual variables that can modify these risks. We have also developed a simple generic risk framework that app users, developers, and other stakeholders can use to assess the likely risks posed by a specific app in a specific context. This should help app commissioners, developers, and users to manage risks and improve patient safety.

(J Med Internet Res 2014;16(9):e210) doi:10.2196/jmir.3133

KEYWORDS

risk assessment; medical app; mobile health; mobile phone; patient safety; smartphone; mHealth; medical informatics applications

Introduction

Overview

The use of mobile medical apps by clinicians, patients, and others has grown dramatically since the introduction of mobile phones and tablet computers. Recent studies show that mobile devices and apps can support a variety of routine medical tasks including clinical reference, drug dose calculation, patient education, accessing medical records, and clinical decision support [1-4]. Mobile phone apps have also been shown to benefit patients in a range of interventions across numerous medical specialties and treatment modalities [5-9]. Medical apps offer clinicians the ability to access medical knowledge and patient data at the point of care with unprecedented ease. However, the intersection of mobile technology, apps, and health care is currently in its most dynamic phase, meaning that there is a need to ensure that patient safety is not compromised before this field matures. For the purposes of this paper, a mobile medical app means any software application created for or used on a mobile device for medical or other health-related purposes. This paper highlights the need for risk assessment to support clinical use of mobile medical apps by critically appraising the existing literature in this field. We identify the different types of risks to which medical apps can contribute and develop a framework that brings together the usage scenarios, contextual factors, and app complexity to estimate the overall probability and severity of harm resulting from use of a mobile medical app.

Evidence of Unsafe Apps

It is important that mobile medical apps used in health care settings are accurate and reliable, especially as health care professionals and patients may make critical decisions based on information from an app. There is limited literature that addresses the accuracy of mobile medical apps, and that which

exists is often highly specialized and not necessarily generalizable to all medical apps [10]. Despite this, several studies have highlighted a number of medical apps that can compromise patient safety and are potentially dangerous in clinical use. For example, certain apps designed for opioid dosage conversion or melanoma detection demonstrate dangerously poor accuracy, while a number of other medical apps do not follow evidence-based guidelines [11-14]. Such risks have led to recent calls for increased regulation before further use and adoption of some apps in clinical practice [15-17]. One issue highlighted by a small number of studies is that many app developers have little or no formal medical training and do not involve clinicians in the development process and may therefore be unaware of patient safety issues raised by inappropriate app content or functioning [18-20]. Another issue is the sheer volume and exponential growth of medical apps, meaning it is practically impossible to assess each and every medical app [21]. The narrow scope of the current evidence base means it is difficult to generalize these statements to all medical and health-related apps. There is sufficient evidence that a small subsection of medical apps presents a risk to patient safety, and therefore it is appropriate to develop a model to help assess these risks.

Regulatory Oversight

Clinicians trying to safely navigate the apps minefield have had relatively little support from regulatory agencies. The Food and Drug Administration (FDA) released their guidance only in July 2013 after a 2-year consultation period and are focusing primarily on apps that transform the mobile platform into a regulated medical device [22], which to date numbers approximately 100 apps [23]. The remainder will be subject to what the FDA calls “enforcement discretion”, that is, no regulation [24]. Other regulatory agencies such as the Medicines and Healthcare Products Regulatory Agency and the Therapeutic Goods Administration of Australia have offered limited guidance to health care practitioners by including apps under their existing regulations for medical devices [25,26]. The lack of clarity regarding when a medical app becomes a formal medical device means that many developers may not recognize that their app requires formal regulation. As a result, the vast majority of medical apps remain without any form of regulation or safety check, and some of these may present a patient safety or other risk.

The Need for a Risk Framework to Support Clinical Use of Medical Apps

To inform the safe clinical use of apps and future professional guidance and regulation, it is important to understand and then quantify the different kinds of risk posed by medical apps. It is generally accepted that two dimensions define risk [27]: (1) the probability of an event occurring that could lead to harm, and (2) the severity of the harm that is likely to follow that event.

As with many aspects of medicine, the decision to use a medical app in a particular clinical context relies on our ability to assess the risk of harm and balance it against the anticipated benefits. These judgments require health care professionals to understand the intended benefits, limitations, and risks associated with medical apps in order to make an informed app usage decision. The first step in this process is to identify the different types of risk to which medical apps can contribute, summarized (in broadly increasing order of severity) in Table 1.

There is currently no clinically relevant risk assessment framework for medical apps, so health care practitioners, patients, and app developers find it challenging to quickly assess the risks posed by a specific app. In order to develop a comprehensive risk assessment framework, and to distinguish the different kinds of risk listed in Table 1, we must understand the key variables that can influence risk in medical apps. These variables can be broken down into those risk factors that are inherent to an app and those that depend on the external context where the app is used. Risk factors inherent to an app may be reduced through appropriate regulation, while managing contextual risk factors may require a formal education program to raise awareness among app users. In our opinion, the main contextual and inherent app risk factors are listed in Table 2 below, in no particular order. Arguably many of these risk variables are applicable to many other sources of medical information such as websites or textbooks, although there are important considerations specific to mobile apps that should be recognized.

Table 1. Different types of risk that medical use of apps may contribute to, and scenarios where these may arise.

Type of risk in increasing order of severity	Main stakeholder affected	Sample scenario where this risk could arise	What can be done to manage this risk
Loss of reputation	Professional/organization	App displays sensitive performance data about professional or service	Good security
Loss of privacy (patient confidentiality)	Patient	Poor security of patient data	Encryption
		Lose phone holding patient data	Avoid holding patient data on mobile device
Poor quality patient data	Patient/professional/ organization (eg, financial data)	App allows bad data to be entered into patient record or retrieved from it at handover	Data validation on entry and retrieval from authenticated source
Poor lifestyle or clinical decision	Patient/professional	Bad patient data used in risk calculation algorithm	Check correct data retrieved
		Bad knowledge or search tool	Check algorithm properly coded
		Bad advice or algorithm	Use proven health behavior change methods
		Poor risk communication	
Inappropriate but reversible clinical action	Patient/professional	Poor medication advice	Test quality of advice on sample data
			Provide facility for user feedback and respond to this
Inappropriate and irreversible clinical action	Patient/professional/ organization (liability exposure)	Bad algorithm controlling insulin pump, surgical robot, radiotherapy machine, etc	Adopt safety critical software design and development methods
			Exhaustively check design and test algorithm & user interface

Table 2. The main inherent and external (contextual) risk variables contributing to the total risk associated with mobile medical apps.

Type of risk variable	Specific risk variable	Explanation
Inherent to the app	Intended function	When the intended function of the app is inherently dangerous, eg, calculating insulin requirements or reprogramming a pacemaker, this will increase risk
	Inaccurate or out of date content	Apps that contain inaccurate or out-of-date content have an increased chance of causing harm
	Complexity of task supported by the app	Apps that carry out complex tasks (eg, drug dosage calculations) have greater potential for harm due to programming errors than simple information display
	Lack of feedback or failsafe mechanism	Apps that do not offer the user a means to report safety issues to the developers are less safe
External factors, depending on context of app use	App user	Use of the app by people other than those intended by the developer may cause harm
	Inappropriate app usage	Apps that are used inappropriately, outside their design envelope, are inherently risky
	Inadequate user training	Even when the app user is as the developer intended, risk can be increased if the user has inadequate training or knowledge to recognize when there is a patient safety hazard, eg, incorrect content or inappropriate advice from the app
	Likelihood of errors being detected	App usage in scenarios with a low error detection capacity (eg, community care versus intensive care) are likely to be riskier
	App usage factor (AUF)	Total number of app users multiplied by the average number of app uses per user per day. Apps with a high usage factor have a greater safety impact on the population than those with a low usage factor

The last two contextual factors are discussed in more detail here. One is the likelihood of a clinical error being detected and averted, which should be high in a well-monitored inpatient or high dependency setting but low when there is only intermittent patient contact, such as in outpatient clinics or primary care. Paradoxically, therefore, the risk of using a faulty app may be lower in an intensive care unit than in general practice. The second is the app usage factor (AUF), which links app risk to the number of users and frequency of use. Risk is proportional to the number of patients affected, so disease prevalence or similar indices of the number of people likely to be affected by an error need to be considered. We have developed the idea of

the AUF to help estimate the risk impact of a particular app on a given population. It thus follows that a popular app with a high number of frequent users will have a high AUF and subsequent high impact on the population.

It is also important to consider the generic clinical safety hazards posed by the hardware, software, and sensors that make up a typical medical software application, not just mobile apps. This includes risks posed by the display, user interface, network issues, and subsequent loss of information. Each of these factors should be taken into account, so that the more complex the app, the greater the risk. Unfortunately, these risks are difficult to assess without formal training, but there is guidance for health organizations and developers that aims to address these factors in more detail [28].For the purposes of our risk assessment framework, these factors have been included within the Complexity of task variable.

Developing a formal risk assessment framework for mobile medical apps should enable us to reduce the "residual risk" (exposure to loss remaining after all other known risks have been countered, factored in, or eliminated) by recognizing and implementing a range of possible safety measures in future app development, procurement, and regulation models.

Bringing Together Usage Scenarios, Contextual Factors, and App Complexity to Estimate Overall Probability and Severity of Harm

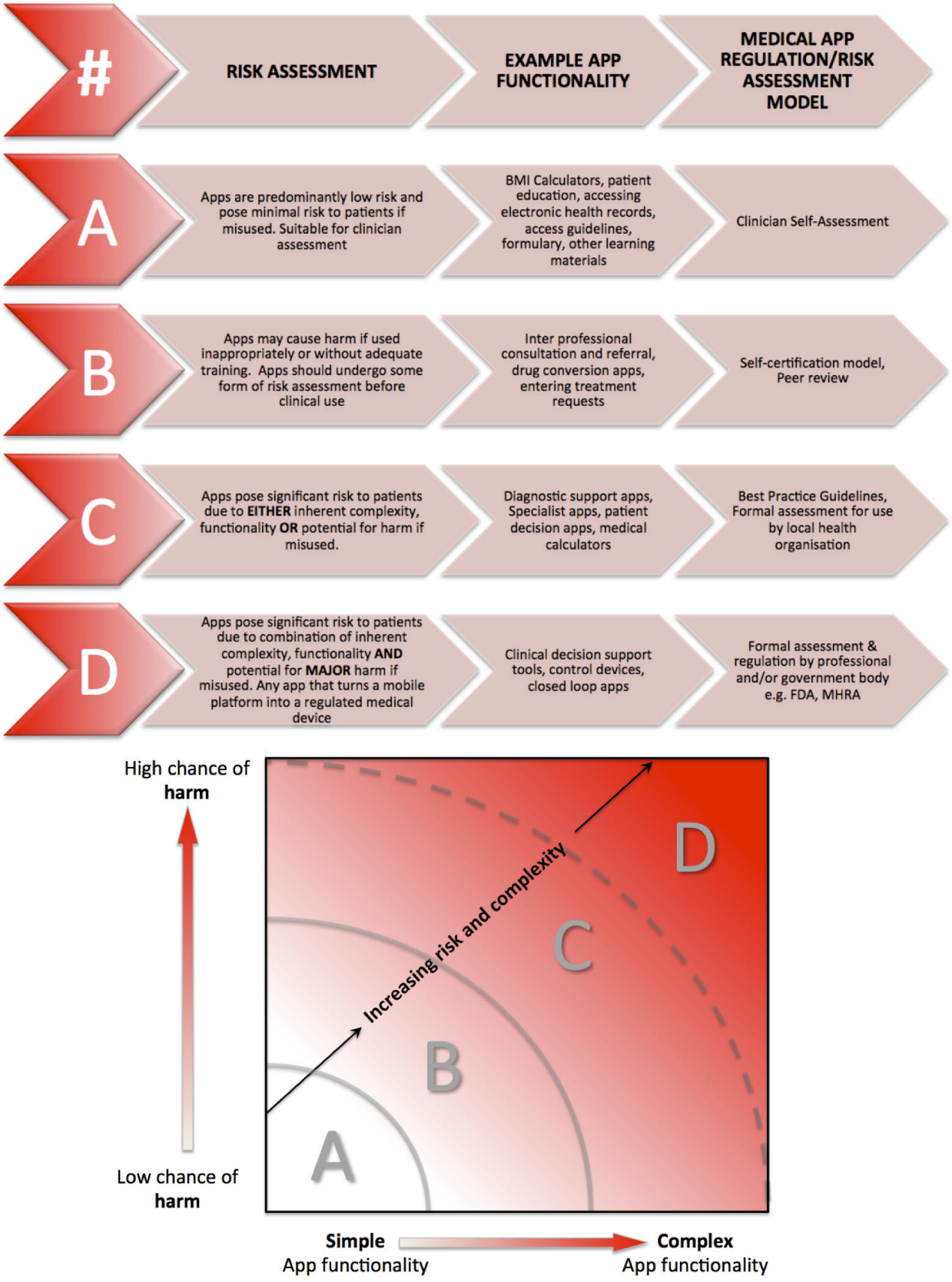
We believe that the risks posed by a specific medical app depend on three main dimensions: (1) the probability and the severity of harm, defined by the risk scenarios listed in Table 1, (2) the inherent complexity of the app, which determines how predictable that risk is, and (3) the external or contextual factors listed above.

Given the wide variety of medical apps, we believe that different approaches to risk assessment and management will be required dependent on app risk. This is illustrated in Figure 1, which shows a 2-dimensional "app-space" where an app can be located depending on its probability of harm, based on the variables above, and its complexity. According to its combined chances of harm and complexity, it will fall into one of four broad zones. Apps in Zone A require only local inspection, those in Zone B require a more formal risk assessment, and those in Zone C require professional review of a full safety case and the use of safety critical development methods. Apps that fall into Zone D should meet the criteria for formal regulation and review by governmental bodies such as the FDA due to their high probability of causing harm. It is not possible to assess the proportion of medical apps in each of the risk categories of A-C given the lack of data on medical apps available. However, based on the total number of medical apps available (approximately 20,000) [29] and the number currently regulated by the FDA (approximately 100) [23], we calculate that the proportion of apps that currently fall into risk category D is approximately 0.5%. This classification into four broad risk zones should help app users, developers, and regulators to evaluate each app using a relevant risk assessment and management model based on the zone where the app is located. It is important to note that these zones form a spectrum rather than discrete entities, hence the gray lines at the boundaries of each zone.

Perhaps the biggest threat to patient safety from medical apps is likely to result from inadequate education and knowledge of health care professionals and patients about their risks. We think in the vast majority of cases, it is probably the actions of a user resulting from a specific app that leads to harm, rather than the app itself. Therefore, an important additional strategy to minimize the risks posed by apps is to develop an educational program to raise awareness of potential patient safety and other risks following inappropriate app use. Developing a single, authoritative, coherent set of guidance and supporting educational materials will require the support of professional bodies such as the Royal Colleges. This will help avoid a confusing plethora of guidance, such as occurred when the harm resulting from some uses of social media was recognized.

In the meantime, there are a range of proposed app regulation models, many of which are highlighted in Figure 1, that may provide some form of protection against hazardous medical apps for patients and health care practitioners [30-33]. Many of these risk management methods are in the early stages of development and have not yet been formally implemented, but they offer a number of advantages for health care professionals, patients, and developers alike, offering some degree of safety check for medical apps not meeting the requirements for formal regulation. A detailed discussion of regulation and regulatory issues for mobile medical apps is beyond the scope of this paper, and interested readers are directed to the references above for further information.

Figure 1. Two-dimensional "App-space" for risk assessment of mobile medical apps with key suggesting appropriate models for app regulation.



Conclusions

While the widespread use of high-quality apps by health care practitioners and patients is to be welcomed, there still remains a significant potential for harm. The risks to patient safety and professional reputation are real, and steps should be taken to mitigate these. Identification of all the different kinds of risk and of key variables that influence risk are key stages in the development of a risk assessment model, which should also take into account app complexity and the probability of harm.

Education of current health practitioners about the risks posed by medical apps should start soon, before the first case reports of patients harmed by a medical app come to light. Further work should focus on the recognition and mitigation of medical app risk, as the outlook for medical apps in health care is bright once their quality and safety can be reliably assessed and managed.

Acknowledgments

We acknowledge helpful comments from the reviewers and the editor of JMIR. The authors would like to thank Warwick University for kindly covering the article processing fees for this article. This paper was not supported by any external funding.

Conflicts of Interest

Author TL is an editor for the iMedicalApps website, dedicated towards providing news on the integration of mobile technology into medical care and the reviewing of medical apps for mobile devices. Neither TL or JW consult or receive reimbursement from app developers or creators.

References

1. 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]
2. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. J Med Internet Res 2012;14(5):e128 [FREE Full text] [doi: 10.2196/jmir.1994] [Medline: 23017375]
3. Sclafani J, Tirrell TF, Franko OI. Mobile tablet use among academic physicians and trainees. J Med Syst 2013 Feb;37(1):9903 [FREE Full text] [doi: 10.1007/s10916-012-9903-6] [Medline: 23321961]
4. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]
5. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]
6. 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]
7. 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;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
8. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
9. Morris R, Javed M, Bodger O, Hemington Gorse S, Williams D. A comparison of two smartphone applications and the validation of smartphone applications as tools for fluid calculation for burns resuscitation. Burns 2014 Aug;40(5):826-834. [doi: 10.1016/j.burns.2013.10.015] [Medline: 24246618]
10. Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. J Med Internet Res 2014;16(2):e32 [FREE Full text] [doi: 10.2196/jmir.3062] [Medline: 24491911]
11. Haffey F, Brady RR, Maxwell S. A comparison of the reliability of smartphone apps for opioid conversion. Drug Saf 2013 Feb;36(2):111-117. [doi: 10.1007/s40264-013-0015-0] [Medline: 23322549]
12. Wolf JA, Moreau JF, Akilov O, Patton T, English JC, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol 2013 Apr;149(4):422-426 [FREE Full text] [doi: 10.1001/jamadermatol.2013.2382] [Medline: 23325302]
13. 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]
14. Ferrero NA, Morrell DS, Burkhart CN. Skin scan: a demonstration of the need for FDA regulation of medical apps on iPhone. J Am Acad Dermatol 2013 Mar;68(3):515-516. [doi: 10.1016/j.jaad.2012.10.045] [Medline: 23394920]
15. Misra S, Lewis TL, Aungst TD. Medical application use and the need for further research and assessment for clinical practice: creation and integration of standards for best practice to alleviate poor application design. JAMA Dermatol 2013 Jun;149(6):661-662. [doi: 10.1001/jamadermatol.2013.606] [Medline: 23783150]
16. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
17. McCartney M. How do we know whether medical apps work? BMJ 2013;346:f1811. [Medline: 23516158]
18. Hamilton AD, Brady RR. Medical professional involvement in smartphone 'apps' in dermatology. Br J Dermatol 2012 Jul;167(1):220-221. [doi: 10.1111/j.1365-2133.2012.10844.x] [Medline: 22283748]

19. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012;10:144 [FREE Full text] [doi: 10.1186/1741-7015-10-144] [Medline: 23171675]
20. Rodrigues MA, Visvanathan A, Murchison JT, Brady RR. Radiology smartphone applications; current provision and cautions. Insights Imaging 2013 Oct;4(5):555-562 [FREE Full text] [doi: 10.1007/s13244-013-0274-4] [Medline: 23912880]
21. van Velsen L, Beaujean DJ, van Gemert-Pijnen JE. Why mobile health app overload drives us crazy, and how to restore the sanity. BMC Med Inform Decis Mak 2013;13:23 [FREE Full text] [doi: 10.1186/1472-6947-13-23] [Medline: 23399513]
22. US Food and Drug Administration. Mobile medical applications: guidance for industry and Food and Drug Administration staff. 2013. URL: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf [accessed 2014-09-08] [WebCite Cache ID 6SQq2TgYl]
23. FDA issues final guidance on mobile medical apps Internet. 2013 Sep 23. URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm369431.htm [accessed 2014-09-08] [WebCite Cache ID 6SQpzUtdf]
24. McCarthy M. FDA will not regulate most mobile medical apps. BMJ 2013;347:f5841. [Medline: 24072875]
25. Medicines and Healthcare Products Regulatory Agency (MHRA). 2010. Medical Device Technology Forum on the use of software as a medical device URL: http://www.mhra.gov.uk/Howweregulate/NewTechnologiesForums/DevicesNewTechnologyForum/Forums/CON084987 [accessed 2014-09-08] [WebCite Cache ID 6SQq3qtCy]
26. Therapeutic Goods Administration. Regulation of medical software and mobile medical apps. 2013 Sep 25. URL: http://www.tga.gov.au/industry/devices-software-mobile-apps.htm [accessed 2014-09-08] [WebCite Cache ID 6SQq76Z5e]
27. International Organization for Standardization (ISO). Medical Devices - Application of risk management to medical devices ISO 14971. URL: http://www.iso.org/iso/catalogue_detail?csnumber=38193 [accessed 2014-09-08] [WebCite Cache ID 6SQqjQ7pz]
28. Health & Social Care Information Centre. Clinical Risk Management: Telehealth/Mobile Health Solutions - Implementation Guidance. 2013. URL: http://www.isb.nhs.uk/documents/isb-0129/amd-39-2012/0129392012tele.pdf [accessed 2014-09-08] [WebCite Cache ID 6SQqIet5M]
29. Aungst TD, Clauson KA, Misra S, Lewis TL, Husain I. How to identify, assess and utilise mobile medical applications in clinical practice. Int J Clin Pract 2014 Feb;68(2):155-162. [doi: 10.1111/ijcp.12375] [Medline: 24460614]
30. Lewis TL. A systematic self-certification model for mobile medical apps. J Med Internet Res 2013;15(4):e89 [FREE Full text] [doi: 10.2196/jmir.2446] [Medline: 23615332]
31. Wyatt JC, Lewis TL, Curtis K. How to Assess and Improve the Quality and Safety of Apps for Medicine and Health Promotion? In: Med 2.0, 6th World Congr Soc Media, Mob Apps, Internet/Web 2.0 2013.: JMIR; 2013 Presented at: Medicine 2.0, 6th World Congress on Social Media, Mobile Apps, Internet/Web2.0; September 23-24, 2013; London p. 442-443.
32. Yetisen AK, Martinez-Hurtado JL, da Cruz Vasconcellos F, Simsekler MC, Akram MS, Lowe CR. The regulation of mobile medical applications. Lab Chip 2014 Mar 7;14(5):833-840. [doi: 10.1039/c3lc51235e] [Medline: 24425070]
33. d4. Regulation of Health Apps: A Practical Guide URL: http://www.d4.org.uk/research/regulation-of-health-apps-a-practical-guide-January-2012.pdf [accessed 2014-09-08] [WebCite Cache ID 6SQqgN5pR]

Abbreviations

AUF: app usage factor
FDA: Food and Drug Administration

Edited by G Eysenbach; submitted 25.11.13; peer-reviewed by L Ferris, D Rhon; comments to author 03.02.14; revised version received 11.02.14; accepted 11.04.14; published 15.09.14.

Please cite as:
Lewis TL, Wyatt JC
mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use
J Med Internet Res 2014;16(9):e210
URL: http://www.jmir.org/2014/9/e210/
doi:10.2196/jmir.3133
PMID:25223398

Viewpoint

Crowdsourcing Knowledge Discovery and Innovations in Medicine

Leo Anthony Celi[1,2*], MD, MPH, MS; Andrea Ippolito[3*], MS, MEng; Robert A Montgomery[4*], MD; Christopher Moses[5*], BS; David J Stone[6*], MD

[1]Institute for Medical Engineering and Science, Laboratory of Computational Physiology, Massachusetts Institute of Technology, Cambridge, MA, United States

[2]Beth Israel Deaconess Medical Center, Division of Pulmonary, Critical Care and Sleep Medicine, Harvard Medical School, Boston, MA, United States

[3]Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA, United States

[4]Beth Israel Deaconess Medical Center, Department of Medicine, Harvard Medical School, Boston, MA, United States

[5]Smart Scheduling, Inc., Cambridge, MA, United States

[6]UVA Center for Wireless Health, Departments of Anesthesiology and Neurosurgery, University of Virginia School of Medicine, Charlottesville, VA, United States

[*]all authors contributed equally

Corresponding Author:
Leo Anthony Celi, MD, MPH, MS
Institute for Medical Engineering and Science
Laboratory of Computational Physiology
Massachusetts Institute of Technology
77 Massachusetts Avenue
E25-505
Cambridge, MA, 02139
United States
Phone: 1 617 253 7937
Fax: 1 617 258 7859
Email: lceli@mit.edu

Abstract

Clinicians face difficult treatment decisions in contexts that are not well addressed by available evidence as formulated based on research. The digitization of medicine provides an opportunity for clinicians to collaborate with researchers and data scientists on solutions to previously ambiguous and seemingly insolvable questions. But these groups tend to work in isolated environments, and do not communicate or interact effectively. Clinicians are typically buried in the weeds and exigencies of daily practice such that they do not recognize or act on ways to improve knowledge discovery. Researchers may not be able to identify the gaps in clinical knowledge. For data scientists, the main challenge is discerning what is relevant in a domain that is both unfamiliar and complex. Each type of domain expert can contribute skills unavailable to the other groups. “Health hackathons” and “data marathons”, in which diverse participants work together, can leverage the current ready availability of digital data to discover new knowledge. Utilizing the complementary skills and expertise of these talented, but functionally divided groups, innovations are formulated at the systems level. As a result, the knowledge discovery process is simultaneously democratized and improved, real problems are solved, cross-disciplinary collaboration is supported, and innovations are enabled.

(J Med Internet Res 2014;16(9):e216) doi:10.2196/jmir.3761

KEYWORDS

knowledge discovery; crowdsourcing; innovation; hackathon

Addressing the Knowledge Gaps in Medicine

On October 30th, 1948, Austin Bradford Hill and his colleagues at England’s Medical Research Council published “Streptomycin treatment of pulmonary tuberculosis” [1]. Using the power of a coin flip (or in this case, a random draw of an envelope), Hill was able to remove selection bias, revealing the clearest possible picture of causality then available. With this simple addition, he established the basic framework of a randomized controlled trial (RCT), a new standard for guiding evidenced-based medicine. In the years since, RCTs have upended much of clinical practice, allowing the medical field to organize into a

system that creates and propagates new knowledge. Yet 65 years after the first RCT, only 10-20% of medical decisions are based on evidence [2]. And, as target populations subdivide along permutations of chronic morbid conditions and countless genetic polymorphisms, as diagnostic tools become more personalized, and as therapeutic options expand beyond the evaluation of individual drugs and devices to encompass the health care delivery network itself, it is increasingly apparent that RCTs cannot scale to match the exponential growth of medical complexity. While there are efforts underway to reduce the waste, cost, and difficulty of conducting research, clinicians and patients alike are currently left coping with a system of unacceptable ambiguity [3-6].

When faced with complex diagnostic or treatment uncertainties, patient and provider alike face several dilemmas. Combinations of diagnostic and therapeutic quandaries create the need for difficult decisions that reside in and are strongly affected by the contexts of individual patient factors and practice settings. One must determine when to probe more deeply and when to back off and observe without intervention. It is difficult, and perhaps a bit embarrassing to the medical profession, to attempt to further involve patients in the decision-making process based on current levels of uncertainty. It is not unlikely that during any clinical interaction one or more questions or problems will arise that cannot be fully addressed due to incomplete translation of research findings or clinical literature to the bedside, but more commonly, due to the incomplete state of medical knowledge.

Much of medical education consists of gaining skill and confidence not only in navigating but also in subsequently guiding others through this trail of ambiguity. One of the key objectives of the research enterprise is driving this ambiguity down so that practice is based more thoroughly on evidence. With the near ubiquitous implementation of digital documentation, we have the potential capability of answering more of these currently unresolved questions and transferring these answers into real-time workflows.

The Full-Time Clinician and Knowledge Discovery

Ideally, from the vast amount of electronic data we already have created and further generate every day, frontline providers should be better empowered to answer the tough questions that pertain to individual patients. Better information should allow clinicians to make better decisions with a more robust element of patient involvement. However, there are real but surmountable barriers to such an approach. The condensed version of the answer is that we need more data-savvy participants as well as more carefully engineered software applications at the core of the clinical data analytic process. Clinicians should not have to become data scientists, but an appropriate awareness and understanding of basic data issues is fast becoming an important element of clinical practice. This does not represent the stumbling block for the current and upcoming generations of medical students (who have grown up in digital environments) that it may represent for older, sometimes resistant clinicians.

It would be unrealistic to expect clinicians to conduct queries of clinical data to generate evidence/data-driven decisions for each patient. This is especially so given the current lack of a technological infrastructure that would allow frontline providers to pose such questions at the point of care in real or near-real time. This reveals a critical fragility in our current knowledge generating system: the divide between the roles of researcher, data systems engineer, and clinician. This separation is detrimental—for researchers, it can make it difficult to identify knowledge gaps in clinical decision making. For engineers, the problems include identifying what is important in a foreign and complex domain, working within non-interoperable, proprietary silos, and the difficulties involved in creating effective and user-friendly clinical information systems. And for practitioners, it diminishes a sense of connection to scientific investigation and disengages busy clinicians from constructive inquiry and participation in their own clinical data systems. Overall, these systematic flaws restrict the raw number of people, ideas, and innovations that are available to address and solve the myriad problems encountered in the day-to-day care of patients.

Democratizing Research I: Crowdsourcing and Open Data

While recognizing the challenge, we believe that it is important to find ways to democratize or "crowdsource" research. The term "crowdsourcing" was first introduced in 2005 by Jeff Howe and Mark Robinson, editors of Wired magazine, after conversations about how businesses were using the Internet to outsource work to individuals [7].

> *Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers.*

Crowdsourcing knowledge discovery in medicine can be vertically approached by lowering the barriers of participation to frontline providers and horizontally approached by extending an input role to non-traditional but interested contributors such as patients themselves. When applied to innovations in general, this process would permit people interacting with the medical system to develop exactly what they want, rather than relying on manufacturers to act as their (often very imperfect) agents. Moreover, individual users should not have to build everything from scratch or query a database on their own: they benefit from collaborating with those who have the skillset that they lack or building on solutions developed by and freely shared by others.

Embedding user-driven research and development into communities can create connections that accelerate and enhance the innovation process, increasing the speed and effectiveness of the dissemination of a solution or new knowledge [8]. This concept has been well documented in other industries. For instance, the Defense Advanced Research Projects Agency

originally developed ARPAnet to create a network of researchers and defense contractors to accelerate the exchange of software and data—this evolved into what we know as the Internet [9]. Linus Torvalds developed LINUX in 1991 as a free and open-source operating system that enabled anyone (albeit, with the requisite skill set) to contribute to its development [10].

Although medicine brings a unique set of challenges to systematic change and improvement, a more democratized approach is needed to support and optimize such processes. There have been numerous projects in the health field that have used crowdsourcing as methodology, including the three that we cite here. "FoodSwitch" is a mobile phone app that provides consumers with nutrition information obtained from users through a crowdsourcing function integrated within the app [11]. In another paper, Brown and colleagues described a method to distribute evaluation of scientific literature, a time-consuming endeavor that requires hours of coding and rating, across a large group through online crowdsourcing using Amazon's "Mechanical Turk" [12]. Finally, Good and company developed and evaluated an online game called "The Cure", which captured information from players regarding genes for use as predictors of breast cancer survival [13]. Their group demonstrated that crowdsourcing games can be developed as a means to address problems involving domain knowledge.

The crowdsourcing and open data movements present an opportunity to involve frontline providers and patients in accelerating innovation, including knowledge creation. Not unlike dispersed crowdsourcing networks, "hackathons" and data marathons provide opportunities for those at the front line of care, who are most familiar with the pain points within and the information gaps that plague day-to-day practice, to contribute to the much needed health care transformation. Knowledge discovery and innovations have been activities traditionally limited to doctors who have devoted their careers to research in academia or consulting in industry, or those who have given up clinical medicine. Funding to pursue academic research and opportunities to direct biotech research are seldom available to clinicians who spend most of their time in practice. Research is deemed exclusive to those with training in experimental methodologies and/or data analytics, while an additional degree in business is favored in biomedical entrepreneurship or consulting. As a result, there is always a level of disconnect between the foci of research and innovation, and solutions that will truly improve care delivery and health outcomes. In addition, practicing clinicians are almost always exhausted by the daily grind, frustrated by the inefficiencies of the health care system, and very seldom given the time, the opportunity, and the incentive to step back and address systems-level problems, including knowledge gaps in the practice of medicine.

Democratizing Research II: Hackathons and Data Marathons

Traditionally, hackathons are 24- to 48-hour events at the front end of the innovation process that provide an accessible forum to pitch complex, difficult problems and develop initial solutions and prototypes in a quick, iterative manner. Health care hackathons and data marathons provide opportunities for providers to collaborate with engineers or data scientists and quickly make an impact by offering the clinical perspective around problems or information gaps. Similarly, these events present a great forum for engineers and data scientists to gain access to real problems in medicine and to engage with stakeholders who have expert domain knowledge. By teaming up engineers and data scientists with clinicians, the hackathons and data marathons provide the non-clinicians a rare opportunity to transform health care. And while a truly novel discovery or a fully functioning solution is a rare outcome at the end of the event, we believe health care-focused hackathons and data marathons enable crowdsourcing of valuable and diverse points of view as well as creating new personal connections that will form the basis for productive longer-term collaborations.

The authors of this article have helped organize numerous hackathons and data marathons that have brought together engineers, data scientists, and clinicians (including nurses, pharmacists, and other allied health personnel) to address problems and questions identified during routine clinical practice, including the Critical Data Marathon held at the Massachusetts Institute of Technology (MIT) in January 2014 (see Multimedia Appendix 1). To date, the MIT Hacking Medicine has organized 17 events in the United States, India, Uganda, and Spain with a diverse set of partners including the Laboratory of Computational Physiology and Sana at MIT, the Consortium for Affordable Medical Technologies (CAMTech) at the Massachusetts General Hospital (MGH), and Brigham and Women's Hospital. These events have resulted in over 600 innovative ideas and over 250 teams that have developed prototypes, launched products, and/or published papers. In addition, a number of other organizations and initiatives are contributing to this hackathon movement in the health care community, including Health 2.0, Hacking Health, and the MedStar Institute for Innovation.

An example of the type of innovation that results from the collaboration fostered during hackathons is the Augmented Infant Resuscitator (AIR) [14]. Dr Data Santorino, a pediatrician in Uganda and researcher at Mbarara University of Science and Technology, presented the problem of newborn deaths from improper resuscitation techniques seen in low-income countries. At one of the hackathons held at MGH, he teamed up with an engineer from MIT, another clinician from MGH, and a business entrepreneur, and developed the prototype for AIR. The project has since gained considerable investment and the device is currently in field trials in Uganda.

For the MIT Critical Data Marathon held in January 2014, participants worked directly on a large, open-access clinical database called MIMIC, short for Multi-parameter Intelligent Monitoring in Intensive Care [15]. The database is the creation of a public-private partnership between the Beth Israel Deaconess Medical Center (BIDMC) in Boston, MIT, and Philips Healthcare. Committed to archiving all available data from the intensive care units at BIDMC—rather than a subset considered relevant at the time—researchers subsequently de-identify and publish the database for easy access at no cost. We believe that providing the data to as many people as possible is the best way to unlock the functionally cryptic information

in electronic health records for translation into valuable information. MIMIC has attracted both clinicians and data scientists who have partnered on many outcome studies that have included examinations of practice variability, the heterogeneity of treatment effects, cost analyses, and predictive modeling, among others. Clinicians who have not typically engaged in research but who possess a deep understanding of the information gaps as well as the elements involved in medical practice are now empowered to contribute to and become part of a data-driven learning system.

Achievable Benefits

We have previously commented on how open data and crowdsourcing may address but at the same time potentially augment the problem of unreliable and wasteful research [16]. The issue stems from the irreproducibility of what gets published and the inability of researchers to know what is not getting published. Our group organized a conference in conjunction with the data marathon held in January 2014 to address these concerns [16]. Thought leaders from academia, government, and industry across disciplines gathered and discussed the pitfalls and challenges of the data revolution sweeping health care. The consensus seemed to be that success will require a systematized and fully transparent data interrogation, where data and methods are freely shared among different groups of investigators addressing the same or similar questions. The added accuracy of the scientific findings is only one of the benefits of the systematization of the open data movement. Another will be the opportunity afforded to individuals of every educational level and area of expertise to contribute to science.

The Critical Data Marathon is one example of the realization of the goal of MIMIC: the democratization of medical research and crowdsourcing of knowledge discovery. We witnessed clinicians excitedly pairing with data scientists to work together in translating and parsing their questions into study designs and methodologies; nurses and doctors providing data scientists with essential but nuanced clinical contexts; and even architects and designers assisting in the visualization of findings. By engaging practicing clinicians as well as future ones, that is, medical students, we enable them to contribute to innovation at the systems level. Open questions remain on how to scale data sets and the infrastructure supporting the use of these data sets. But despite such challenges, the crowdsourcing movement is slowly transforming the medical culture into one where there is no divide between research and practice. These hackathons and data marathons provide a platform for frontline health care workers to create solutions to the problems in which they are immersed, democratize innovations and research in health care by engaging those who may not see themselves as academics or entrepreneurs, and harness the power of cross-disciplinary collaboration at a much larger scale.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Inaugural MIT Critical Data Marathon, January 3-5, 2014. Photos courtesy of Andrew Zimolzak.

[JPG File, 228KB - jmir_v16i9e216_app1.JPG]

References

1. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. Br Med J 1948;2(4582):769-782.
2. Committee on the Learning Health Care System in America, Institute of Medicine. In: Smith M, Saunders R, Stuckhardt L, McGinnis JM, editors. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC: The National Academies Press; Sep 06, 2012.
3. The Economist. 2013 Oct 19. Unreliable research: Trouble at the lab URL: http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble [accessed 2014-09-15] [WebCite Cache ID 6ScXiZ9J1]
4. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet 2014 Jan 11;383(9912):166-175. [doi: 10.1016/S0140-6736(13)62227-8] [Medline: 24411645]
5. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: increasing value, reducing waste. Lancet 2014 Jan 11;383(9912):101-104. [doi: 10.1016/S0140-6736(13)62329-6] [Medline: 24411643]
6. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet 2014 Jan 18;383(9913):257-266. [doi: 10.1016/S0140-6736(13)62296-5] [Medline: 24411650]
7. Howe J. Crowdsourcing: a definition. URL: http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html [accessed 2014-08-31] [WebCite Cache ID 6SEC6ae4b]
8. von Hippel E. Democratizing innovation: the evolving phenomenon of user innovation. JfB 2005 Mar;55(1):63-78 [FREE Full text] [doi: 10.1007/s11301-004-0002-8]
9. Ryan J. A History of the Internet and the Digital Future. London, UK: Reaktion Books Ltd; 2010.
10. DiBona C, Ockman S. Open Sources: Voice from the Open Source Revolution. Sebastopol, CA: O'Reilly Media; 1999.

11. Dunford E, Trevena H, Goodsell C, Ng KH, Webster J, Millis A, et al. FoodSwitch: a mobile phone app to enable consumers to make healthier food choices and crowdsourcing of national food composition data. JMIR Mhealth Uhealth 2014;2(3):e37 [FREE Full text] [doi: 10.2196/mhealth.3230] [Medline: 25147135]
12. Brown AW, Allison DB. Using crowdsourcing to evaluate published scientific literature: methods and example. PLoS One 2014 Jul;9(7):e100647 [FREE Full text] [doi: 10.1371/journal.pone.0100647] [Medline: 24988466]
13. Good BM, Loguercio S, Griffith OL, Nanis M, Wu C, Su AI. The Cure: design and evaluation of a crowdsourcing game for gene selection for breast cancer survival prediction. JMIR Serious Games 2014 Jul 29;2(2):e7. [doi: 10.2196/games.3350]
14. DePasse JW, Carroll R, Ippolito A, Yost A, Santorino D, Chu Z, et al. Less noise, more hacking: how to deploy principles from MIT's hacking medicine to accelerate health care. Int J Technol Assess Health Care 2014 Aug 6:1-5. [doi: 10.1017/S0266462314000324] [Medline: 25096225]
15. Celi LA, Mark RG, Stone DJ, Montgomery RA. "Big data" in the intensive care unit. Closing the data loop. Am J Respir Crit Care Med 2013 Jun 1;187(11):1157-1160. [doi: 10.1164/rccm.201212-2311ED] [Medline: 23725609]
16. Badawi O, Brennan T, Celi LA, Feng M, Ghassemi M, Ippolito A, et al. Making big data useful for health care: a summary of the inaugural MIT Critical Data Conference. JMIR Med Inform 2014 Aug 22;2(2):e22. [doi: 10.2196/medinform.3447]

Abbreviations

AIR: Augmented Infant Resuscitator
BIDMC: Beth Israel Deaconess Medical Center
MGH: Massachusetts General Hospital
MIMIC: Multi-parameter Intelligent Monitoring in Intensive Care
MIT: Massachusetts Institute of Technology
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 06.08.14; peer-reviewed by A Sarafi Nejad, A Holzinger, SH Tee; comments to author 27.08.14; revised version received 31.08.14; accepted 31.08.14; published 19.09.14.

Please cite as:
Celi LA, Ippolito A, Montgomery RA, Moses C, Stone DJ
Crowdsourcing Knowledge Discovery and Innovations in Medicine
J Med Internet Res 2014;16(9):e216
URL: http://www.jmir.org/2014/9/e216/
doi:10.2196/jmir.3761
PMID:25239002

Original Paper

Prevention of Generalized Anxiety Disorder Using a Web Intervention, iChill: Randomized Controlled Trial

Helen Christensen[1], BA (Hons), MClinPsych, PhD; Philip Batterham[2], PhD; Andrew Mackinnon[3], PhD; Kathleen M Griffiths[2], PhD; Kanupriya Kalia Hehir[4], BPsych; Justin Kenardy[5], PhD; John Gosling[2], BPsych; Kylie Bennett[2], BSc

[1]Black Dog Institute, University of New South Wales, Sydney, Australia

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

[3]Orygen Youth Mental Health Research Centre, University of Melbourne, Melbourne, Australia

[4]Australian Primary Health Care Research Institute, Australian National University, Canberra, Australia

[5]Centre of National Research on Disability and Rehabilitation Medicine, University of Queensland Schools of Medicine and Psychology, Brisbane, Australia

Corresponding Author:
Helen Christensen, BA (Hons), MClinPsych, PhD
Black Dog Institute
University of New South Wales
Hospital Road, Prince of Wales Hospital
Randwick
Sydney, 2031
Australia
Phone: 61 293829288
Fax: 61 293828207
Email: h.christensen@blackdog.org.au

Abstract

Background: Generalized Anxiety Disorder (GAD) is a high prevalence, chronic disorder. Web-based interventions are acceptable, engaging, and can be delivered at scale. Few randomized controlled trials evaluate the effectiveness of prevention programs for anxiety, or the factors that improve effectiveness and engagement.

Objective: The intent of the study was to evaluate the effectiveness of a Web-based program in preventing GAD symptoms in young adults, and to determine the role of telephone and email reminders.

Methods: A 5-arm randomized controlled trial with 558 Internet users in the community, recruited via the Australian Electoral Roll, was conducted with 6- and 12-month follow-up. Five interventions were offered over a 10-week period. Group 1 (Active website) received a combined intervention of psycho-education, Internet-delivered Cognitive Behavioral Therapy (ICBT) for anxiety, physical activity promotion, and relaxation. Group 2 (Active website with telephone) received the identical Web program plus weekly telephone reminder calls. Group 3 (Active website with email) received the identical Web program plus weekly email reminders. Group 4 (Control) received a placebo website. Group 5 (Control with telephone) received the placebo website plus telephone calls. Main outcome measures were severity of anxiety symptoms as measured by the GAD 7-item scale (GAD-7) (at post-test, 6, and 12 months). Secondary measures were GAD caseness, measured by the Mini International Neuropsychiatric Interview (MINI) at 6 months, Centre for Epidemiologic Studies-Depression scale (CES-D), Anxiety Sensitivity Index (ASI), Penn State Worry Questionnaire (PSWQ), and Days out of Role.

Results: GAD-7 symptoms reduced over post-test, 6-month, and 12-month follow-up. There were no significant differences between Group 4 (Control) and Groups 1 (Active website), 2 (Active website with telephone), 3 (Active website with email), or 5 (Control with telephone) at any follow-up. A total of 16 cases of GAD were identified at 6 months, comprising 6.7% (11/165) from the Active groups (1, 2, 3) and 4.5% (5/110) from the Control groups (4, 5), a difference that was not significant. CES-D, ASI, and PSWQ scores were significantly lower for the active website with email reminders at post-test, relative to the control website condition.

Conclusions: Indicated prevention of GAD was not effective in reducing anxiety levels, measured by GAD-7. There were significant secondary effects for anxiety sensitivity, worry, and depression. Challenges for indicated prevention trials are discussed.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 76298775; http://www.controlled-trials.com/ISRCTN76298775 (Archived by WebCite at http://www.webcitation.org/6S9aB5MAq).

(J Med Internet Res 2014;16(9):e199) doi:10.2196/jmir.3507

KEYWORDS

anxiety disorders; prevention; early intervention; Internet; online systems; cognitive behavioral therapy

Introduction

Approximately 5% of the general population experiences General Anxiety Disorder (GAD) at least once in their lifetime [1], with population surveys indicating a lifetime prevalence rate of between 4.3-5.9% and a 12-month prevalence rate of between 1.2-1.9% [2,3]. The cost of GAD to the community is high as a result of its chronic course [4]. GAD frequently presents early in the lifespan and affects the individual throughout adulthood, with an estimated lag time to treatment of between 9 and 23 years [5]. If the prevalence of GAD is to be lowered, prevention, particularly focusing on the early adult and adolescence years when the illness emerges [6], will reduce the prevalence of mental disorder by up to 23% [7-9].

There is some evidence that GAD can be prevented. However, the conduct of research trials has not been optimal either because the researchers have been unable to exclude those with a diagnosis at the onset of the intervention or because the trials are too small or too short to investigate the number of incident cases following the intervention [10]. We reviewed the research literature, but found that only four trials excluded a diagnosis of GAD in adults at baseline. Two of the studies found preventative effects, but both trials had limitations [11,12]. Van't Veer Tazelaar and colleagues [12] reported that depression and anxiety caseness could be halved in elderly people who were provided with a stepped care intervention of problem solving and cognitive behavioral therapy (CBT) bibliotherapy. However, the investigators did not report data separately for GAD, so the effect of the intervention on GAD compared to depression diagnosis could not be determined. Pitceathly and colleagues [13] reported a protective effect of a brief coping intervention on GAD in cancer patients. The effect was not detectable in the full sample, but was evident for those identified at the start of the trial with high risk of anxiety or depression. No preventative effects were found in the other two studies [14,15]. In the first, a stepped care intervention in elderly people living in residential care did not result in reduced incidence of combined anxiety or depression. In the second, carers of patients with Alzheimer's disease did not show lower levels of anxiety or depression as a result of an intervention involving a Family Meetings intervention. In effect, no genuine prevention trials have been conducted with younger adults and none with adults without a cancer diagnosis.

A key challenge to delivering prevention interventions is the low level of engagement by those at risk; if symptoms are not disabling, motivation may be low and seeking help from doctors seen as inappropriate. Web-based interventions provide a potentially very useful delivery medium because they are accessible, acceptable, globally disseminable, and have been found to be effective in delivering CBT in clinical settings for both depression and anxiety [16-18]. Engagement may be enhanced by "push" factors (ie, factors that encourage involvement or engagement, such as reminders or coaching) [19]. However, inconsistent findings are reported [20-23]. Because prevention programs are delivered to large numbers at a population level, the costs associated with different push factors are critical to the feasibility of prevention efforts. Hence, there is a need to know the extent to which email reminders and telephone communication with the research team will improve adherence and effectiveness. The present study aimed to evaluate the effectiveness of a Web-based multimedia CBT intervention in individuals aged 18 to 30 years with symptoms of anxiety, who did not meet diagnostic criteria at baseline. The intervention was a website that provided psychoeducation, CBT, physical activity promotion, and relaxation training, with the majority of the sessions focusing on CBT. Each component of the intervention was found to be effective in GAD treatment [24-27]. Our rationale to include all four components was based on the view that combining a range of evidence-based interventions provides potential for maximal impact, as well as the opportunity for participants' preferences (for example, see [28]). We know very little about engagement in prevention (compared to treatment) programs, and offering a range of interventions, all of which have evidence-based support from treatment settings, would potentially optimize effectiveness and uptake. We focused on young adults, as GAD develops during adolescence and early adulthood and any improvement would be likely to provide benefits over years. The intervention's "e-couch" website is open following registration and the CBT "worry" program can be experienced.

Five interventions were offered over a 10-week period. Group 1 (Active website) received the combined intervention as described above. Group 2 (Active website with telephone) received the same Web program plus weekly telephone reminder calls. Group 3 (Active website with email) received the Web program plus automated weekly email reminders. Group 4 (Control) received a placebo website, matched in length to the active website. Group 5 (Control with telephone) received the placebo website plus telephone reminder calls. This design allowed us to compare the effectiveness of the active interventions to the control condition (1, 2, 3 vs 4) and also to determine the independent effect of phone contact in the control conditions (5 vs 4). Email reminders are less expensive than person-made telephone reminders, an important consideration for a prevention trial. To our knowledge, no similar prevention trials have been conducted with each of these inclusions: the use of an online intervention for GAD, targeting of young adults in the community, and excluding existing GAD diagnosis.

Methods

Study Design

A randomized controlled trial with 5 arms, called the "iChill" trial, with post-test, 6- and 12-month follow-up, was conducted. The study was approved by the ANU Human Ethics Committee (Protocol 2008/548).

Setting, Participants, and Eligibility Criteria

The study protocol [29] describes trial details. A screening questionnaire was emailed to 120,000 randomly chosen Australians aged 18-30 years registered on the Australian Electoral Roll. Individuals meeting inclusion criteria were invited to a Web portal where they provided consent and undertook screening and baseline surveys. They were then interviewed via telephone to determine current GAD diagnosis using the MINI International Neuropsychiatric Interview (MINI) [30] and randomized to the trial. Inclusion criteria were willingness to consent, an active email and phone number, English language proficiency, Internet access, and a score above 5 on the GAD-7 [31]. In order to specify the target population for whom the intervention might be effective, participants were excluded if they were currently undergoing CBT or seeing a psychologist or a psychiatrist, had a current or previous diagnosis of bipolar disorder, schizophrenia, or psychosis, were at risk of self-harm or suicide based on the MINI depression module, or had a current diagnosis of panic disorder, social phobia, or post-traumatic stress disorder (PTSD) on the MINI. Participants were not excluded if taking antidepressants or benzodiazepines. A total of 40 (7.2%, 40/558) were taking antidepressants and 8 (1.4%, 8/558) were taking benzodiazepines at baseline; 510 were not taking either antidepressants or benzodiazepines.

Randomization

The algorithm for randomization consisted of a stratified block design with eight strata (2 x 2 x 2) corresponding to gender, past GAD diagnosis, and severity of GAD symptoms with a block size of 10. "Past GAD diagnosis" was obtained from the MINI—a proportion of participants (24.2%, 135/558) met lifetime criteria for GAD on the MINI but did not meet MINI criteria for current GAD. To minimize imbalance between such participants, stratification was conducted on the basis of meeting lifetime GAD criteria (along with gender and current GAD symptoms). Allocation was administered via software architecture, and participants were informed of condition allocation after baseline interview.

Interventions

The Active website intervention was a 10-week structured version of the Anxiety and Worry modules of the "e-couch" program (ecouch.anu.edu.au), which consisted of an integrated program of psychoeducation (weeks 1-2), CBT (weeks 3-7), relaxation (weeks 8-9), and physical activity promotion (week 10). The psychoeducation section (Modules 1 and 2) provides information on worry, stress, fear, and anxiety; a description of anxious thinking; differentiation of GAD from other anxiety disorders; risk factors for GAD; comorbidity; and consequences of anxiety and available treatments. This section is based on interventions for mental health literacy that have succeeded in reducing symptoms of depression and anxiety, and improving mental health attitudes [32]. The CBT toolkits (Modules 3-7) addressed typical anxious thoughts and included sections on dealing with the purpose and meaning of worry, the act of worrying, and the content of worry. The information is derived from materials that have been found to reduce anxious cognitions in at-risk people [33,34]. Progressive muscle relaxation (PMR) (Module 8) instructs participants on how to progressively tense and relax different muscle groups to induce relaxation and help to identify tension early. PMR has been trialed in a previous website program for depression in adults [35] and adolescents [36]. The mindfulness meditation module (Module 9) helps participants become aware of their breathing and body, acknowledging thoughts and external distractions but remaining focused on the present. The final module, physical activity (Module 10), tailors advice about physical activity based on the stages of change theory [37]. The control website was an adapted version of the HealthWatch control condition developed for the Australian National University WellBeing study [38]. This website provided information about general health (nutrition, heart health, etc), and invited responses to questions about anxiety. Scripted telephone reminder calls in the Active plus telephone condition were made on a weekly basis to check on participants' progress and to remind them to complete the module and/ or to keep completing the program. The phone reminders were intended to serve purely as reminders, had no therapeutic input, and were made by casual phone interviewers. Telephone calls were scripted and were based on the email scripts. Any technical issues were referred to the Trial Manager. Phone calls generally lasted between 30 seconds to 2 minutes. Phone calls were made regardless of whether the participant completed the program or not. Participants in the Active plus email and Control plus email conditions were sent a weekly reminder email. These were similar in content to the phone calls. There was no therapeutic input.

Outcome Measures

The primary outcome was the Generalized Anxiety Disorder 7-item scale (GAD-7) [31]. Secondary outcomes were GAD caseness based on MINI; worry, measured by the Penn State Worry Questionnaire (PSWQ) [39]; anxiety sensitivity, as measured by the Anxiety Sensitivity Index (ASI) [40]; depression symptoms, measured by the Centre for Epidemiologic Studies-Depression Scale (CES-D) [41]; and disability measured by Days out of Role from the US National Comorbidity Study [42]. GAD caseness was measured at 6 months. GAD caseness at 36 months will be determined by proxy using GAD-7 cut-off scores. Other measures not analyzed in this paper focused on comorbidities, such as harmful/hazardous alcohol use as measured by the Alcohol Use Disorders Identification Test (AUDIT) [43], or duplicate measures of depression caseness estimated by the Patient Health Questionnaire 9-item (PHQ-9) [44], and other behaviors such as help seeking and perceived need for treatment. Outcomes were assessed at baseline, 6, and 12 months, with the exception of MINI caseness, which was assessed at 6 months.

Sample Size

We aimed, conservatively, to find an effect of 0.3 between each Active website group and the Control, based on effect sizes of 0.6 found for previous treatment and indicated prevention trials (0.6) [33,34]. This assumes a pre-post correlation of .7 between scores. With 600 participants, we would have 80% power to detect effects, allowing for 15% attrition.

Statistical Analysis

Primary analyses were undertaken on an intent-to-treat basis (ITT). Mixed model repeated measures (MMRM) [45] were used to include all available data, including that from participants who subsequently withdrew from the trial. This approaches yields unbiased estimates of intervention effects under the assumption that data are missing at random (MAR). Unlike conventional approaches to analysis, MAR allows observations to be missing conditional on observed variables appearing in the analytic model [46]. Non-linear mixed models were used to analyze caseness.

Results

A total of 558 people were randomized to a trial condition of whom 360 (64.5%) completed post-test, 303 (54.3%) completed 6-month follow-up, and 264 (47.3%) completed the 12-month follow-up. Figure 1 shows the flow of participants. Sample characteristics are presented in Table 1. Although mean GAD-7 scores were above the cut-point of 5 at screening for all participants (mean 8.3, SD 3.3), anxiety symptoms had decreased by the time the baseline was completed (mean 6.7, SD 3.8).

There were no differences on any of the baseline measures with the exception of “preference for active condition” and “employment status”. Across all conditions, the most preferred program was the Active website with email reminders (37.8%, 211/558), followed by the Control (31.4%, 175/558), with few participants stating a preference for telephone reminders (5.6%, 31/558 and 11.6%, 65/558, for Active and Control conditions, respectively). Lower preference for the Active website was found among those in the Control and Active with email conditions. Higher rates of full-time work were found among those receiving the Active website with telephone reminders and lower rates among those receiving the Control website with telephone reminders.

Adherence to the intervention differed significantly according to condition. Participants in the reminder conditions completed the majority of the 10 modules (Active/email: 5.5 modules, Active/telephone: 7.3, Control/telephone: 8.3) while those who did not receive reminders completed a little over one-third (3.7 modules for both active and control; $F_{4,\ 477.5}$=38.1, P<.001).

GAD-7 symptoms reduced at post-test and 6-month follow-up, but returned to baseline levels by 12 months. There were no significant differences between Group 4 (Control) and Groups 1 (Active website), 2 (Active website with telephone), 3 (Active website with email), or 5 (Control with telephone) at any follow-up. Outcomes were unchanged after adjusting for employment status and preferences for condition. Likewise, accounting for adherence by adjusting for module completion (ie, testing the efficacy of the intervention) did not change these outcomes. Figure 2 shows estimated marginal means of GAD-7 scores from the mixed model shown in Table 2.

Data on secondary outcomes analyses are displayed for continuous variables in Table 3. Based on the MINI assessment at 6 months, 16 cases of GAD were identified, comprising 6.7% from the Active groups (1, 2, and 3) and 4.5% from the Control groups (4, 5). There was no significant difference in the number of cases across these collapsed groups.

The mixed effect model repeated measure (MMRM) analyses for secondary outcomes were as follows. As for the primary outcome, there were no significant overall interactions between condition and time for CES-D, PSWQ, or Days out of Role. However, there was a significant interaction between condition and time for the ASI. Furthermore, there were significant effects of specific conditions at specific time points. CES-D, ASI, and PSWQ scores were significantly lower for the active website with email reminders at post-test, relative to the control website condition ($t_{389.2}$= −2.5, P=.015; $t_{368.7}$=−3.4, P<.001; $t_{371.9}$=−2.4, P=.017 respectively). The decrease in ASI scores for the active/email condition remained significant at 6 months ($t_{343.1}$=−2.3, P=.021). In addition, Days out of Role due to anxiety was significantly decreased at 12 months (but not at post-test or 6 months) for the active/email condition ($t_{398.3}$=−2.4, P=.016). There was also reduced worry in the active website with phone reminders at post-test and 6 months, relative to the control condition ($t_{368.6}$=−2.0, P=.047; $t_{340.2}$= −2.1, P=.035 respectively).

Table 1. Characteristics of the sample at baseline by trial arm.

	Active website (n=111)	Active website with email (n=113)	Active website with phone (n=110)	Control website (n=111)	Control website with phone (n=113)	F[a] or $\chi 2$	P
	mean (SD) or n (%)						
Characteristic							
Age	25.7 (3.2)	25.4 (3.3)	25.5 (3.1)	26.0 (3.0)	25.6 (3.5)	0.623	.646
GAD-7[d] score	6.8 (3.9)	6.2 (3.9)	6.8 (3.6)	6.9 (3.8)	6.6 (3.7)	0.604	.660
CES-D[e] score	16.7 (10.1)	17.7 (10.9)	16.7 (9.8)	18.8 (10.6)	17.3 (8.8)	0.802	.524
Anxiety sensitivity	19.3 (11.0)	19.9 (11.5)	17.5 (10.6)	18.7 (9.9)	18.8 (9.9)	0.753	.556
PSWQ[f]	40.5 (12.2)	37.9 (12.5)	39.5 (11.6)	40.3 (12.0)	39.2 (10.8)	0.835	.503
AUDIT[g] score	7.2 (5.1)	7.2 (5.7)	6.7 (4.8)	6.9 (4.7)	7.0 (5.3)	0.189	.944
DOR[h] due to anxiety	0.6 (1.5)	0.9 (2.3)	0.6 (1.6)	0.6 (1.4)	0.6 (1.2)	0.859	.488
Self-rated health[i]	2.4 (0.9)	2.5 (0.8)	2.4 (0.8)	2.5 (0.9)	2.5 (0.7)	0.879	.476
Childhood adversity[j]	1.5 (1.4)	1.9 (1.5)	1.8 (1.6)	1.7 (1.5)	1.7 (1.5)	0.919	.452
Traumatic events[k]	1.4 (1.5)	1.6 (1.6)	1.5 (1.5)	1.4 (1.5)	1.6 (1.6)	0.498	.737
Positive beliefs Internet therapy[l]	5.6 (1.2)	5.4 (1.3)	5.4 (1.3)	5.5 (1.3)	5.8 (1.3)	2.064	.084
Female gender	92 (82.9%)	90 (79.6%)	88 (80.0%)	89 (80.2%)	91 (80.5%)	0.474[b]	.976
Completed university degree	37 (33.3%)	32 (28.6%)	28 (25.9%)	36 (32.4%)	30 (26.8%)	2.352[b]	.671
Prefer active condition[m]	67 (60.4%)	51 (45.1%)	70 (63.6%)	60 (54.1%)	70 (61.9%)	10.501[b]	.033[c]
Employment status						16.297[b]	.038[c]
Full-time employment	69 (62.2%)	68 (61.3%)	81 (75.7%)	65 (59.1%)	58 (51.8%)		
Part-time employment	30 (27.0%)	27 (24.3%)	16 (15.0%)	34 (30.9%)	36 (32.1%)		
Not in labor force	12 (10.8%)	16 (14.4%)	10 (9.3%)	11 (10.0%)	18 (16.1%)		

[a] F tests are from one-way analysis of variance (ANOVA) for continuous variables

[b] $\chi 2$ tests for categorical variables

[c] $P<.05$

[d] GAD: Generalized Anxiety Disorder

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

[f] PSWQ: Penn State Worry Questionnaire

[g] AUDIT: Alcohol Use Disorders Identification Test

[h] DOR: Days out of Role

[i] Self-rated health assessed on a 5-point scale from 1 (excellent) to 5 (poor)

[j] Childhood adversity based on aggregate of 6 items assessing paternal/maternal mental health problems and substance use problems, high familial conflict, and parental separation/divorce

[k] Traumatic life events based on count from list of 14 traumatic events

[l] Positive beliefs in Internet therapy based on 2 items regarding confidence in learning skills about anxiety and better understanding anxiety using the Internet

[m] Preference for active condition based on a single item asking which intervention would be preferred

Table 2. Repeated measures mixed model of GAD-7 scores at post-test, 6, and 12 months.

Parameter		Estimate	Standard error	df	*t* / *F*	*P*
Intercept		6.946	0.359	553.0	19.3	<.001
Condition					1.0	.410
	Active website	−0.117	0.508	553.0	−0.2	.818
	Active website with email reminders	−0.716	0.506	553.0	−1.4	.158
	Active website with phone reminders	−0.173	0.510	553.0	−0.3	.734
	Control website with phone reminders	−0.335	0.506	553.0	−0.7	.508
	Control website	0.000	0.000			
Time					29.5	<.001
	Baseline	0.000	0.000			
	Post-test	−1.066	0.494	389.2	−2.2	.032
	6-month follow-up	−2.337	0.512	372.8	−4.6	<.001
	12-month follow-up	−1.178	0.612	315.8	−1.9	.055
Condition × time interaction					0.8	.622
	Active vs control at baseline	0.000	0.000			
	Active vs control at post-test	0.331	0.698	390.6	0.5	.636
	Active vs control at 6 months	0.179	0.723	374.2	0.2	.805
	Active vs control at 12 months	−0.274	0.853	313.9	−0.3	.748
	Active email vs control at baseline	0.000	0.000			
	Active email vs control at post-test	−0.517	0.718	393.9	−0.7	.472
	Active email vs control at 6 months	0.370	0.730	372.3	0.5	.612
	Active email vs control at 12 months	−0.157	0.896	319.4	−0.2	.861
	Active phone vs control at baseline	0.000	0.000			
	Active phone vs control at post-test	−0.689	0.682	384.5	−1.0	.314
	Active phone vs control at 6 months	0.736	0.709	369.4	1.0	.300
	Active phone vs control at 12 months	−1.082	0.851	315.4	−1.3	.205
	Control phone vs control at baseline	0.000	0.000			
	Control phone vs control at post-test	−0.137	0.655	379.2	−0.2	.835
	Control phone vs control at 6 months	0.394	0.683	364.4	0.6	.565
	Control phone vs control at 12 months	−0.150	0.808	310.2	−0.2	.853

Table 3. Primary and secondary outcome data.

Outcome		Active	Active / phone	Active / email	Control	Control / phone			
		mean (SD)/n	mean (SD)/n	mean (SD)/n	mean (SD)/n	mean (SD)/n	*F* [a]	df	*P*
Sample size									
	Baseline	111	110	113	111	113			
	Post-test	66	75	58	66	93			
	6-month	55	62	54	55	77			
	12-month	53	52	40	48	70			
GAD-7 [c]							0.8	12, 323.4	.622
	Baseline	6.8 (3.9)	6.8 (3.6)	6.2 (3.9)	7.0 (3.8)	6.6 (3.7)			
	Post-test	6.1 (4.7)	4.7 (3.6)	4.6 (2.9)	6.1 (4.1)	5.3 (4.2)			
	6-month	4.3 (3.0)	5.0 (4.0)	4.3 (3.7)	4.6 (3.6)	4.7 (3.1)			
	12-month	5.1 (4.6)	4.0 (3.4)	5.1 (4.1)	5.9 (4.5)	5.1 (3.6)			
PSWQ [d]							1.0	12, 345.4	.419
	Baseline	40.5 (12.2)	39.5 (11.6)	37.9 (12.5)	40.3 (12.0)	39.2 (10.8)			
	Post-test	39.0 (13.2)	37.4 (10.6)	33.8 (11.5)	41.0 (12.3)	38.4 (12.8)			
	6-month	33.9 (13.2)	38.2 (11.3)	35.9 (11.5)	38.9 (13.4)	37.9 (13.5)			
	12-month	34.1 (14.0)	33.2 (11.2)	34.4 (13.1)	38.3 (13.9)	37.2 (12.1)			
ASI [e]							1.7	12, 345.9	.057
	Baseline	19.3 (11.0)	17.5 (10.6)	19.9 (11.5)	18.7 (9.9)	18.8 (9.9)			
	Post-test	18.9 (11.8)	14.6 (10.6)	17.1 (11.1)	18.5 (12.2)	16.7 (11.0)			
	6-month	16.0 (12.3)	15.0 (11.5)	15.4 (9.5)	17.4 (10.2)	15.1 (9.8)			
	12-month	19.5 (12.4)	14.7 (10.0)	19.4 (11.6)	20.3 (11.1)	21.1 (10.6)			
CES-D [f]							1.9	12, 328.4	.036 [b]
	Baseline	16.7 (10.1)	16.7 (9.8)	17.7 (10.9)	18.8 (10.6)	17.3 (8.8)			
	Post-test	14.0 (10.8)	12.4 (8.6)	10.9 (8.4)	17.5 (11.3)	13.8 (9.6)			
	6-month	10.7 (7.7)	12.6 (10.5)	11.8 (8.9)	14.4 (11.3)	13.4 (8.1)			
	12-month	12.0 (9.2)	9.7 (5.9)	12.3 (11.3)	15.3 (9.3)	12.6 (8.7)			
DOR [g]							1.1	12, 369.9	.324
	Baseline	0.6 (1.5)	0.6 (1.6)	0.9 (2.3)	0.6 (1.4)	0.6 (1.2)			
	Post-test	0.5 (1.3)	0.2 (0.5)	0.4 (0.9)	0.5 (1.4)	0.2 (0.8)			
	6-month	0.3 (1.4)	0.4 (1.1)	0.1 (0.6)	0.5 (1.8)	0.5 (2.5)			
	12-month	0.3 (0.8)	0.2 (0.6)	0.1 (0.5)	0.7 (1.8)	0.4 (1.9)			

[a]statistics are omnibus *F* tests from mixed models repeated measures for each outcome, based on time × condition interaction terms

[b]*P*<.05

[c]GAD: Generalized Anxiety Disorder

[d]PSWQ: Penn State Worry Questionnaire

[e]ASI: Anxiety Sensitivity Index

[f]CES-D: Center for Epidemiologic Studies Depression scale

[g]DOR: days out of role due to anxiety

Figure 1. CONSORT diagram showing flow of participants through the study.

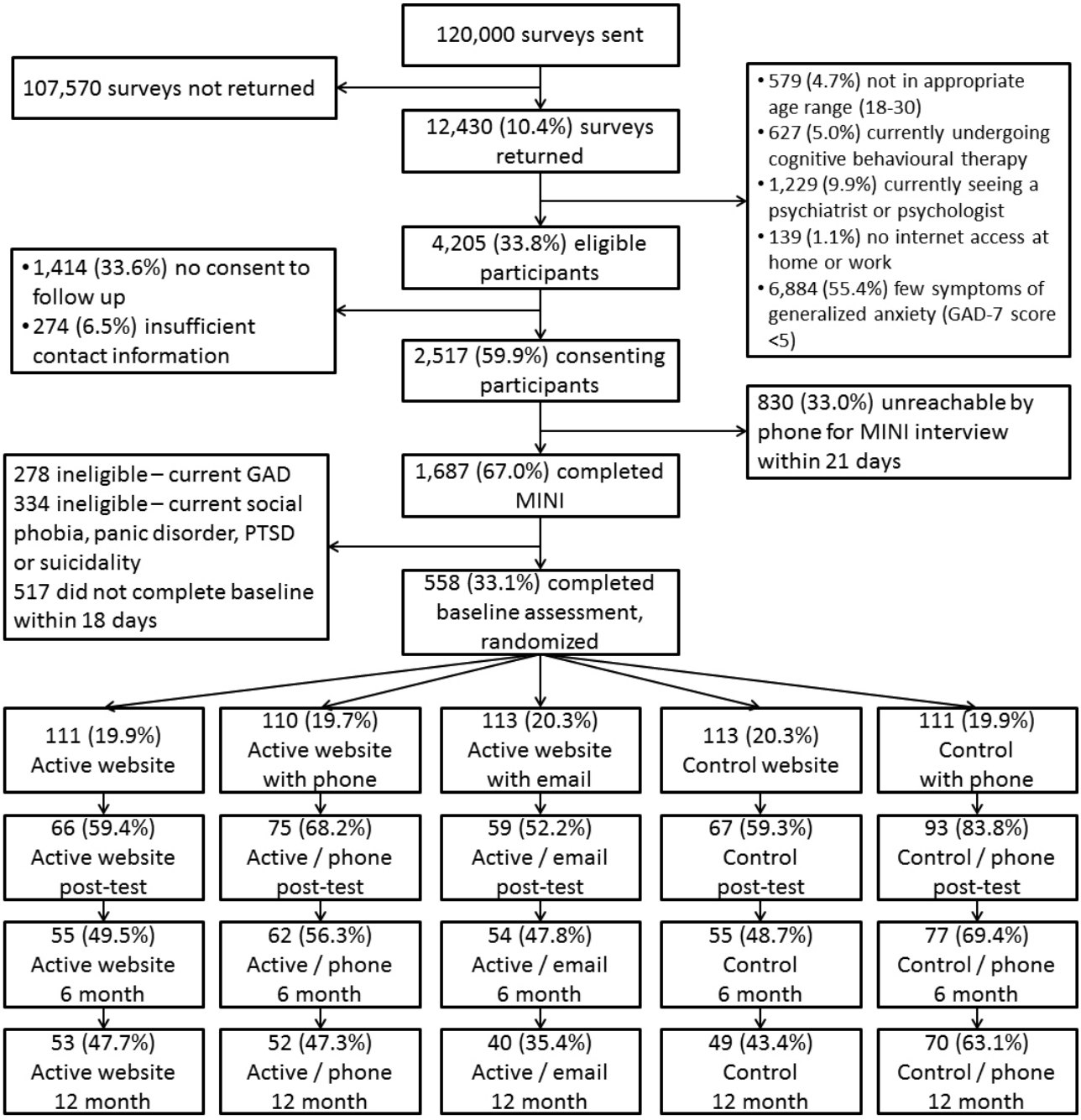
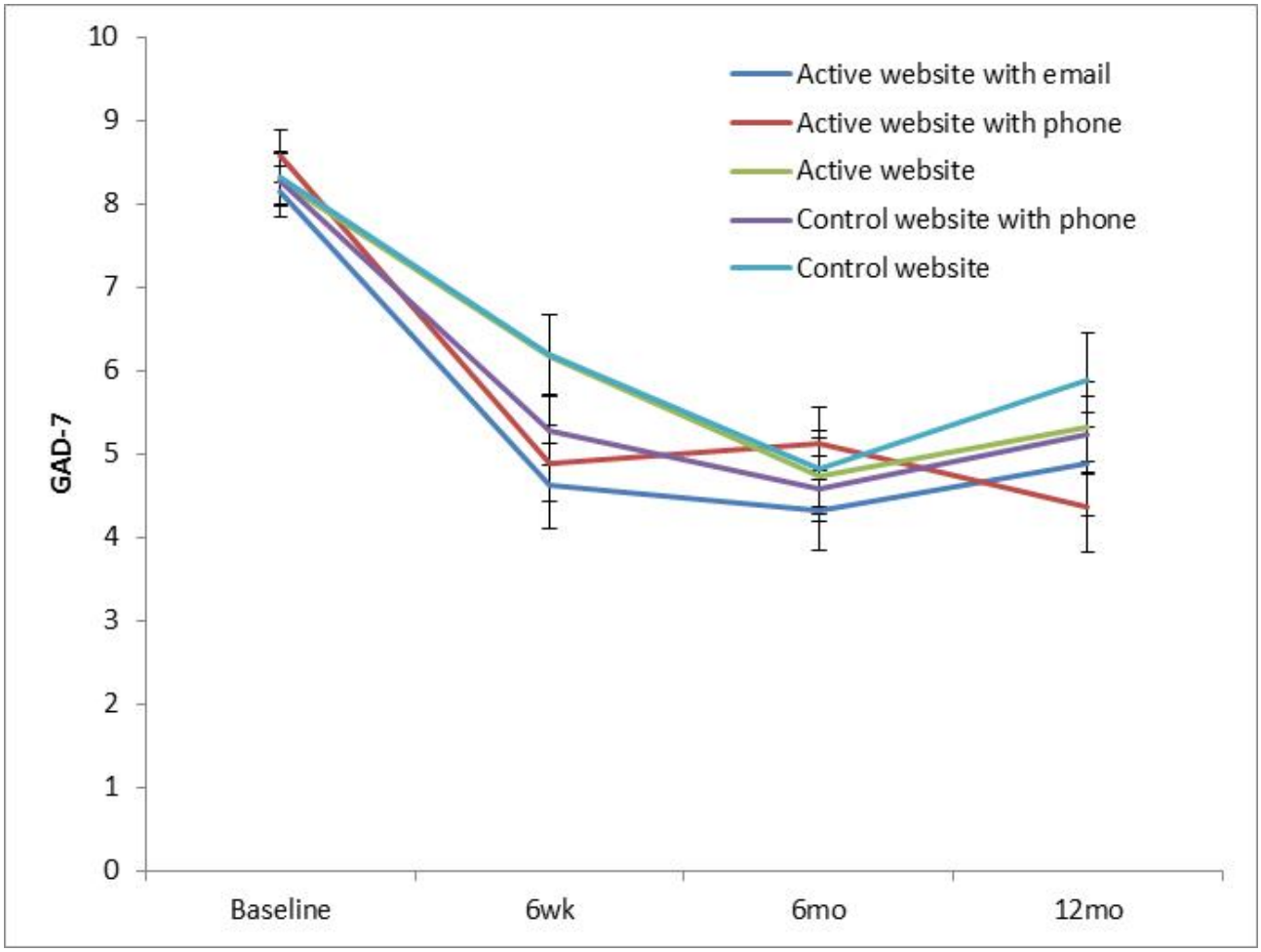


Figure 2. Outcome data on GAD-7 at post-test, 6-month and 12-month follow-ups (error bars represent standard error).



Discussion

Principal Findings

We found no evidence that a range of brief active interventions were associated with improved anxiety outcomes at post-test or at 6 or 12 months as measured by GAD-7 and by caseness at 6 months as measured by the MINI. We found that there were effects on secondary outcomes, most strongly found for anxiety sensitivity, where there was a significant interaction between condition and time for the ASI. In addition, ASI scores were significantly lower for the active website with email reminders at post-test, and at 6 months. There were also small effects at various follow-up intervals for the other measures mostly associated with the active website with email reminders. For instance, CES-D and PSWQ scores were significantly lower for the active website with email reminders at post-test, relative to the control website condition. In addition, Days out of Role due to anxiety was significantly decreased at 12 months for the active/email condition.

At best, these findings suggest that an active website with email may have small effects on a number of secondary outcomes. It is also possible that the GAD-7 may be a poorer measure of anxiety change than the ASI, and that genuine prevention effects operate, but were not discoverable because of our choice of outcome measure. The GAD-7 may be sensitive to a range of anxiety disorders [47] and may not have been the most robust outcome measure.

Limitations

In addition, a number of limitations to the present study need to be considered. There was differential dropout in a number of conditions relative to the control. Nevertheless, the completer analyses (not reported in the paper but undertaken) produced comparable effects to the main ITT analysis. The condition with the strongest secondary outcome effects was also the condition that was associated with the highest preference rating, suggesting that preference for it might have influenced the findings. In addition, the intervention itself may not have been optimal. We argued that combining a range of evidence-based interventions would provide maximal impact as well as provide opportunity for participants to make choices within the program content. It might be suggested that this diluted the effects of the individual components. However, we disagree, given that (1) the combined intervention is highly efficacious (effect sizes greater than 1.0) for patients with a diagnosis of GAD [48], (2) the multimedia intervention in the current paper was associated with effects on a range of secondary measures, and (3) evidence from earlier online trials indicates that much shorter interventions (eg, over 5 sessions rather than 8) are associated with positive outcomes, and that very short interventions can be effective [49]. Consequently, we doubt that the intervention website itself was the reason for the lack of effect on the primary outcome

measure. Another criticism is that choosing a multi-modal treatment results in the findings not being clear-cut, as any component might be the effective one. However, we would argue that prevention trials, in direct comparison to treatment trials, are in their infancy. Trials that demonstrate the effectiveness of multi-component interventions represent the first stage of a prevention research program, where, once an effect can be demonstrated, further research would normally then investigate the effect of the subcomponents.

The number of individuals in the current prevention study developing a diagnosis over the 12-month period following the intervention was unexpectedly low. However, elsewhere, comparable rates of 8.6% for the intervention and 4.44% in usual care groups have been reported [14]. These low rates may be due to regression to the mean and to the low threshold of anxiety for recruitment to the trial. The 6-month interval to determine caseness is relatively short with respect to prevention trials in the sense that a longer interval permits more opportunity to develop a disorder and thus to judge the effectiveness of the interventions. The choice of the 6-month interval for the MINI was made in order to maximize follow-up numbers (ie, to determine caseness) as dropouts were expected to increase over the 12-month follow-up period. We reasoned that we could determine proxy measures of caseness at 12 months via the GAD-7. A planned 3-year follow-up using GAD-7 as the primary outcome has commenced and we will report anxiety and depression outcomes at this time. Attrition was higher than expected at about 35% at post-test, although similar rates of dropout have been reported for face-to-face CBT [50].

Consistent with previous eHealth trials, data completion was higher in the control condition and was lowest for the active website condition with automated emails—a finding consistent with earlier trials [35,51]). The lower attrition in the control groups has been attributed to participant burden. Whether online or face-to-face, psychological interventions can be hard work, even threatening, and are often associated with dropout [52]. We also had substantial loss to recruitment between an expression of interest to the trial, and enrollment following invitation to undertake baseline measures. The reasons for failure to take up enrollment are not clear, but the trial was configured to allow enrollment within a week of consent. Delays in telephone and email contact often stretched the recruitment process, although there were strict time limits on each of the processes involved. The multi-modal nature of recruitment made identification of the reasons for non-response very complex, as the screener was conducted by post, the MINI assessment by telephone call, and the baseline invitation by email to an online survey.

A related issue is the role of contact in promoting adherence. In the present trial, adherence to the website was increased by contact via email or telephone. However, increased contact does not always result in improved adherence. A study we undertook with crisis call centers showed that adherence was much lower in participants who were provided with a website and telephone support compared to individuals without such support [53]. More research is required to examine for whom and under what circumstances telephone contact can increase adherence, and the factors that lead to increased dropout. We have reviewed factors that predict adherence for online programs [54]. We also acknowledge that the effects of the intervention in an adolescent rather than a young adult sample might have been more evident, since GAD or worry might emerge in this period. Levels of attrition need to be considered, since these were high. In addition, the sample excluded concurrent other diagnoses such as social phobia and PTSD, reducing generalizability. The trial raises the important question of how best to keep symptomatic people engaged in interventions. One possibility is to change orientation toward healthy living and offer prevention for GAD by stealth. An alternative approach is to constrain participants from dropping out through structure (eg, curriculum activities in schools or induction programs in workforces). These possibilities are currently being pursued in other research projects.

Conclusions

Despite a number of limitations, the present trial represents a methodologically rigorous, well-executed prevention trial, which for the first time examines the effectiveness of the prevention of GAD in symptomatic 18-30 year olds in the community using online technologies. Diagnosis was established using a telephone interview at baseline and 6 months, and GAD diagnosis was an exclusion factor at commencement, ensuring that it was a genuine prevention trial. A post-hoc power analysis found that we had approximately 95% power to find a between-groups effect of d=0.3 between the Active website alone and Control website alone, indicating that the trial was sufficiently powered. Preference for trial condition was measured and assessed for its effect, and the study aimed to determine push and pull factors that might influence uptake and efficaciousness. In this trial, we were not able to demonstrate the preventative effects of the website on anxiety symptoms as measured by the GAD-7. There were indications that prevention was operating in one of the five conditions (email plus active website) on a number of the secondary measures. The 3-year follow-up will provide a stronger test of whether secondary outcomes such as anxiety sensitivity are modifiable in response to a website with email reminders and to determine whether anxiety symptoms and caseness are averted with a longer lapse of time.

Acknowledgments

The trial was funded by NHMRC Project Grant 525419. Helen Christensen is supported by NHMRC Fellowship 1056964. Philip Batterham is supported by NHMRC Fellowship 1035262. Kathleen Griffiths is supported by NHMRC Fellowship 1059620. Both Philip Batterham and Andrew Mackinnon had full access to the data in the study and had input into the final analyses. Philip Batterham takes full responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

Griffiths is the Director of Australian National University e-hub self-help services, which delivers the public access version of e-couch GAD and co-authored the GAD stream of e-couch; Bennett is the development manager of e-hub services. Neither derives a personal financial benefit from e-couch.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [55].

[PDF File (Adobe PDF File), 994KB - jmir_v16i9e199_app1.pdf]

References

1. 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]
2. Judd LL, Kessler RC, Paulus MP, Zeller PV, Wittchen HU, Kunovac JL. Comorbidity as a fundamental feature of generalized anxiety disorders: results from the National Comorbidity Study (NCS). Acta Psychiatr Scand Suppl 1998;393:6-11. [Medline: 9777041]
3. Tyrer P, Baldwin D. Generalised anxiety disorder. Lancet 2006 Dec 16;368(9553):2156-2166. [doi: 10.1016/S0140-6736(06)69865-6] [Medline: 17174708]
4. Greenberg PE, Sisitsky T, Kessler RC, Finkelstein SN, Berndt ER, Davidson JR, et al. The economic burden of anxiety disorders in the 1990s. J Clin Psychiatry 1999 Jul;60(7):427-435. [Medline: 10453795]
5. Feldner MT, Zvolensky MJ. Prevention of anxiety psychopathology: a critical review of the empirical literature. Clinical Psychology: Science and Practice 2004;11(4):405-424. [doi: 10.1093/clipsy.bph098]
6. 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]
7. 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]
8. Kessler RC, Greenberg PE, Mickelson KD, Meneades LM, Wang PS. The effects of chronic medical conditions on work loss and work cutback. J Occup Environ Med 2001 Mar;43(3):218-225. [Medline: 11285869]
9. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Br J Psychiatry 2004 Jun;184:526-533 [FREE Full text] [Medline: 15172947]
10. Institute of Medicine (US) Committee on Prevention of Mental Disorders. In: Mrazek PJ, Haggerty RJ, editors. Reducing risks for mental disorders: Frontiers for preventive intervention research. Washington, DC: National Academy Press; 1994.
11. van't Veer-Tazelaar PJ, van Marwijk HW, van Oppen P, van Hout HP, van der Horst HE, Cuijpers P, et al. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. Arch Gen Psychiatry 2009 Mar;66(3):297-304. [doi: 10.1001/archgenpsychiatry.2008.555] [Medline: 19255379]
12. van't Veer-Tazelaar PJ, van Marwijk HW, van Oppen P, van der Horst HE, Smit F, Cuijpers P, et al. Prevention of late-life anxiety and depression has sustained effects over 24 months: a pragmatic randomized trial. Am J Geriatr Psychiatry 2011 Mar;19(3):230-239. [Medline: 21425519]
13. Pitceathly C, Maguire P, Fletcher I, Parle M, Tomenson B, Creed F. Can a brief psychological intervention prevent anxiety or depressive disorders in cancer patients? A randomised controlled trial. Ann Oncol 2009 May;20(5):928-934 [FREE Full text] [doi: 10.1093/annonc/mdn708] [Medline: 19126633]
14. Dozeman E, van Marwijk HW, van Schaik DJ, Smit F, Stek ML, van der Horst HE, et al. Contradictory effects for prevention of depression and anxiety in residents in homes for the elderly: a pragmatic randomized controlled trial. Int Psychogeriatr 2012 Aug;24(8):1242-1251. [doi: 10.1017/S1041610212000178] [Medline: 22436082]
15. Joling KJ, van Marwijk HW, Smit F, van der Horst HE, Scheltens P, van de Ven PM, et al. Does a family meetings intervention prevent depression and anxiety in family caregivers of dementia patients? A randomized trial. PLoS One 2012;7(1):e30936 [FREE Full text] [doi: 10.1371/journal.pone.0030936] [Medline: 22303473]
16. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
17. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression: systematic review. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
18. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. Med J Aust 2010 Jun 7;192(11 Suppl):S4-11. [Medline: 20528707]
19. Chiu TM, Eysenbach G. Stages of use: consideration, initiation, utilization, and outcomes of an internet-mediated intervention. BMC Med Inform Decis Mak 2010;10:73 [FREE Full text] [doi: 10.1186/1472-6947-10-73] [Medline: 21092275]

20. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
21. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. J Med Internet Res 2013;15(1):e23 [FREE Full text] [doi: 10.2196/jmir.2151] [Medline: 23363466]
22. Woodall WG, Buller DB, Saba L, Zimmerman D, Waters E, Hines JM, et al. Effect of emailed messages on return use of a nutrition education website and subsequent changes in dietary behavior. J Med Internet Res 2007;9(3):e27 [FREE Full text] [doi: 10.2196/jmir.9.3.e27] [Medline: 17942389]
23. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Internet-based CBT for depression with and without telephone tracking in a national helpline: randomised controlled trial. PLoS One 2011;6(11):e28099 [FREE Full text] [doi: 10.1371/journal.pone.0028099] [Medline: 22140514]
24. Linden M, Zubraegel D, Baer T, Franke U, Schlattmann P. Efficacy of cognitive behaviour therapy in generalized anxiety disorders. Results of a controlled clinical trial (Berlin CBT-GAD Study). Psychother Psychosom 2005;74(1):36-42. [doi: 10.1159/000082025] [Medline: 15627855]
25. Jayakody K, Gunadasa S, Hosker C. Exercise for anxiety disorders: systematic review. Br J Sports Med 2014 Feb;48(3):187-196. [doi: 10.1136/bjsports-2012-091287] [Medline: 23299048]
26. Herring MP, Jacob ML, Suveg C, Dishman RK, O'Connor PJ. Feasibility of exercise training for the short-term treatment of generalized anxiety disorder: a randomized controlled trial. Psychother Psychosom 2012;81(1):21-28. [doi: 10.1159/000327898] [Medline: 22116310]
27. Borkovec TD, Costello E. Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. J Consult Clin Psychol 1993 Aug;61(4):611-619. [Medline: 8370856]
28. McHugh RK, Whitton SW, Peckham AD, Welge JA, Otto MW. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. J Clin Psychiatry 2013 Jun;74(6):595-602. [doi: 10.4088/JCP.12r07757] [Medline: 23842011]
29. Christensen H, Griffiths KM, Mackinnon AJ, Kalia K, Batterham PJ, Kenardy J, et al. Protocol for a randomised controlled trial investigating the effectiveness of an online e health application for the prevention of Generalised Anxiety Disorder. BMC Psychiatry 2010;10:25 [FREE Full text] [doi: 10.1186/1471-244X-10-25] [Medline: 20302678]
30. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
31. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
32. McIntosh A, Cohen A, Turnbull N, Esmonde L, Dennis P, Eatock J. Clinical guidelines and evidence review for panic disorder and generalised anxiety disorder. Sheffield: University of Sheffield/London: National Collaborating Centre for Primary Care; 2004.
33. 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]
34. 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]
35. 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]
36. 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]
37. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
38. Griffiths KM, Crisp D, Christensen H, Mackinnon AJ, Bennett K. The ANU WellBeing study: a protocol for a quasi-factorial randomised controlled trial of the effectiveness of an Internet support group and an automated Internet intervention for depression. BMC Psychiatry 2010;10:20 [FREE Full text] [doi: 10.1186/1471-244X-10-20] [Medline: 20211025]
39. Fresco DM, Mennin DS, Heimberg RG, Turk CL. Using the Penn State Worry Questionnaire to identify individuals with generalized anxiety disorder: a receiver operating characteristic analysis. J Behav Ther Exp Psychiatry 2003;34(3-4):283-291. [doi: 10.1016/j.jbtep.2003.09.001] [Medline: 14972674]
40. Peterson R, Reiss R. Anxiety Sensitivity Index Manual (2nd ed.). Worthington, OH: IDS Publishing Corporation; 1992.
41. 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]
42. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003 Jun 18;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095] [Medline: 12813115]

43. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
44. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
45. Verbeke G, Molenberghs G. Linear mixed models for longitudinal data. New York: Springer; 2000.
46. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
47. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [Medline: 17339617]
48. Christensen H, Mackinnon A, Batterham P, O'Dea B, Guastella AJ, Griffiths KM, et al. The effectiveness of an online e-health application compared to attention placebo or sertraline in the treatment of generalised anxiety disorder. J Med Internet Res 2014:In Press (forthcoming).
49. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
50. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
51. Donker T, Bennett K, Bennett A, Mackinnon A, van Straten A, Cuijpers P, et al. Internet-delivered interpersonal psychotherapy versus internet-delivered cognitive behavioral therapy for adults with depressive symptoms: randomized controlled noninferiority trial. J Med Internet Res 2013;15(5):e82 [FREE Full text] [doi: 10.2196/jmir.2307] [Medline: 23669884]
52. Bados A, Balaguer G, Saldaña C. The efficacy of cognitive-behavioral therapy and the problem of drop-out. J Clin Psychol 2007 Jun;63(6):585-592. [doi: 10.1002/jclp.20368] [Medline: 17457848]
53. Christensen H, Farrer L, Batterham PJ, Mackinnon A, Griffiths KM, Donker T. The effect of a web-based depression intervention on suicide ideation: secondary outcome from a randomised controlled trial in a helpline. BMJ Open 2013;3(6):e002886 [FREE Full text] [doi: 10.1136/bmjopen-2013-002886] [Medline: 23811172]
54. 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]
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

ASI: Anxiety Sensitivity Index
AUDIT: Alcohol Use Disorders Identification Test
CBT: Cognitive Behavioral Therapy
CES-D: Centre for Epidemiologic Studies-Depression
GAD: generalized anxiety disorder
ITT: intent-to-treat
MAR: missing at random
MINI: Mini International Neuropsychiatric Interview
PHQ-9: Patient Health Questionnaire 9-item
PMR: progressive muscle relaxation
PSWQ: Penn State Worry Questionnaire

Edited by G Eysenbach; submitted 01.05.14; peer-reviewed by C Reynolds III, A Parks, N Titov, C Donovan; comments to author 27.05.14; revised version received 30.06.14; accepted 04.08.14; published 02.09.14.

Please cite as:
Christensen H, Batterham P, Mackinnon A, Griffiths KM, Kalia Hehir K, Kenardy J, Gosling J, Bennett K
Prevention of Generalized Anxiety Disorder Using a Web Intervention, iChill: Randomized Controlled Trial
J Med Internet Res 2014;16(9):e199
URL: http://www.jmir.org/2014/9/e199/
doi:10.2196/jmir.3507
PMID:25270886

Review

Online and Social Networking Interventions for the Treatment of Depression in Young People: A Systematic Review

Simon M Rice[1,2], BBSc, BSc(Hons), DipEd, MPsych(Clin), GradCertClinEpidemiology, PhD; Joanne Goodall[1], BA, PGDipPsych; Sarah E Hetrick[1,2], BA, PGCertHSc, MA, DPsych; Alexandra G Parker[1,2], BA(Hons), MPsych(Clin), PhD; Tamsyn Gilbertson[1], BA (Hons), MSc; G. Paul Amminger[1,3], MD, PhD, FRANZCP; Christopher G Davey[1], MBBS(Hons), MPsychiatry, FRANZCP, PhD; Patrick D McGorry[1], MD, FRCP, FRANZCP, PhD; John Gleeson[4], BA(Hons), MPsych(Clin), PhD; Mario Alvarez-Jimenez[1], BSc(Hons), DClinPsy, MAResearchMeth, PhD

[1]Orygen Youth Health Research Centre, Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

[2]headspace Centre of Excellence in Youth Mental Health, University of Melbourne, Melbourne, Australia

[3]Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria

[4]School of Psychology, Australian Catholic University, Melbourne, Australia

Corresponding Author:
Simon M Rice, BBSc, BSc(Hons), DipEd, MPsych(Clin), GradCertClinEpidemiology, PhD
Orygen Youth Health Research Centre
Centre for Youth Mental Health
University of Melbourne
35 Poplar Rd
Parkville
Melbourne, 3052
Australia
Phone: 61 419497599
Fax: 61 93873003
Email: simon.rice@unimelb.edu.au

Abstract

Background: Major depression accounts for the greatest burden of all diseases globally. The peak onset of depression occurs between adolescence and young adulthood, and for many individuals, depression displays a relapse-remitting and increasingly severe course. Given this, the development of cost-effective, acceptable, and population-focused interventions for depression is critical. A number of online interventions (both prevention and acute phase) have been tested in young people with promising results. As these interventions differ in content, clinician input, and modality, it is important to identify key features (or unhelpful functions) associated with treatment outcomes.

Objective: A systematic review of the research literature was undertaken. The review was designed to focus on two aspects of online intervention: (1) standard approaches evaluating online intervention content in randomized controlled designs (Section 1), and (2) second-generation online interventions and services using social networking (eg, social networking sites and online support groups) in any type of research design (Section 2).

Methods: Two specific literature searches were undertaken. There was no date range specified. The Section 1 search, which focused on randomized controlled trials, included only young people (12-25 years) and yielded 101 study abstracts, of which 15 met the review inclusion criteria. The Section 2 search, which included all study design types and was not restricted in terms of age, yielded 358 abstracts, of which 22 studies met the inclusion criteria. Information about the studies and their findings were extracted and tabulated for review.

Results: The 15 studies identified in Section 1 described 10 trials testing eight different online interventions, all of which were based on a cognitive behavioral framework. All but one of the eight identified studies reported positive results; however, only five of the 15 studies used blinded interviewer administered outcomes with most trials using self-report data. Studies varied significantly in presentation of intervention content, treatment dose, and dropout. Only two studies included moderator or clinician input. Results for Section 2 were less consistent. None of the Section 2 studies reported controlled or randomized designs. With the exception of four studies, all included participants were younger than 25 years of age. Eight of the 16 social networking studies

reported positive results for depression-related outcomes. The remaining studies were either mixed or negative. Findings for online support groups tended to be more positive; however, noteworthy risks were identified.

Conclusions: Online interventions with a broad cognitive behavioral focus appear to be promising in reducing depression symptomology in young people. Further research is required into the effectiveness of online interventions delivering cognitive behavioral subcomponents, such as problem-solving therapy. Evidence for the use of social networking is less compelling, although limited by a lack of well-designed studies and social networking interventions. A range of future social networking therapeutic opportunities are highlighted.

(J Med Internet Res 2014;16(9):e206) doi:10.2196/jmir.3304

KEYWORDS

Internet; depression; young adult; adolescent; social networking; support groups; review

Introduction

Major depressive disorder is reported as the leading cause of disability worldwide [1] and accounts for the greatest burden of all diseases globally [2-4]. The effects of depression are considerable and are associated with a range of negative outcomes [5]. Depression typically first manifests during adolescence or young adulthood (up to 25 years) [6] and tends to display a worsening pattern over the course of repeated episodes, including a lack of responsiveness to initially effective treatments [7]. Relapse rates after a first episode of depression are 20-30%, compared to 70-80% for those who have experienced two or more episodes [8], and research has established that relapse rates in children and adolescents range between 34% and 75% within 1-5 years after the first depressive episode [9].

Cognitive behavioral therapy (CBT) is recommended as a first line treatment for depression [10-13]. However, research has suggested that CBT may be only modestly effective in treating depression in young people [14]. Furthermore, mental health service use is low among young people for a variety of reasons, including distrust of health professionals and stigma [15]. Hence, there is a need for better youth depression treatment strategies that maximize engagement and reduce likelihood of relapse or recurrence [16,17].

Given young people's enthusiasm for Web-based communication, the development of innovative, online psychosocial interventions may help improve treatment acceptability and engagement for young people experiencing depressive symptoms [13]. Further, social networking sites (SNS) are rapidly becoming an essential avenue for social communication and support [18], especially among young people [19], and could be pivotal in engagement with mental health services [20]. Using such avenues of support in the treatment of depression may also have some benefit.

In recent years, a range of online interventions have been successfully tested for the management of a number of mental disorders, with research supporting the efficacy of these interventions in alleviating anxiety and depressive symptoms [21-23]. Online interventions have been reported as being as effective as face-to-face therapy [15], and some countries now recommend the use of online interventions within clinical guidelines for the treatment of depression [24].

To date, there has been a relative lack of online intervention studies focusing on young people with depression [25]. While young people tend to choose face-to-face support as their preferred mode of help-seeking for depression [26], a significant number also indicate preference for online intervention given the added anonymity and immediacy associated with the online environment [27]. Furthermore, online interventions have the potential to reach young people who may not be inclined or able to seek help from traditional sources given their ability to transcend geographical boundaries and provide 24-hour accessibility and reach [15].

Service delivery gaps in the provision of youth specific psychotherapeutic services [28], in addition to modest treatment outcomes [14], have resulted in unmet need with adverse consequences for both the affected young people and their wider communities [15]. Consequently, there has been an increase in research evaluating the plausibility of online interventions for the treatment and relapse prevention of depression. The aim of the present study was to systematically appraise the relevant literature of online approaches to depression treatment in young people. This approach was undertaken using separate search strategies to identify studies of (1) standard online interventions, and (2) second-generation online interventions using social networking functions.

Methods

Searching

Electronic searches were conducted in June 2013 of the PsycINFO, Cochrane Library, Embase, and MEDLINE databases accessed via the University of Melbourne library. Search strategies were devised using relevant subject headings for each database. In order to capture all relevant online interventions, additional free text words were identified by experts who have published in youth mental health online interventions and systematic review methodology. All studies retained for analysis needed to meet the study inclusion criteria outlined below. Further manual searching of reference lists was also undertaken, in addition to a "soft search" in the Google Scholar database for Section 1 for "Online interventions" and Section 2 for "Social networking and depression". Search terms are listed in Multimedia Appendix 1. No date range was specified.

Inclusion Criteria

Section 1: Online Interventions for Depression

Included studies had to meet each of the following criteria: (1) use a randomized controlled trial (RCT) design, (2) test any therapeutic modality delivered online, (3) use either an open access (not restricted) or referred (participation through invitation only and supplied with a password to access site) intervention, (4) focus on prevention (using a universal or targeted approach), treatment, or relapse prevention, and (5) include participants aged between 12 and 25 years.

Section 2: Social Networking and Depression

Included studies met each of the following criteria: (1) focus on social networking services, defined as online interactions where the reader is also able to write, and where the social network is predefined, (2) either examine social networking websites that aim to prevent, treat, or provide relapse prevention for depression, or describe associations between the use of general social networking sites and depressive symptoms, and (3) include user-to-user contact. Given the early development stage of social networking interventions, no restrictions were placed on participant age or study design.

Data Synthesis

For each section, search results were independently screened by 2 researchers to establish their relevance for inclusion in the appropriate section. Any discrepancy between the researchers was referred to a third researcher involved in establishing the search strategy for the final decision. In order to assist the narrative synthesis of the included studies, a predefined data extraction template was designed for Sections 1 and 2.

Results

Section 1: Online Interventions for Depression

Study Characteristics

We identified 101 articles, of which 15 studies met inclusion criteria (see Multimedia Appendix 2). According to characteristics of samples recruited, these studies were categorized as either prevention (ie, participants identified as at risk of depression or not meeting current diagnostic criteria for major depressive disorder [MDD], n=9), intervention (ie, participants identified as meeting clinical diagnostic criteria for MDD, n=5), or combined prevention and intervention (ie, mixed sample of participants meeting and not meeting diagnostic criteria for MDD, n=1). None of the identified 15 studies focused on relapse prevention post acute-phase treatment.

With the exception of the one combined prevention and intervention study [29], all included trials reported positive findings. However, only four trials used blinded observer ratings [24,30-32]. The majority (n=9) of trials used participant self-report data as outcome variables.

The 15 identified studies reported on data from 10 separate trials. Of these, there were nine separate interventions tested: Cognitive Behavioral Analysis [33], MoodGYM [21,34], Cognitive Behavioral Skills Training Program [35], SPARX [31,36], CATCH-IT [24,30,32,37,38], Internet Problem Solving Therapy [29], Blues Blaster [39], Computerized CBT [40], and Master Your Mood [41]. There was one secondary publication for MoodGYM and four secondary publications for CATCH-IT. All of these interventions used a cognitive behavioral treatment approach, targeting either depression, or depression and comorbid anxiety. With the exception of the Cognitive Behavioral Analysis intervention [33], each of the interventions were designed to be completed in modules. The number of modules ranged from 4 (Cognitive Behavioral Skills Training Program) to 14 (CATCH-IT). Based on descriptions of the intervention modules presented in the published papers, varied key CBT content tended to be present (eg, psychoeducation, behavioral activation, thought monitoring), though there was some variation across interventions. None of the included studies incorporated ongoing social networking.

The included MoodGYM, Cognitive Behavioral Analysis, and SPARX interventions were delivered via computer in a group classroom setting, with supervision from a classroom teacher/tutor. The remaining interventions were designed to be delivered in a self-paced and self-directed manner. Studies varied in duration (eg, dose) both in terms of total length of intervention access, which ranged from 5 weeks (eg, MoodGYM) to 32 weeks (eg, Cognitive Behavioral Skills Training Program), and individual module length, which ranged from 20-40 minutes (MoodGYM) to 90 minutes (Master Your Mood).

The majority of included studies involved a fully automated intervention, with only two of the eight interventions including moderator involvement, or feedback to participants. Of the moderated interventions, the Internet Problem Solving Therapy intervention used email support from the moderator (with messages followed up within a 3-day period) [29], plus several phone calls to monitor safety. The Master Your Mood intervention adopted a closed chat-room moderation style. This was facilitated by trained professionals, and the intervention content was presented in an online group format (6 sessions, 90 minutes duration, 6 participants maximum) [41]. The CATCH-IT intervention included several motivational phone calls, provided only to those in the motivational interviewing condition.

Dropout rates were reported for 14 of the 15 studies and varied significantly, ranging from 3%-41%. The only study with a dropout rate of less than 5% used the SPARX avatar-based gaming intervention platform [36] with the accompanying RCT of SPARX [31] also reporting a low dropout rate (9.6%). Four studies [29,35,39,41] reported attrition rates above 15%. These studies varied in their level of moderator and peer contact: two were fully automated (Cognitive Behavioral Skills Training; Blues Blaster) [35,39], one used occasional moderator phone contact (Internet Problem Solving Therapy) [29], and one included direct peer-to-peer contact via chat room sessions (Master Your Mood) [41].

Efficacy: Prevention Studies

The prevention studies (Blues Blaster, Cognitive Behavioral Analysis, MoodGYM, and CATCH-IT) demonstrated intervention efficacy among university students [33,39], secondary students (although effects were significant only for

male participants) [21], and adolescents identified as at-risk of depressive disorder [24,37,38,42,43]. Several studies using the CATCH-IT intervention indicated that motivational interviewing provided by a primary care practitioner prior to intervention commencement aided the young person's engagement and symptom improvement [24,37,38]. However, the CATCH-IT intervention had not been compared to no intervention or a comparison intervention. For the prevention studies, comparison groups varied from waitlist control, attention placebo, and those not receiving motivational interviewing.

Efficacy: Intervention Studies

The intervention studies demonstrated superiority of the online intervention to the comparison treatment. These involved the Cognitive Behavioral Skills Training Program, SPARX, Computerized CBT, and Master Your Mood. Comparison groups were either treatment as usual (pharmacotherapy and psychosocial services) [35], waitlist control [36,41], or brief psychoeducation [40]. One of the more novel intervention studies (the large SPARX RCT intervention that integrated CBT concepts within an online fantasy game) demonstrated non-inferiority relative to treatment as usual provided by youth clinics, school-based counseling services, or general practitioners [31].

Combined Prevention and Intervention

The trial that targeted its Internet Problem Solving Therapy intervention to young people who were either at risk of depression or had depression showed no difference between the intervention group and the waitlist control. This trial recruited a relatively small (n=45) yet heterogeneous sample [29]. Participants were 12-18 years, with mild/moderate depressive and/or anxiety symptoms. Depressive and anxiety symptoms declined in both the intervention and waitlist control groups, with no added benefit of the Internet-delivered problem solving intervention.

Section 2: Social Networking and Depression

Study Characteristics

A total of 358 abstracts were screened, yielding 22 studies meeting inclusion criteria (see Multimedia Appendix 2). Articles were grouped into three overarching categories: (1) social networking sites (SNS, n=16) [44-59] defined as sites with a non-specific user purpose, (2) online support groups (OSG; n=4) [22,60-62] defined as sites with a specific purpose of seeking and sharing health-related information and related personal experiences, and (3) general commentary (n=1) [63]. One additional study undertook a combined analysis of SNS and OSG use [64]. There were 8 studies with a clinical focus (eg, recruiting help-seeking individuals) and 14 studies recruiting from a general population. None of the SNS or OSG studies focused on relapse prevention post acute-phase treatment.

The majority of included studies reported on data collected via online questionnaires. All but five of the identified SNS studies reported on participant data, with the remaining studies including one systematic review, one general literature review, two pieces discussing SNS in the context of clinical practice, and a descriptive networking study. Ages ranged between 12-85 years; however, with the exception of the systematic review paper [22], the two case study reports [54,59], and the analysis of online support group postings [60], all included studies focused on young people (eg, those aged 25 years or younger).

Studies Focusing on Social Networking Sites

Findings related to the benefits of SNS on depression were inconsistent. Eight of the SNS studies reported positive findings [46,48-50,55,56,58,59], five reported mixed or unclear findings [45,47,53,54,64], and three studies reported negative findings [44,51,52]. The SNS studies are discussed relative to potential benefits and harms.

Potential Benefits of Using Social Networking Sites

Direct one-to-one interactive communication (as opposed to passive SNS use) was associated with greatest well-being and least depression [45]. SNS use was associated with potential mental health benefits including socialization, facilitation of supportive relationships, belongingness, self-esteem, communication, and learning [46,63]. Positive findings highlighted key opportunities in the relationship between SNS use and depression for those from marginalized groups at risk of social isolation [46]. This study also highlighted the ubiquitous nature of young people's use of SNS and benefits of educational outcomes, facilitating supportive relationships, identity formation, and self-esteem. Three studies using an online questionnaire design identified the potential for SNS to aid in the prevention of symptom onset and symptom remission by functioning as a screening tool for at-risk individuals via status update posts [49,50,55]. Potential interventions, such as targeted advertising triggered by keywords posted on social networking sites, could present relevant cost-effective options for mental health support [55]. Social networking was also identified as a means to re-establish friendships following social withdrawal for those experiencing depression [59].

Potential Harms of Using Social Networking Sites

Identified risks included cyberbullying, harassment, sexting, privacy concerns, and SNS-induced depression in the association between young people's use of SNS and depression [63]. As expected, negative interactions with SNS were associated with greater depressive symptomology and negative affect [44,57]. However, at the general population level, there was little evidence of a direct relationship between depression and general SNS use [47,48]. This was highlighted by a weak relationship (r=.15) between high school students' use of SNS and depression symptoms [51]. Although use of SNS was identified as both a possible causal contagion factor in a cluster of suicides among 15-18 year olds [52], suicidal posts on SNS were also identified as offering an opportunity for identification and intervention [54].

Studies Focusing on Online Support Groups

The included systematic review of OSG use reported a lack of high-quality research examining the effects of OSGs on depression. The included systematic review reported that only two low-quality studies examined the effectiveness or efficacy of OSG use on depression with mixed results [22]. This finding

is echoed in the present review, with no RCTs found that test OSG interventions for depression.

Of the included OSG studies in our review, three reported overall positive findings [60-62] and recommended OSGs for people experiencing depression. These studies reported that regular participation in OSGs was associated with participants receiving greater online and subsequent offline emotional support for depression and that OSG use may complement formal care, provide an environment for knowledge exchange, and inspiration for coping with depression [60].

In contrast, one study reported that use of OSGs (but not SNS) was associated with increased suicidal ideation for young people 14-24 years over a 12-month period [64]. Other general disadvantages of OSG membership included reports that some individuals experienced participation as burdensome (and experienced symptom exacerbation) following negative comments posted by other site members [57]. Further, excessive reliance on online support was also apparent for some individuals. Preference for online support tended to be evident in cases where external supportive networks were not apparent [65].

Discussion

Principal Findings

Young people are eager users of new technology, and online interventions offer an opportunity for mental health support that is immediate and cost-effective [15]. This review draws together two threads of emerging inquiry related to young people's use of online technology for managing symptoms of depression. Findings related to RCTs of online interventions in young people indicated largely positive effects. This is consistent with broader findings from reviews of Web-based interventions undertaken with adult populations [66] and university students [67,68]. The overall impact of SNS use was less clear but may in part be due to significant heterogeneity in research design, target population, and intervention assessed.

Online Interventions and Depression

With the exception of Internet Problem Solving Therapy, all studies highlighted the benefits of online interventions for prevention and treatment of major depression in young people. However, caution is required in interpreting this finding given that few studies used blinded interviewer outcome assessments. General recommendations included the feasibility of widespread roll-out of online interventions across school-based populations [21], comparative low cost [35], applicability to group-based delivery [41], appeal to young people outside of mainstream educational settings [36], and the ability of online interventions to assist with unmet clinical need [31]. Furthermore, 6-month [30] and 12-month [32] follow-up of the CATCH-IT intervention suggests enduring clinical benefits. However, the CATCH-IT intervention has not been evaluated against a competing intervention or control group.

Based on sample characteristics, studies were classified as prevention, intervention, or both. However, in practice, included studies tended to make little distinction regarding this and were broad in their disorder of focus. For example, the MoodGYM intervention is designed to both prevent and decrease depression symptoms [21], the Blues Blaster intervention was evaluated as a prevention intervention, despite being adapted from the Adolescent Coping With Depression clinical intervention [39,69], and the Internet Problem Solving Therapy intervention was designed to treat both depression and anxiety [29]. A logical advance for the field would be greater specificity in the development and testing of online interventions that specially match treatment focus to stage of disorder (eg, primary prevention, clinical intervention, relapse prevention).

Participant attrition is an important issue in the use of online interventions. Included studies that promoted greater engagement (either through use of motivational interviewing or content innovation) tended to report lower attrition rates. Automated self-help services require significant motivation and self-discipline [29], and this may be expecting too much from young people experiencing depression. Indeed, low motivation of young people and possible dissatisfaction with delayed moderator/clinician input were offered as an explanation for outcomes of the only included Section 1 study to report no benefit of the online intervention [29].

Ongoing engagement is also an important factor for online interventions. The effect of high intervention adherence (relative to low intervention adherence) was found to improve treatment outcomes of the MoodGYM intervention [34]. This is a noteworthy finding given some interventions report relatively low levels of engagement (eg, the CATCH-IT study reported that 9% and 22% of those allocated to respective treatment groups failed to access the intervention at all [24]).

Social Networking—Therapeutic Opportunities

The primary focus of the studies reviewed in Section 2 was the exploration of the relationship between engagement in SNS (including OSG use) and depressive symptomology. This is an emerging research area, and there was significant heterogeneity across study designs. Hence, results must be interpreted with caution.

Research suggests that some young people may be more willing to disclose information on an SNS than in person [70]. While SNS use may result in unhealthy online interactions for some [63], for many vulnerable young people SNSs may provide needed opportunities for social support and identity formation [45]. Indeed, longitudinal research has shown that Internet use for social support is associated with declines in depression [71]. Because of the noted risks (alongside identified benefits) of SNS use for depression, researchers and clinicians must work towards the development of protocols to integrate such functionality within a safe and supportive framework and provide positive alternatives to non-therapeutic SNS.

Given the early phase of mental health–related research on the use of SNS, significant knowledge gaps in research are to be expected. While some of the SNS studies reported positive findings, others identified adverse outcomes or no associations at all. It is possible that SNSs exert a bidirectional effect on depression symptoms or that the relationship is mediated by the intervention content, safety, and type of interaction. For example, positive online interactions may lead to increased

social support and reduced depression, and negative online interactions (or with a negative focus) may lead to increased depression and perceived burden.

There may also be qualitative differences between young people who use particular websites for help seeking or discussion of mental health issues. For example, previous research has shown that those higher in hopelessness (a key predictor of suicidal ideation) may be more likely to engage in blogging type sites (eg, OSGs) versus sites focused on briefer posts and content [64]. OSGs are communities designed to provide support, acceptance, and knowledge exchange between members overcoming similar adversities and may attract young people with particular personality characteristics. Given the overall study design limitations (eg, lack of control or comparison group), well-designed trials are urgently needed in this area [22].

Participation in online social networking has become ubiquitous and routine in the lives of young people. As such, there is both a great opportunity and an urgent need to identify and make use of the potential benefits of such sites for young people's well-being. Provided that the SNS experience is primarily positive, time spent on SNSs has the potential to operate as a preventative or therapeutic medium for individuals with depression and may complement traditional therapy for more severe forms of the disorder.

Implications and Limitations

Adherence and engagement are essential to effective and efficient delivery of online interventions [34]. Although the population reach and cost-effectiveness of online interventions makes their use attractive, significant work remains to be done in refining and better targeting online interventions to maximize their effectiveness. As online interventions evolve, they will incorporate greater dynamism and functionality. It is possible that automated self-help interventions may lose their appeal, and the development of future interventions must consider features and treatment strategies that promote engagement. A further and necessary advancement, given potential risk issues associated with unmonitored suicidality or symptom deterioration, will likely include the use of real-time moderator input and integrated crisis support within the online environment. This was lacking in the studies identified in our review. Emerging models of online moderation are evolving, including improving the integration of clinician input into online interventions [25]. For example, the moderated online social therapy (MOST) framework [72-74] provides a methodology for promoting ongoing engagement through online therapy, social networking, and regular clinician support for relapse prevention of serious mental health problems in young people.

The next generation of online interventions will include refinements and functionality not possible in previous technologies. There is significant scope for greater responsiveness and immediacy, including real-time clinician input and customized feedback. In addition, the next generation of online interventions offers an opportunity for better matching of intervention content to phase of illness, and the role of online social and peer support [26,72].

Literature searching for the present review was current at June 2013, and it is likely that higher-quality studies evaluating SNS use will start to soon appear in scholarly journals. Promising new lines of research are emerging for the treatment of depression in young people using mobile phone-based interventions [75,76], live chat interventions [77], Web-based positive psychology interventions [78], and interventions focusing on comorbid depression and alcohol misuse [79]. Importantly, the current review did not include studies reporting on data from Twitter. Recent work using sentiment analysis and qualitative methodologies has highlighted that relative to those without depressive symptoms, Twitter users who experience depressive symptoms are more likely to post references to negative emotions or anger [80] and are more likely to view Twitter as a tool for emotional interaction and consoling oneself [81]. Complementing this work, emerging evidence suggests that salient emotional content from Twitter posts may effectively predict depression before actual symptom onset [82]. Taken together, these findings suggest potential broad implications for scalable early detection and intervention, and they highlight the need for a comprehensive review of studies using Twitter generated data.

The present review was limited by a lack of studies focusing on relapse prevention. Given that relapses in depression are high [7,8] the lack of relapse prevention studies using SNS or OSGs is concerning and is an area requiring urgent research and clinical attention. Illustrating this, a recent Cochrane review highlighted the need for rigorous RCTs to assess intervention efficacy and psychotherapy interventions aimed at preventing relapse given the lack of progress in this area [3]. Further, all included studies in Section 1 focused on cognitive behavioral interventions. Given that research suggests that CBT may be only modestly effective with young people [14], there is value in developing and evaluating relapse prevention interventions that draw on other therapeutic modalities (eg, strengths-based approaches) [83]. Hence, researchers and clinicians should focus their energies on the development and evaluation of innovative online treatments for depression relapse prevention.

Conclusions

Mental health interventions will increasingly make use of online technologies. A key challenge for these interventions will be having broad appeal and engagement. There is clear evidence that online interventions using a cognitive behavioral focus are promising in reducing depression symptomology in young people. Further study is required to identify the effectiveness of online interventions delivering cognitive behavioral subcomponents, such as problem solving therapy or other therapeutic approaches. Although evidence for the use of online social networking is less robust, such interventions exist in a dynamic space with significant opportunity for methodological advancement and rigor.

Acknowledgments

The authors gratefully acknowledge the HCF Health and Medical Research Foundation for providing funding for this research. MA-J is supported by a CR Roper Fellowship, Faculty of Medicine, Dentistry, and Health Science, the University of Melbourne, Australia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms.

[PDF File (Adobe PDF File), 33KB - jmir_v16i9e206_app1.pdf]

Multimedia Appendix 2

Summary of studies, Sections 1 and 2.

[PDF File (Adobe PDF File), 45KB - jmir_v16i9e206_app2.pdf]

References

1. World Health Organization, Depression. 2012. URL: http://www.who.int/mediacentre/factsheets/fs369/en/ [accessed 2014-07-17] [WebCite Cache ID 6R7cHOuOO]
2. Depression. Australian: Government Department of Health and Ageing; 2009. URL: http://www.health.gov.au/internet/publications/publishing.nsf/Content/mhSum-toc~mhSum-psy~mhSum-psy-dep [accessed 2014-07-17] [WebCite Cache ID 6R7cpexex]
3. Cox GR, Fisher CA, De Silva S, Phelan M, Akinwale OP, Simmons MB, et al. Interventions for preventing relapse and recurrence of a depressive disorder in children and adolescents. Cochrane Database Syst Rev 2012;11:CD007504. [doi: 10.1002/14651858.CD007504.pub2] [Medline: 23152246]
4. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. Lancet 2011 Jun 18;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6] [Medline: 21652063]
5. Cole DA, Peeke LG, Martin JM, Truglio R, Seroczynski AD. A longitudinal look at the relation between depression and anxiety in children and adolescents. J Consult Clin Psychol 1998 Jun;66(3):451-460. [Medline: 9642883]
6. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
7. Kendler K, Thornton L, Gardner C. Stressful life events and previous episodes in the etiology of major depression in women: an evaluation of the "kindling" hypothesis. Am J Psychiatry 2000 Aug;157(8):1243-1251. [Medline: 10910786]
8. Keller MB, Klerman GL, Lavori PW, Coryell W, Endicott J, Taylor J. Long-term outcome of episodes of major depression. Clinical and public health significance. JAMA 1984 Aug 10;252(6):788-792. [Medline: 6748178]
9. Kennard BD, Emslie GJ, Mayes TL, Hughes JL. Relapse and recurrence in pediatric depression. Child Adolesc Psychiatr Clin N Am 2006 Oct;15(4):1057-79, xi. [doi: 10.1016/j.chc.2006.05.003] [Medline: 16952775]
10. National Collaborating Centre for Mental Health (UK). Depression in children and young people: Identification and management in primary, community and secondary care. Leicester, UK: The British Psychological Society; 2005.
11. Birmaher B, Brent D, AACAP Work Group on Quality Issues, Bernet W, Bukstein O, Walter H, et al. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. J Am Acad Child Adolesc Psychiatry 2007 Nov;46(11):1503-1526. [doi: 10.1097/chi.0b013e318145ae1c] [Medline: 18049300]
12. Cheung AH, Zuckerbrot RA, Jensen PS, Ghalib K, Laraque D, Stein RE, GLAD-PC Steering Group. Guidelines for Adolescent Depression in Primary Care (GLAD-PC): II. Treatment and ongoing management. Pediatrics 2007 Nov;120(5):e1313-e1326 [FREE Full text] [doi: 10.1542/peds.2006-1395] [Medline: 17974724]
13. McDermott B, Baigent M, Chanen A, Fraser L, Graetz B, Hayman N, et al. Clinical practice guidelines: Depression in adolescents and young adults. Melbourne, Australia: beyondblue the National Depression Initiative; 2010.
14. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychol Bull 2006 Jan;132(1):132-149 [FREE Full text] [doi: 10.1037/0033-2909.132.1.132] [Medline: 16435960]
15. King R, Bickman L. Use of the internet for provision of better counselling and psychotherapy services to young people, their families and carers. Psychotherapy in Australia 2010;17(1):66-74.

16. Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley JR. Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. J Am Acad Child Adolesc Psychiatry 1999 Mar;38(3):272-279. [doi: 10.1097/00004583-199903000-00014] [Medline: 10087688]
17. Emslie GJ, Kennard BD, Mayes TL, Nightingale-Teresi J, Carmody T, Hughes CW, et al. Fluoxetine versus placebo in preventing relapse of major depression in children and adolescents. Am J Psychiatry 2008 Apr;165(4):459-467 [FREE Full text] [doi: 10.1176/appi.ajp.2007.07091453] [Medline: 18281410]
18. Burns JM. Game On: Exploring the Impact of Technologies on Young Men's Mental Health and Wellbeing Young and Well Cooperative Reserach Centre, Melbounrne. Melbourne: Young and Well Cooperative Reserach Centre; 2013. URL: http://www.youngandwellcrc.org.au/knowledge-hub/publications/game-exploring-impact-technologies-young-mens-mental-health-wellbeing/ [accessed 2014-08-31] [WebCite Cache ID 6SFahls32]
19. Australian Government. Click and connect: Young Australians' use of online social media. Canberra: Australian Government; 2009. URL: http://www.acma.gov.au/webwr/aba/about/recruitment/click_and_connect-01_qualitative_report.pdf [accessed 2014-08-31] [WebCite Cache ID 6SFar8ANj]
20. Appelbaum PS, Kopelman A. Social media's challenges for psychiatry. World Psychiatry 2014 Feb;13(1):21-23 [FREE Full text] [doi: 10.1002/wps.20085] [Medline: 24497238]
21. 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]
22. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
23. Emslie GJ, Heiligenstein JH, Hoog SL, Wagner KD, Findling RL, McCracken JT, et al. Fluoxetine treatment for prevention of relapse of depression in children and adolescents: a double-blind, placebo-controlled study. J Am Acad Child Adolesc Psychiatry 2004 Nov;43(11):1397-1405. [doi: 10.1097/01.chi.0000140453.89323.57] [Medline: 15502599]
24. 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]
25. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
26. Bradford S, Rickwood D. Adolescent's preferred modes of delivery for mental health services. Child Adolesc Ment Health 2012 Oct 11;19(1):39-45. [doi: 10.1111/camh.12002]
27. King R, Bambling M, Lloyd C, Gomurra R, Smith S, Reid W, et al. Online counselling: The motives and experiences of young people who choose the Internet instead of face to face or telephone counselling. Counselling and Psychotherapy Research 2006 Sep;6(3):169-174. [doi: 10.1080/14733140600848179]
28. Rickwood DJ, Telford NR, Parker AG, Tanti CJ, McGorry PD. headspace - Australia's innovation in youth mental health: who are the clients and why are they presenting? Med J Aust 2014 Feb 3;200(2):108-111. [Medline: 24484115]
29. 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]
30. 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]
31. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012;344:e2598 [FREE Full text] [Medline: 22517917]
32. Saulsberry A, Marko-Holguin M, Blomeke K, Hinkle C, Fogel J, Gladstone T, et al. Randomized Clinical Trial of a Primary Care Internet-based Intervention to Prevent Adolescent Depression: One-year Outcomes. J Can Acad Child Adolesc Psychiatry 2013 May;22(2):106-117 [FREE Full text] [Medline: 23667356]
33. 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]
34. Calear AL, Christensen H, Mackinnon A, Griffiths KM. Adherence to the MoodGYM program: outcomes and predictors for an adolescent school-based population. J Affect Disord 2013 May;147(1-3):338-344. [doi: 10.1016/j.jad.2012.11.036] [Medline: 23245469]
35. 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]

36. Fleming T, Dixon R, Frampton C, Merry S. A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. Behav Cogn Psychother 2012 Oct;40(5):529-541. [doi: 10.1017/S1352465811000695] [Medline: 22137185]
37. 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]
38. Van Voorhees BW, Fogel J, Pomper BE, Marko M, Reid N, Watson N, et al. Adolescent Dose and Ratings of an Internet-Based Depression Prevention Program: A Randomized Trial of Primary Care Physician Brief Advice versus a Motivational Interview. J Cogn Behav Psychother 2009;9(1):1-19 [FREE Full text] [Medline: 20694059]
39. Makarushka MM. Efficacy of an Internet-based intervention targeted to adolescents with subthreshold depression, Doctoral Thesis.: University of Oregon; 2012. URL: https://scholarsbank.uoregon.edu/xmlui/handle/1794/12091 [accessed 2014-07-17] [WebCite Cache ID 6R7muHytZ]
40. Stasiak K, Hatcher S, Frampton C, Merry SN. A pilot double blind randomized placebo controlled trial of a prototype computer-based cognitive behavioural therapy program for adolescents with symptoms of depression. Behav Cogn Psychother 2014 Jul;42(4):385-401. [doi: 10.1017/S1352465812001087] [Medline: 23253641]
41. 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]
42. 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]
43. Saulsberry A, Corden ME, Taylor-Crawford K, Crawford TJ, Johnson M, Froemel J, et al. Chicago Urban Resiliency Building (CURB): An Internet-Based Depression-Prevention Intervention for Urban African-American and Latino Adolescents. J Child Fam Stud 2012 Aug 14;22(1):150-160. [doi: 10.1007/s10826-012-9627-8]
44. Feinstein BA, Bhatia V, Hershenberg R, Davila J. Another Venue for Problematic Interpersonal Behavior: The Effects of Depressive and Anxious Symptoms on Social Networking Experiences. Journal of Social and Clinical Psychology 2012 Apr;31(4):356-382. [doi: 10.1521/jscp.2012.31.4.356]
45. Burke M. Reading, Writing, Relationships: The Impact of Social Network Sites on Relationships and Well-Being, Doctoral Thesis.: Carnegie Mellon University URL: http://reports-archive.adm.cs.cmu.edu/anon/anon/usr/ftp/hcii/CMU-HCII-11-107.pdf [accessed 2014-07-17] [WebCite Cache ID 6R7pCt8rq]
46. Collin P, Rahilly K, Richardson I, Third A. The Benefits of Social Networking Services, Orygen Youth Health Research Centre, Centre for Youth Mental Health, Univeristy of Melbourne. Melbourne: Young and Well Cooperative Research Centre; 2011 Apr. URL: http://www.fya.org.au/app/theme/default/design/assets/publications/The-Benefits-of-Social-Networking-Services.pdf [accessed 2014-09-01] [WebCite Cache ID 6SFb19gMn]
47. Davila J, Hershenberg R, Feinstein BA, Gorman K, Bhatia V, Starr LR. Frequency and Quality of Social Networking Among Young Adults: Associations With Depressive Symptoms, Rumination, and Corumination. Psychol Pop Media Cult 2012 Apr 1;1(2):72-86 [FREE Full text] [doi: 10.1037/a0027512] [Medline: 24490122]
48. Jelenchick LA, Eickhoff JC, Moreno MA. "Facebook depression?" social networking site use and depression in older adolescents. J Adolesc Health 2013 Jan;52(1):128-130. [doi: 10.1016/j.jadohealth.2012.05.008] [Medline: 23260846]
49. Moreno MA, Christakis DA, Egan KG, Jelenchick LA, Cox E, Young H, et al. A pilot evaluation of associations between displayed depression references on Facebook and self-reported depression using a clinical scale. J Behav Health Serv Res 2012 Jul;39(3):295-304 [FREE Full text] [doi: 10.1007/s11414-011-9258-7] [Medline: 21863354]
50. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety 2011 Jun;28(6):447-455 [FREE Full text] [doi: 10.1002/da.20805] [Medline: 21400639]
51. Pantic I, Damjanovic A, Todorovic J, Topalovic D, Bojovic-Jovic D, Ristic S, et al. Association between online social networking and depression in high school students: behavioral physiology viewpoint. Psychiatr Danub 2012 Mar;24(1):90-93 [FREE Full text] [Medline: 22447092]
52. Robertson L, Skegg K, Poore M, Williams S, Taylor B. An adolescent suicide cluster and the possible role of electronic communication technology. Crisis 2012;33(4):239-245. [doi: 10.1027/0227-5910/a000140] [Medline: 22562859]
53. Rogers VL, Griffin MQ, Wykle ML, Fitzpatrick JJ. Internet versus face-to-face therapy: emotional self-disclosure issues for young adults. Issues Ment Health Nurs 2009 Oct;30(10):596-602. [doi: 10.1080/01612840903003520] [Medline: 19742368]
54. Ruder TD, Hatch GM, Ampanozi G, Thali MJ, Fischer N. Suicide announcement on Facebook. Crisis 2011;32(5):280-282. [doi: 10.1027/0227-5910/a000086] [Medline: 21940257]
55. Holleran S. The early detection of depression from social networking sites.: Doctoral thesis: University of Arizona; 2010. URL: http://gradworks.umi.com/34/02/3402028.html [accessed 2014-07-17] [WebCite Cache ID 6R7qUdRe9]

56. Silenzio VM, Duberstein PR, Tang W, Lu N, Tu X, Homan CM. Connecting the invisible dots: reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks. Soc Sci Med 2009 Aug;69(3):469-474 [FREE Full text] [doi: 10.1016/j.socscimed.2009.05.029] [Medline: 19540641]
57. Takahashi Y, Uchida C, Miyaki K, Sakai M, Shimbo T, Nakayama T. Potential benefits and harms of a peer support social network service on the internet for people with depressive tendencies: qualitative content analysis and social network analysis. J Med Internet Res 2009;11(3):e29 [FREE Full text] [doi: 10.2196/jmir.1142] [Medline: 19632979]
58. Wright KB, Rosenberg J, Egbert N, Ploeger NA, Bernard DR, King S. Communication competence, social support, and depression among college students: a model of facebook and face-to-face support network influence. J Health Commun 2013;18(1):41-57. [doi: 10.1080/10810730.2012.688250] [Medline: 23030518]
59. Veretilo P, Billick SB. Psychiatric illness and facebook: a case report. Psychiatr Q 2012 Sep;83(3):385-389. [doi: 10.1007/s11126-012-9207-5] [Medline: 22274630]
60. Nimrod G. From knowledge to hope: online depression communities. nternational Journal on Disability and Human Development 2012 Jan 01;11(1):23-30. [doi: 10.1515/ijdhd.2012.009]
61. Nimrod G. Challenging the internet paradox: Online Depression Communities and Well-Being. International Journal of Internet Science 2013;8(1):30-48 [FREE Full text]
62. Nimrod G. Online depression communities: members' interests and perceived benefits. Health Commun 2013 Jul;28(5):425-434. [doi: 10.1080/10410236.2012.691068] [Medline: 22809441]
63. O'Keeffe GS, Clarke-Pearson K, Council on Communications and Media. The impact of social media on children, adolescents, and families. Pediatrics 2011 Apr;127(4):800-804 [FREE Full text] [doi: 10.1542/peds.2011-0054] [Medline: 21444588]
64. Dunlop SM, More E, Romer D. Where do youth learn about suicides on the Internet, and what influence does this have on suicidal ideation? J Child Psychol Psychiatry 2011 Oct;52(10):1073-1080. [doi: 10.1111/j.1469-7610.2011.02416.x] [Medline: 21658185]
65. Chung JE. Social interaction in online support groups: Preference for online social interaction over offline social interaction. Computers in Human Behavior 2013 Jul;29(4):1408-1414. [doi: 10.1016/j.chb.2013.01.019]
66. Barak A, Hen L, Boniel-Nissim M, Shapira N. A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. Journal of Technology in Human Services 2008 Jul 03;26(2-4):109-160. [doi: 10.1080/15228830802094429]
67. 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 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
68. 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]
69. Lewinsohn PM, Clarke GN, Hops H, Andrews J. Cognitive-behavioral treatment for depressed adolescents. Behavior Therapy 1990 Sep;21(4):385-401. [doi: 10.1016/S0005-7894(05)80353-3]
70. Christofides E, Muise A, Desmarais S. Information disclosure and control on Facebook: are they two sides of the same coin or two different processes? Cyberpsychol Behav 2009 Jun;12(3):341-345. [doi: 10.1089/cpb.2008.0226] [Medline: 19250020]
71. Bessière K, Pressman S, Kiesler S, Kraut R. Effects of internet use on health and depression: a longitudinal study. J Med Internet Res 2010;12(1):e6 [FREE Full text] [doi: 10.2196/jmir.1149] [Medline: 20228047]
72. Alvarez-Jimenez M, Bendall S, Lederman R, Wadley G, Chinnery G, Vargas S, et al. On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis. Schizophr Res 2013 Jan;143(1):143-149. [doi: 10.1016/j.schres.2012.10.009] [Medline: 23146146]
73. Alvarez-Jimenez M, Gleeson JF. Connecting the dots: twenty-first century technologies to tackle twenty-first century challenges in early intervention. Aust N Z J Psychiatry 2012 Dec;46(12):1194-1196. [doi: 10.1177/0004867412464067] [Medline: 23212140]
74. Gleeson JF, Alvarez-Jimenez M, Lederman R. Moderated online social therapy for recovery from early psychosis. Psychiatr Serv 2012 Jul;63(7):719. [doi: 10.1176/appi.ps.20120p719] [Medline: 22752039]
75. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
76. Whittaker R, Merry S, Stasiak K, McDowell H, Doherty I, Shepherd M, et al. MEMO--a mobile phone depression prevention intervention for adolescents: development process and postprogram findings on acceptability from a randomized controlled trial. J Med Internet Res 2012;14(1):e13 [FREE Full text] [doi: 10.2196/jmir.1857] [Medline: 22278284]
77. 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 [FREE Full text] [doi: 10.2196/jmir.3261] [Medline: 24874006]

78. Manicavasagar V, Horswood D, Burckhardt R, Lum A, Hadzi-Pavlovic D, Parker G. Feasibility and effectiveness of a web-based positive psychology program for youth mental health: randomized controlled trial. J Med Internet Res 2014;16(6):e140 [FREE Full text] [doi: 10.2196/jmir.3176] [Medline: 24901900]
79. Deady M, Teesson M, Kay-Lambkin F, Mills KL. Evaluating a brief, internet-based intervention for co-occurring depression and problematic alcohol use in young people: protocol for a randomized controlled trial. JMIR Res Protoc 2014;3(1):e6 [FREE Full text] [doi: 10.2196/resprot.3192] [Medline: 24583824]
80. Park M, Cha C, Cha M. Depressive moods of users portrayed in Twitter. 2012 Presented at: ACM SIGKDD Conference on Knowledge Discovery and Data Mining; Aug. 12-16, 2012; Beijing, China URL: http://wan.poly.edu/KDD2012/forms/workshop/HI-KDD12/doc/paper_16.pdf
81. Park M, McDoald DW, Cha M. Perception differences between the depressed and non-depressed users in Twitter. 2013 Presented at: Seventh International AAAI Conference on Weblogs and Social Media; July 8-11, 2013; Massachusetts, USA URL: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6114/6387
82. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting Depression via Social Media. 2013 Presented at: Seventh International AAAI Conference on Weblogs and Social Media; July 8-11, 2013; Massachusetts, USA.
83. Kennard BD, Emslie GJ, Mayes TL, Nightingale-Teresi J, Nakonezny PA, Hughes JL, et al. Cognitive-behavioral therapy to prevent relapse in pediatric responders to pharmacotherapy for major depressive disorder. J Am Acad Child Adolesc Psychiatry 2008 Dec;47(12):1395-1404 [FREE Full text] [doi: 10.1097/CHI.0b013e31818914a1] [Medline: 18978634]

Abbreviations

CBT: cognitive behavioral therapy
MDD: major depressive disorder
OSG: online support group
RCT: randomized controlled trial
SNS: social networking site

Edited by G Eysenbach; submitted 06.02.14; peer-reviewed by D Kavanagh, M Park; comments to author 16.07.14; revised version received 22.07.14; accepted 04.08.14; published 16.09.14.

Please cite as:
Rice SM, Goodall J, Hetrick SE, Parker AG, Gilbertson T, Amminger GP, Davey CG, McGorry PD, Gleeson J, Alvarez-Jimenez M
Online and Social Networking Interventions for the Treatment of Depression in Young People: A Systematic Review
J Med Internet Res 2014;16(9):e206
URL: http://www.jmir.org/2014/9/e206/
doi:10.2196/jmir.3304
PMID:25226790

Original Paper

Internet Search and Krokodil in the Russian Federation: An Infoveillance Study

Andrey Zheluk[1], BAppSc, GDip(Hist), GDip(IT), MBA; Casey Quinn[2*], BCom, Mphil, PhD; Peter Meylakhs[3*], BA(psyc), BA(soc), PhD

[1]Menzies Centre for Health Policy, University of Sydney, University of Sydney, Australia

[2]PRMA Consulting Ltd, Fleet, Hampshire, United Kingdom

[3] Laboratory for Internet Studies, National Research University, Higher School of Economics, St Petersburg, Russian Federation

[*]these authors contributed equally

Corresponding Author:
Andrey Zheluk, BAppSc, GDip(Hist), GDip(IT), MBA
Menzies Centre for Health Policy
University of Sydney
D02 Victor Coppleson Building
University of Sydney, 2006
Australia
Phone: 61 29036522
Fax: 61 293512818
Email: andreyzheluk@gmail.com

Related Articles:

This is a corrected version. See correction statement: http://www.jmir.org/2015/3/e76/

This is a corrected version. See correction statement: http://www.jmir.org/2016/1/e16/

Abstract

Background: Krokodil is an informal term for a cheap injectable illicit drug domestically prepared from codeine-containing medication (CCM). The method of krokodil preparation may produce desomorphine as well as toxic reactants that cause extensive tissue necrosis. The first confirmed report of krokodil use in Russia took place in 2004. In 2012, reports of krokodil-related injection injuries began to appear beyond Russia in Western Europe and the United States.

Objective: This exploratory study had two main objectives: (1) to determine if Internet search patterns could detect regularities in behavioral responses to Russian CCM policy at the population level, and (2) to determine if complementary data sources could explain the regularities we observed.

Methods: First, we obtained krokodil-related search pattern data for each Russia subregion (oblast) between 2011 and 2012. Second, we analyzed several complementary data sources included krokodil-related court cases, and related search terms on both Google and Yandex to evaluate the characteristics of terms accompanying krokodil-related search queries.

Results: In the 6 months preceding CCM sales restrictions, 21 of Russia's 83 oblasts had search rates higher than the national average (mean) of 16.67 searches per 100,000 population for terms associated with krokodil. In the 6 months following restrictions, mean national searches dropped to 9.65 per 100,000. Further, the number of oblasts recording a higher than average search rate dropped from 30 to 16. Second, we found krokodil-related court appearances were moderately positively correlated (Spearman correlation=.506, $P \leq .001$) with behaviors consistent with an interest in the production and use of krokodil across Russia. Finally, Google Trends and Google and Yandex related terms suggested consistent public interest in the production and use of krokodil as well as for CCM as analgesic medication during the date range covered by this study.

Conclusions: Illicit drug use data are generally regarded as difficult to obtain through traditional survey methods. Our analysis suggests it is plausible that Yandex search behavior served as a proxy for patterns of krokodil production and use during the date range we investigated. More generally, this study demonstrates the application of novel methods recently used by policy makers to both monitor illicit drug use and influence drug policy decision making.

(J Med Internet Res 2014;16(9):e212) doi:10.2196/jmir.3203

KEYWORDS

Russia; search engine; surveillance; controlled substances; designer drugs; street drugs

Introduction

Overview

Krokodil, otherwise known as desomorphine, is a cheap injectable drug easily synthesized in household kitchens from codeine-containing medication (CCM). The first confirmed report of krokodil use in Russia occurred in 2004. In 2012, reports of horrific krokodil-related injection injuries began to appear beyond Russia in Western Europe [1] and the United States [2]. We conducted this exploratory study to determine if several complementary data sources may provide insight into the relative scale and spatial patterns of behaviors consistent with an interest in the production and use of krokodil before and after the imposition of Russian federal restrictions on CCM sales in 2012.

Review of Literature on Krokodil

Current scientific literature on krokodil is limited. We reviewed international literature available through PubMed and Google Scholar. In addition, we searched the four most popular Russian online news sources [3] for the term “desomorphine” in the date range January 2009 to December 2012 using Yandex News (see Table 1). Across the four sources, we identified 929 Russian language articles associated with the term “desomorphine” in this date range, which was bounded by a period of increasing public interest in 2009, and the 6-month period following federal restrictions on CCM sales across Russia in June 2012.

Table 1. Russian news sources reviewed via Yandex News (Jan 1, 2009 to Dec 31, 2012).

Source	Website	Count (N=929)	Orientation
RIA Novosti	rian.ru	103	State-owned
Vesti.ru (Website of Russia 24 TV)	vesti.ru	38	State-owned
Komsomlskaya Pravda	kp.ru	748	Private, tabloid
Russia business consulting	rbk.ru	40	Private, business focus

The Origins of Krokodil

Current literature describing the origins of krokodil in Russia is vague. Time magazine reported the first appearance of krokodil in the Siberian and the Far East Federal Regions of Russia in the early 2000s [4]. We identified the first Russian news report of krokodil use in the Komi Republic in the western part of the Siberian Federal Region in May 2004 [5]. A police report from 2004 described the seizure of a new illicit drug never before seen in Russia called desomorphine. Conversely, a 2010 video produced by the Russian Drug Control Service (FSKN) suggested krokodil first appeared in the Komi Republic in 2002 [6] and that by 2006, 19 Russian oblasts (subregional administrative units) were affected. These affected oblasts were primarily in the Siberian, Volga, Northwestern, and Central Federal Regions. From 2006 onwards, krokodil use increased dramatically according to Russian news reports [7]. The ease of access to low-cost CCM and ease of domestic manufacture were widely reported as contributing to the spread of krokodil use [1]. Shortages of heroin during 2010 have been described as a further factor contributing to krokodil use. Several authors suggested krokodil largely displaced traditional opiates as a consequence of Afghan heroin shortages after 2010 [1,4,8,9]. A 2012 police report stated seizures of krokodil grew by 40 times from 2 kg in 2006 to 100 kg in 2011. By comparison, a mean of 2922 kg of heroin were seized each year between 2006 and 2010 in the Russian Federation [10,11].

Prevalence of Krokodil Use

Estimates of the scale of krokodil use diverged markedly. In 2011, a senior Russian addiction medicine specialist reported 5000 krokodil users were receiving treatment nationally, out of a total estimated national population of 20,000-30,000 users [12]. In June 2012, the FSKN estimated between 5000 and 7000 deaths were attributable to krokodil in Russia over the preceding 2 years [7]. Also in 2012, international researchers estimated 100,000 people were krokodil dependent in Russia, while suggesting the actual number could be higher [13]. Pharmacy sales of CCM provided a further indicator of the scale and spatial distribution of krokodil use. For example, in the Urals Region, the FSKN reported an increase in annual CCM sales from 4.2 million in 2007 to 12 million packets in 2010 [14].

There were limited data describing the spatial distribution of krokodil use before the federal restrictions on sales of CCM in June 2012. Krokodil use had been widely reported across Russia and bordering Ukrainian regions [15,16]. A December 2011 news report citing “various sources” described the results of toxicology tests conducted in several Russian oblasts [17]. According to this unreferenced news report, toxicology tests suggested 20% of people who inject drugs (PWID) in Chechnya were krokodil users, compared with 60% in Kazan and Ryazan, and 90% in Yaroslavl oblast. Overall, the publicly available data on the scale and prevalence were both limited and fragmented [18] before the introduction of restrictions in June 2012.

Researching Krokodil

A series of articles reviewing the current scientific understanding of krokodil appeared in the International Journal of Drug Policy in 2013. One of these articles by Grund et al consolidated current scientific information, including interview data from Russian

key informants [13]. The authors pointed to the paucity of scientific research into krokodil prevalence and use. Several commentary articles accompanied Grund et al's review. In one accompanying commentary piece, Heimer described the difficulties of investigating the prevalence of krokodil use without ethnographic and epidemiological data [18]. Heimer further noted the time lag between submitting research proposals and fieldwork, describing how his personal efforts to set up field interviews were thwarted by changes in krokodil use patterns resulting from federal restrictions on the sale of CCM in June 2012. In summary, conducting research into krokodil was more complex than with many established illicit drugs because of the novel nature of the substance and policy changes.

Morbidity and Mortality Associated With Krokodil in Russia

Desomorphine was originally developed as a morphine substitute. It was first synthesized in the United States in 1932, with the aim of producing a low-cost substitute with minimal side effects [19]. However, laboratory-synthesized desomorphine was regarded as an unsuccessful substitute, being shorter lasting, stronger, and more addictive than morphine. In Russia, krokodil, or what is termed desomorphine, is an illicit injectable drug domestically manufactured from codeine, iodine, phosphorus, paint thinner, and lighter fuel [13]. The resulting substance is regarded as impure, creating the potential for severe injury among PWID. It is the presence of impurities that has produced consistent reports of injuries characteristic of krokodil use. Characteristic injuries have included vascular damage, skin and soft tissue infections, necrosis and gangrene [20,21], as well as burns associated with domestic manufacture [22]. The short duration of narcotic effects, strong dependence, and chemical instability of the domestically produced drug has led to reports of binges of frequent injecting among krokodil users [23].

Frequent injecting is generally regarded as a risk factor for human immunodeficiency virus (HIV) and other injecting-related harms [24,25]. Further, PWID experience multiple comorbidities, creating a disproportionate need for health services [13,26]. However, as a consequence of punitive drug laws, Russian PWID are often particularly unwilling to seek medical assistance, exacerbating injecting-related injuries [27-29]. In summary, the characteristics of drug preparation and use, as well as the legal environment in which PWID use illicit drugs increased the morbidity and mortality associated with krokodil use in Russia before the restrictions on CCM in 2012.

Russian Policy Responses to Krokodil

The easing of restrictions on access to CCM may have increased the production and use of krokodil in Russia. During the Soviet period up to 1991, CCM was available only through pharmacies with a medical prescription [2]. Following the fall of USSR in 1991, restrictions on sales of CCM were removed. In 2004, Russian manufacturers introduced new CCMs targeted at Russian consumers. Following advocacy by the largest Russian manufacturer, Pharmstandard, these new CCMs remained accessible to consumers without medical prescriptions [30]. While a two-packet-per-person limit on CCM sales formally existed, this was routinely ignored by pharmacists [31]. From the mid-2000s, Russian government agencies explicitly linked the unrestricted access to CCM with the illicit consumption of krokodil [32]. After 2009, krokodil became increasingly recognized as a public health and policy problem. Increasing public concern was accompanied by Russian media reports of conflicts of interest. In particular, the relationship between Pharmstandard and the Russian Minister of Health became the target of media scrutiny [33,34]. Media reports suggested Pharmstandard's advocacy had prevented the imposition of restrictions on the sale of CCM. In summary, several failures in CCM regulation may have facilitated the expansion of krokodil use in Russia.

National Restrictions on Codeine-Containing Medication

At a national drug control conference in 2011, Russian President Medvedev announced restrictions on the sale of CCM without medical prescriptions [35]. Federal restrictions on CCM sales were original scheduled to start mid-2011. However, the Russian Ministry of Health effectively delayed the implementation of restrictions for 12 months. Ministry of Health officials argued that 40 million individuals using CCM for intended analgesic purposes would be disadvantaged by premature restrictions [36]. In opposing immediate restrictions, the Ministry of Health also drew on widespread public opposition to restrictions on CCM sales [37] (see Table 2). The delays in restrictions led to conflict in Russian national media between the FSKN, academics [38], and the Ministry of Health [36]. The public debate over CCM restrictions thus illuminated interagency tensions in an otherwise generally opaque Russian federal health policy landscape.

Table 2. Public opinion survey into consequences of proposed federal CCM restrictions May 2011 [34].

	How will this affect the battle against drugs in Russia?	How will this affect the needs of ordinary patients?
Generally positive	32%	11%
No effect	49%	21%
Generally negative	5%	56%
Difficult to answer	14%	12%

The Effect of Restrictions on Codeine-Containing Medication

From 2011 to June 2012, several Russian oblast governments implemented interim local restrictions on CCM sales [39,40]. These interim restrictions were directed at reducing the production and consumption of krokodil in advance of federal bans [41,42]. On June 1, 2012, a federal law restricted CCM sales across all of Russia. Russian media subsequently reported decreased sales volumes of CCM [43]. However, media reports of krokodil production and use and krokodil-related arrests continued [44,45]. These media reports suggested the federal restrictions on CCM sales had been only partially effective in curtailing krokodil production and use.

Krokodil and the Internet

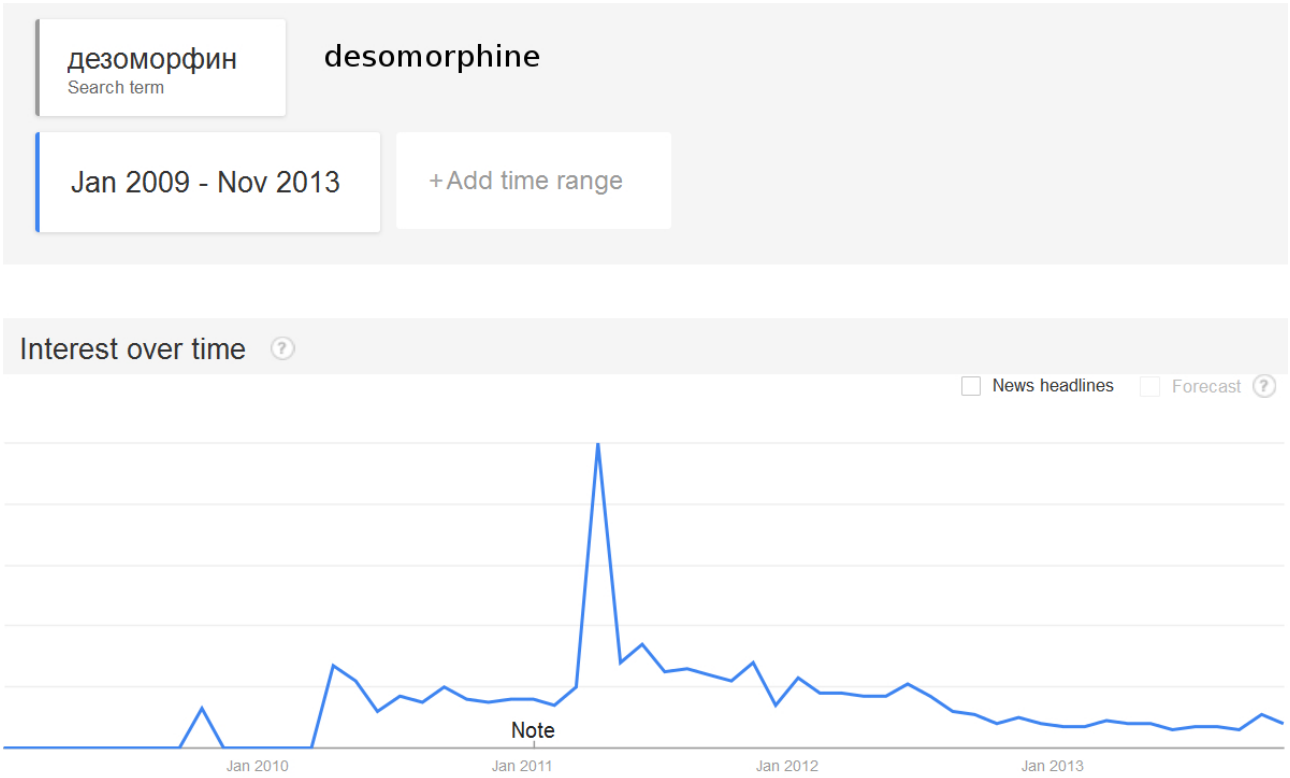
Between 2010 and 2012, Russian policy makers emphasized the negative influence of the Internet in disseminating krokodil-related information. In April 2011, the FSKN presented the results of its research into Internet search patterns for krokodil-associated terms [46]. The FSKN described a marked increase in Internet searches for methods of preparing and purchasing krokodil, from 3000 searches during 2010, to 50,000 in the first 3 months of 2011. Several days later, Russian President Dmitry Medvedev reiterated Russian government concerns about the relationship between krokodil and the Internet at a national forum dedicated to illicit drug use prevention. President Medvedev demonstrated that when the term “desomorphine” was entered into the Yandex search engine, the first results revealed information on how to prepare the drug [35]. Medvedev suggested these results proved that Internet users were most interested in producing desomorphine, rather than simply searching for general information about the drug. This widely reported demonstration by the Russian President served as the catalyst for subsequent federal restrictions on CCM.

Russian Government Statements About Krokodil

President Medvedev’s demonstration stimulated increased public interest in krokodil. The FSKN had reported steadily increasing Internet searches in the 12 months before President Medvedev’s speech. However, extensive media coverage and the highest recorded volume of searches for “desomorphine” emerged in the week following the President’s speech (see Figure 1). Further, searches remained consistently elevated after April 2011 until the Russian CCM restrictions in June 2012. Medvedev’s speech may thus have contributed to the Russian government’s unintentional amplification of popular interest in krokodil production and use [47-49]. Both before and after President Medvedev’s speech, Russian media consistently reported the harms of krokodil alongside the low price and ease of access to ingredients, and ease of drug synthesis in domestic laboratories.

Political and public concern over illicit drug use preceded President Medvedev’s speech. Public opinion polls since 2005 consistently rated illicit drug use as one of the most serious social problems in Russia [50]. Between 2010 and 2012, the Russian government progressively tightened restrictions on all illicit drug information available online. For example, in 2011, a website operated by the nongovernmental organization, Andrey Rylkov Foundation, was suspended for publishing public health information about opioid substitution therapy and harm reduction [51]. Russian government policy towards illicit drug use and drug users is regarded as punitive. Moreover, in the post-Soviet period, drug use came to represent an existential threat, implicated in national spiritual and demographic decay (eg, [52]). The intensity of public concern and political action directed at illicit drug use in Russia has been described as a moral panic [53]. Moral panics are a sociological concept describing the disproportionate public response to issues represented in media [54,55]. Over two decades, intense social and political pressures led to the demonization of drug users in Russia. After 2009, krokodil increasingly served as a focus for illicit drug-related concerns. However, President Medvedev’s April 2011 speech may have triggered a moral panic surrounding krokodil, amplified public interest, and stimulated Internet search behavior.

Figure 1. Desomorphine searches Google Trends Russia - 2009-2013.



Public Curiosity Stimulated by Media Reporting on Illicit Drugs

Exposure to stories about illicit drugs in traditional and online media has been found to increase public curiosity and the use of illicit drugs [56,57]. Dasgupta et al found mortality from prescription opiates increased significantly in the months following media reporting [58]. In addition to the news effects, online media has increased information about manufacture, sales, and use of novel illicit drugs, and undermined international drug control efforts [59-61]. Further, researchers have noted the cohort most likely to experiment with new drugs is also most likely to use the Internet [62]. The Internet is now regarded as the main medium from spreading information about novel illicit drugs [63-65]. Notably, Forsythe directly linked news reporting to increased rates of Internet search to increased use of the novel illicit drug mephedrone in the United Kingdom during 2010 [66].

We conducted this exploratory infodemiology study in order to better understand if the relative scale and spatial distribution of search behavior was consistent with an interest in the production and use of krokodil in Russia before and after the imposition of federal restrictions on CCM sales in 2012. In conducting this study, we examined “the science of distribution and determinants of information...(on) the Internet (and) a population, with the ultimate aim to inform public health and public policy” [67].

Methods

Objectives

This study had two main objectives: (1) to determine if Internet search patterns could detect regularities in behavioral responses to Russian CCM policy at population level, and (2) if complementary data sources could explain the regularities we observed.

Search Patterns and Illicit or Stigmatized Behaviors

Each Internet search is a behavioral measure of an issue’s importance to an individual [68]. If individuals are concerned or interested in an issue, they are more likely to search for information related to that issue. The relative importance of an issue can thus be inferred from the volume of search queries for a specific term or terms representing that issue. The use of aggregated Internet search patterns may be considered a method of unobserved, real time behavioral field research at population scale [69]. Researchers have used Internet search patterns to investigate illegal or stigmatized behaviors. For example, researchers analyzed search patterns to monitor US cigarette tax avoidance [70] and the use of racist terms when searching for information about Barak Obama prior to the 2012 US presidential elections [71]. Other investigators have studied global abortion patterns, finding higher search rates in those geographic regions where abortions were illegal or restricted [72]. In each of these instances, the authors suggested Internet search provided insights into behavior at a population level, beyond that available to traditional survey research.

Google and Yandex Search in Russia

Most Internet search pattern studies have used Google Trends as the data source. Google Trends has been deployed in studies of influenza [73,74], dengue [75], and HIV checks [66,76]. Globally, Google provided 84% of global Internet search queries in May 2011 [77]. In Russia, Google’s market share in Russia was only 25% in 2011 [78], whereas Yandex had a 60% share. The proportion of the Russian population using the Internet

grew from 43% in 2010 to 55% in 2012 [79]. Further, there was near universal Internet use among the age groups most likely to use illicit drugs [80]. Among both males and females aged under 35 years and living in major Russian urban centers, more than 90% were regular Internet users during 2012.

Case Study: Why Is Search-Based Krokodil Surveillance Important in Russia?

Illicit drug use data are generally regarded as difficult to obtain. Drug use estimates are imperfect even in high-income countries with adequate resources [81]. Researchers have generally divided illicit drug use estimation methods into direct and indirect approaches. Direct methods involve surveying household members about patterns of drug use. However, this method is expensive and may not produce truthful responses, particularly in countries such as Russia, where illicit drug use carries severe criminal penalties and stigma [82]. Further, household surveys may fail to reach drug-using populations such as prisoners and the homeless. Indirect methods aim to estimate the size of drug-using populations through comparing data sources. Examples of indirect methods include the multiplier method, based on estimates of the proportion of drug users receiving treatment each year [83]. Additional methods include capture-recapture and back-projection [84], using data sources such as arrest, overdose, and needle exchange data [85,86]. Drug use researchers suggest the main advantage of indirect methods is their lower cost and greater accuracy, as several measures from different data sources are generally combined to produce a single aggregated measure.

In addition, we identified two references to national drug agencies using novel methods for estimating illicit drug use prevalence. The UK government made use of Google Trends when considering restrictions on the novel substance mephedrone in 2010 [66]. Similarly, the Russian FSKN used Internet searches to develop a case for restricting CCM sales in Russia in 2011 [46]. These two instances suggest online search pattern data have been used by national governments in shaping national drug policies. Through conducting this study, we aimed to develop search patterns as an additional method to complement indirect estimates of illicit drug-using populations in Russian-speaking countries.

RQ1: Were There Regularities in the Internet Search Patterns for “Desomorphine” Across Russia in 2011-2013?

To answer this question, we examined search patterns using Yandex and Google Internet search engines. First, we determined the most appropriate search term to represent the informal term “krokodil” in Internet searches. We initially selected two search terms that we believed reflected the majority of searches for the concept central to this study [87]. These two search terms we selected were “desomorphine” (д е з о м о р ф и н in Russian) and “Krokodil” (к р о к о д и л). We referred to the Google Trends related terms (Top Searches) feature to ensure each term referred to the subject of this study. Google related terms provide information on the relative importance of searches related to the specific search term entered into Google Trends. In the case of “desomorphine”, all terms were related to the drug desomorphine, whereas the term “krokodil” revealed primarily unrelated terms (see Table 3). The term “krokodil” also means “crocodile” in Russian, leading to considerable ambiguity in search results. For example, the most popular term associated with “krokodil” (crocodile) referred to a 1960s Soviet-era children’s animated character “Gena the Crocodile”. Conversely, there is anecdotal evidence to suggest that some Russian PWID may not immediately associate the term krokodil with desomorphine [15]. Overall, based on information available from Google Trends and Russian respondents, we anticipated that the search term “desomorphine” would be more likely to reflect search patterns consistent with an interest in the production and use of krokodil.

Table 3. Google Trends related terms for desomorphine, Russian Federation.

Search terms	Russian	Value
Desomorphine-related		
Desomorphine how to prepare	дезоморфин как приготовить	100
Prepare desomorphine	приготовить дезоморфин	100
Krokodil desomorphine	крокодил дезоморфин	80
Krokodil	крокодил	80
Drug desomorphine	наркотик дезоморфин	60
Desomorphine recipe	дезоморфин рецепт	50
Krokodil drug	крокодил наркотик	30
Krokodil-related		
Gena the crocodile (Children's animation)	крокодил гена	100
Crocodile game (Children's game)	игра крокодил	50
Krokodil drug	крокодил наркотик	50
Crocodile/krokodil online	крокодил онлайн	45
Crocodile Dundee (Australian film)	крокодил данди	35
Dundee (Australian film)	данди	35
Cheburashka (Children's animation)	чебурашка	30

Publicly available Google Trends data for Russia has several limitations. First, Google did not provide complete results, returning only oblasts with the highest search volume. Google data for the term desomorphine was available for only 8 of Russia's 83 oblasts and 3 cities during the date range 2011-2013. Second, Google did not provide raw search data. This made direct comparisons between oblasts using Google data impossible. We thus used WordStat as the primary data source. Yandex made publicly available a complete raw search dataset for all Russian regions and oblasts for 6 months before and after the implementation of federal CCM restrictions in June 2012. We used Google Trends as a secondary source of aggregated search results for validation purposes.

Second, we obtained desomorphine search data for each Russia oblast from September 1, 2011, to August 31, 2013. Yandex provides 2 years of publicly available monthly search pattern data at any time. Additionally, we had 6 months previously downloaded monthly search pattern data for the term desomorphine for each Russian oblast, from February to August 2011.

Third, we converted raw search figures for the term desomorphine to population prevalences. This allowed direct comparison across regions and oblasts. We used 2010 federal Russian census data [88] for our population prevalence calculation. We then multiplied each result by 100,000 to increase ease of comprehension, and to provide a population prevalence measure.

Fourth, we analyzed search patterns before and after federal restrictions on CCM sales in June 2012. We obtained the mean search volume for 6 months before the restrictions, as well as 6 and 12 months after (ie, to August 31, 2013). We excluded June 2012 data, as we anticipated atypical search patterns in the immediate post-restriction period. Overall, we segmented the available data to examine the effects of a federal policy change on the relative scale and geographic patterns of krokodil search across the Russia.

Fifth, we obtained all available Google data for the term desomorphine from September 2011 to August 2012. Google search data for the term desomorphine was available for 8 of 83 oblasts only (see Table 4). We did not consider this sample adequate to conduct correlations. Similar limitations with regional Russian Google search results have been reported in earlier studies [69,89].

Table 4. All available Google Trends results in Russia from September 2011 to September 2013.

Region		Search volume
Oblast		
	Chelyabinsk	100
	Novosibirsk	88
	Sverdlovsk	86
	Samara	85
	Rostov	83
	Saint Petersburg city	71
	Moscow city	68
	Krasnodar	67
City		
	Yekaterinburg	100
	Nizhny Novgorod	98
	Chelyabinsk	89
	Samara	87
	Novosibirsk	87
	Rostov-on-Don	85
	Saint Petersburg	81
	Moscow	77
	Kazan	68
	Krasnodar	61

RQ2: Can Complementary Data Sources Explain the Observed Regularities for "Desomorphine" Across Russia in 2011-2013?

To answer this question, we initially reviewed the approaches used to validate search pattern data and drug population data. Search pattern studies have generally validated against an offline measure. For example, the initial search pattern studies established correlations between search patterns and epidemiological surveillance data for influenza [70,90]. Other studies focusing on issue salience, established a correlation between search patterns, traditional media, and opinion polling [91]. Each of these search studies revealed regular patterns of behavior corresponding to a valid offline measure. However, in this case, there was no analogous source of illicit drug use data available. We therefore combined several complementary data sources with a view to providing a plausible explanation for the observed regularities in Internet search patterns.

First, we obtained first court appearance data available for krokodil-related criminal charges for the 77 of 83 Russian oblast data available from the Rospravosudie website. The site is a publicly available, non-government Russian criminal justice research project displaying criminal court case data across all Russian oblasts [16]. One part of the project is dedicated to court appearance data for popular illicit recreational drugs. In addition to krokodil, Rospravosudie provides arrest data on 24 illicit drugs, including krokodil marijuana, amphetamines, JWH ("spice"), and heroin. The available krokodil data is a single per-oblast figure covering the date range 2010 to 2012. The Rospravosudie website publishers note several limitations. First, the arrest data are based on sentencing documents. Only 50% of complete sentencing data are published, and only 50% of cases appear in courts. Second, the database allows comparison of the relative popularity of various illicit drugs. It is not possible to describe the absolute prevalence of illicit drug use based on these data. Third, there is an assumption that the detection rates for illicit drug crimes were on average the same nationally. Fourth, only the first court appearances for illicit drug cases are recorded. Subsequent appeals and cassations for illicit drugs are excluded. In summary, the site authors suggest that despite these limitations, the data reflect the differences in access to illicit drugs across Russian regions.

To analyze krokodil-related court data, we first obtained arrest rates for krokodil for 2010-2012 as a single figure for each Russian oblast. We then converted the arrest rates for each oblast to a per 100,000 population measure. This allowed us to investigate the relationship between court appearances and krokodil searches. We used the mean searches for "desomorphine" per 100,000 population from November 2011 to May 2012 to represent pre-CCM restriction searches. We then conducted Spearman correlation between arrest rates and searches per 100,000 population for "desomorphine" for the 77 regions for which court data were available.

Second, we used Google Trends visual data to provide indicative national search results for popular CCMs and "desomorphine" from January 2009 to January 2013. We identified several

popular CCM available in Russia prior to the June 2012 ban [92]. The four CCM we analyzed with Google Trends were kaffetin, solpadeine, pentalgin, and codelac. Historical Yandex data were not available for this complete date range.

Third, we used Google Trends related searches to analyze several popular CCM available in Russia prior to the June 2012 restrictions. Through analyzing these related searches, we sought to obtain additional information on the characteristics of public interest in CCM and the term desomorphine before and after federal restrictions. Historical Yandex data were not available for this complete date range.

Fourth, we used Yandex keyword feature nationally to confirm that searches for the term desomorphine were associated with illicit drug use. Yandex provides a keyword function that lists word combinations associated with a specified search term. Keywords are analogous to the Google related terms (Top Searches) feature [93]. We identified 85 Yandex keyword combinations incorporating the term desomorphine. These combinations included "desomorphine prepare" and "desomorphine recipe" (see Multimedia Appendix 1).Yandex keywords are available nationally, for each Russian region and 83 oblasts, and many smaller intra-oblast cities. However, the date range is limited to the preceding 30 days only. As the search patterns from month to month are likely to be volatile within smaller geographic units, we obtained and analyzed the results for Yandex national keywords only. In order to identify the main themes present in keyword results, we hand coded the 85 keyword combinations from the national level Yandex keyword feature in the latest available date range, November 2013. Two Russian-speaking researchers then coded each keyword combination into one of three primary themes.

Results

RQ1: Were There Regularities in the Internet Search Patterns for "Desomorphine" Across Russia in 2011-13?

In the 6 months before the CCM restrictions in June 2012, 21 of Russia's 89 oblasts had Internet search rates higher than the national average (mean) of 16.67 per 100,000 (see Multimedia Appendix 2). In the 6 months immediately after restrictions, national average search rates dropped to 9.65 per 100,000. Further, the number of oblasts with a higher than average search rate dropped from 30 to 16. In the 6-month date range from March to August 2013, search rates dropped further to 8.75 per 100,000, with 11 oblasts recording higher than average search rates. However, there were a number of oblasts where searches for "desomorphine" persisted after the federal restriction on CCM sales. These included Sverdlovsk oblast (146.898 before CCM restrictions vs 81.098 post restrictions), Moscow city (31.245 vs 20.586), and Vologda oblast (34.061 vs 17.998). See Table 5. A further detailed analysis of subnational search pattern results appears in narrative form in Multimedia Appendix 3.

Table 5. Yandex search patterns for "desomorphine" in selected Russian subregional cities.

Cities	Pre-ban, 6 months, Dec 2011-June 2012	Post-ban, 6 months, July-Dec 2012	Post-ban, 6 months, Feb-Sept 2013	% change post-ban, 6 months	% change, Feb-Sept 2013
Vologda Oblast					
Vologda city	80.916	28.389	33.857	64.915	58.157
Cherepovets	50.324	28.337	25.829	43.690	48.674
Sverdlovsk Oblast					
Yekaterinburg	16.200	10.372	10.261	35.976	36.662
Kamensk-uralskiy	20.226	13.262	6.583	34.434	67.453
Pervouralsk	13.116	6.290	6.157	52.041	53.061
Rostov Oblast					
Rostov-na-donu	56.460	43.409	42.353	23.117	24.986
Kamensk-Shakhtinsky	17.850	13.002	5.950	27.160	66.667
Shakhty	2.778	0.764	1.875	72.500	32.500
Volgodonsk	26.633	11.219	8.780	57.875	67.033
Taganrog	13.324	6.080	8.861	54.369	33.495
Novocherkassk	6.420	4.247	3.852	33.846	40.000
Samara Oblast					
Samara city	25.329	16.442	24.227	35.085	4.350
Togliatti	30.131	18.736	15.147	37.817	49.731
Krasnodar Oblast					
Sochi	16.068	9.320	8.495	41.994	47.130
Novorossiysk	11.986	11.159	3.582	6.897	70.115
Krasnodar city	32.774	20.917	14.586	36.177	55.495

RQ2: Can Complementary Data Sources Explain the Observed Regularities for "Desomorphine" Across Russia in 2011-13?

To answer this question, we used several complementary sources of krokodil-related data. We found a Spearman correlation of .506 ($P \leq .001$) between searches for the term "desomorphine" and first court appearance data for krokodil related charges for all Russian oblasts. That is a moderately strong positive correlation (see Table 6).

Table 6. Correlation between searches for the term desomorphine and court appearances.

No. of subregions	Data source	Date range	Spearman correlation
83 (77 correlated)	"desomorphine" searches, Yandex WordStat	Dec 2011-May 12	—
77	desomorphine court appearances	2010-2012	.506 ($P \leq .001$)

Second, we examined national Google Trends results for four CCMs and "desomorphine". Overall, search volumes for both CCM decreased in the 6 months before the June 2012 federal restrictions, as did searches for the term "desomorphine". Public interest in CCM and the term desomorphine was roughly similar in the 6 months before the implementation of restrictions. The exception was an increase in search for the CCM pentalgin immediately before the June 2012 restrictions (see Table 7). We attribute this marked increase in interest due to public concern over access to CCM for therapeutic analgesic purposes.

Third, we examined Google Trends related terms for CCMs and desomorphine. We found related terms for CCMs consistent with therapeutic and analgesic uses (see Table 7). By contrast, related terms for desomorphine were consistent with an interest in the production and use of krokodil. The Google related terms data did not record all search results. We attribute this to the Google "threshold effect" described in earlier analysis of drug policy [69]. That is, below an unspecified threshold value, Google records a nil value.

Table 7. Google related search terms for "desomorphine" from 2009-2013.

Date range	Pentalgin (пенталгин)	Value	Codelac (коделак)	Value	Desomorphine (дезоморфин)	Value
2009-2013	Pentalgin N	100	Codelac broncho	100	Desomorphine how to prepare	100
	Pentalgin instructions	65	Codelac phyto	75	Desomorphine krokodil	75
	Pentalgin composition	50	Codelac instructions	75	Krokodil	70
	Pentalgin price	45	Codelac price	65		
	Nurofen	40	Codelac syrup	60		
Dec 2011-May 2012	Pentalgin N	100	Codelac phyto	100	How to prepare desomorphine	100
	Pentalgin instructions	85	Codelac instructions	100	Krokodil	55
July 2012-Aug 2013	Insufficient search volume	Nil	Insufficient search volume	Nil	Insufficient search volume	Nil

Fourth, we used the Yandex keyword feature to analyze the word combinations used with the search term desomorphine. We found combinations associated with krokodil preparation and use accounted for 46.613% of searches, images, and general information for 24.175%, and ambiguous terms for 29.212% (see Multimedia Appendix 1 and Table 8). We used Cohen's Kappa to assess intercoder reliability on all 85 search combinations across three categories using two coders (kappa=.772).

The preparation and use category included all terms associated with drug preparation and use. Images and entertainment included visual material and terms unlikely to be associated with drug use and preparation (eg, "YouTube desomorphine", "junkies desomorphine"). In summary, we found the combination of search patterns with complementary methods useful for identifying behaviors consistent with an interest in the production and use of krokodil.

Table 8. Main themes identified in WordStat keyword combined word searches for "desomorphine" (excluding non-combined word searches for the single term "desomorphine".

Code		n (N=6338)	Percentage
1	Preparation & Use	2952	46.613
2	Images & information	1531	24.175
3	Ambiguous	1850	29.212

Discussion

Principal Findings

We found federal CCM restrictions in June 2012 coincided with changes in the relative scale and spatial patterns of Internet search behaviors consistent with an interest in the production and use of krokodil. These changes in Internet search appeared consistent with behaviors that may be anticipated in the production and use of krokodil in response to changed access to CCM.

We observed marked reductions in searches for the term desomorphine following CCM sales restrictions in June 2012. By comparison with the 6 months preceding federal restrictions, searches dropped by 42.095% nationally (see Multimedia Appendix 2). However, the patterns of decreased search for "desomorphine" varied considerably. This suggests that CCM restrictions changed but did not extinguish behaviors consistent with an interest in the production and use of krokodil.

Third, we found the Google data available were inadequate for statistical analysis. Insufficient Google Trends data were available to conduct statistical analysis to identify oblasts where krokodil use may be prevalent. Google Trends data were available for only 8 of 83 regions (see Table 4).

We identified several complementary data sources that provided a plausible explanation for the observed regularities in Internet search data. First, we found a moderately strong positive correlation (Spearman correlation=.506) between the geographic distribution of court appearances for krokodil-related charges, and Internet searches for the term desomorphine. This result should be treated with some caution. Court appearance data were available for 78 of 83 statistical regions. This may have affected the strength of correlations. More significantly, international researchers generally regard Russian policing as predatory and beyond the rule of law [94,95]. A 2010 study of PWID found reports of evidence planted by police, extortion money, arbitrary arrests, and violence to be widespread. Further, Russian court processes are regarded by researchers and Russian public opinion as likely to produce outcomes favoring police and prosecutors [96-98]. Both policing and judicial practices may be expected to distort court appearance data. It is possible that in the absence of these law enforcement practices the strength of correlation may differ. Conversely, the uncertainty surrounding Russian law enforcement data is consistent with descriptions of other data sources used for indirect drug population estimation by international researchers.

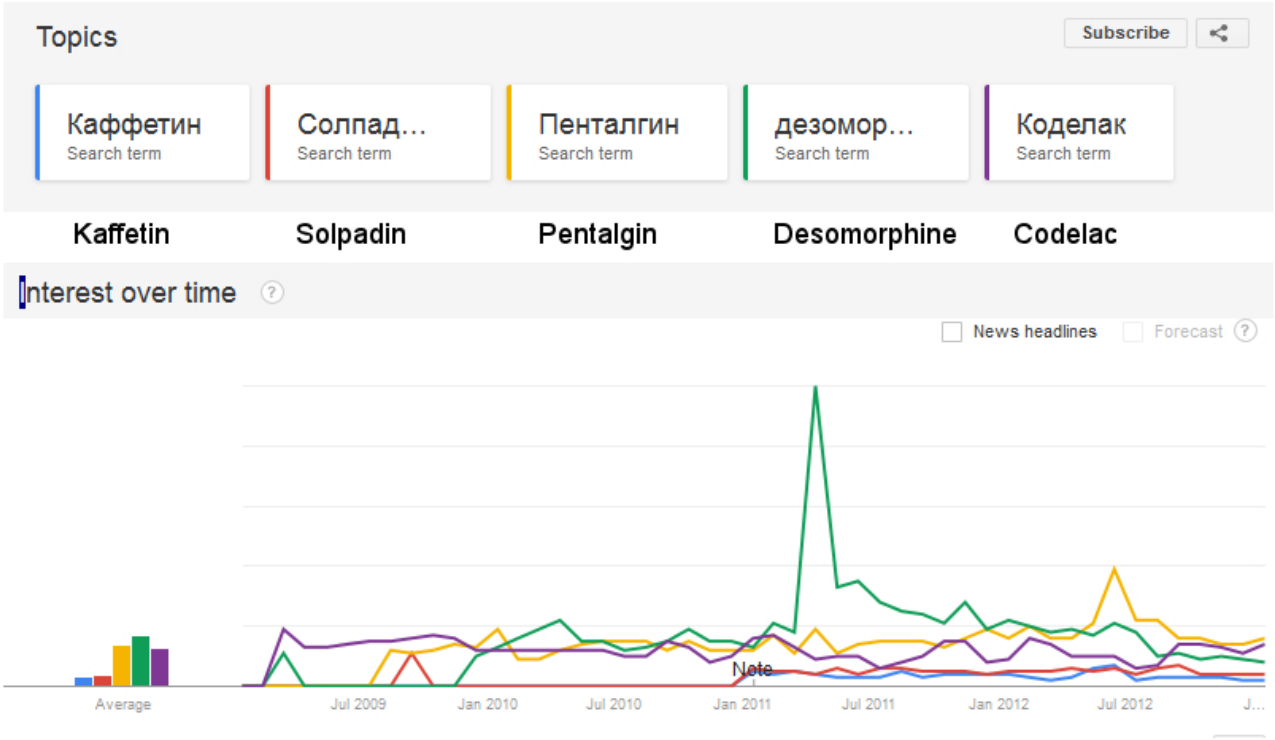
Second, the available Google Trends data suggested public interest in CCM and the term desomorphine was roughly similar in the 6 months prior to federal restrictions. However, the

searches for CCM and desomorphine-related terms were not identical. The interest in CCM and in desomorphine manifested as different national level search patterns over the date range. While we had insufficient Google data to conduct correlations, this difference is evident on visual inspection (see Figure 2). Similarly, Google Trends related terms results suggest different themes for "desomorphine" and CCM (see Table 7). Whereas searches for the term desomorphine were associated with illicit drug use themes, the CCM search themes were primarily associated with therapeutic use of drugs. Finally, there is anecdotal evidence from Russian informants that CCM was also widely consumed orally rather than injected prior to the introduction of federal restrictions (P. Meylakhs, Personal Communication, 2014). The restrictions on CCM sales thus also affected oral use of CCM as a recreational drug.

Yandex keyword analysis revealed a consistent pattern of behavior that was consistent with an interest in the production and use of krokodil. Yandex keyword data also revealed a strong popular interest in visual images of desomorphine use (see Table 8). Overall, 23.59% of searches were coded as pertaining to the images and general information theme. These graphic images were actively employed in Russian government campaigns aimed at reducing krokodil use. However, the wide distribution of images may also have created a popular demand for viewing necrotic injuries as voyeurism or entertainment. While the available Yandex keyword data were outside of the date range of the study, they reveal persistent interest in behavior consistent with an interest in the production and use of krokodil in Russia.

In summary, we used complementary data sources in order to investigate behaviors consistent with an interest in the production and use of krokodil. Our analysis suggests that these combined complementary sources, including online news sources, provided a useful addition to the conventional approaches used to analyze krokodil use in Russia. Further, our analysis also suggests it is plausible that Yandex search behavior served as a proxy for krokodil production and use in the date range 2011-2012.

Figure 2. Google Trends results for CCM and desomorphine search 2009 - 2013.



Further Research and Limitations

Our research suggests that further research into the use of search patterns for investigating illicit drug use prevalence is warranted. First, search patterns offer researchers and non-government groups an additional source of indirect data with which to track the prevalence of traditional and emerging synthetic drugs at low cost and in near real time. We identified two references to national drug agencies in United Kingdom and Russia using Internet search methods to research patterns of illicit drug use [46,66]. Search pattern methods offer non-government groups similar capabilities. Further, the methods described in this paper are directly applicable to the study of other illicit drugs and in Russian-speaking countries where Yandex is widely used alongside Google.

Second, the krokodil case represents an example of a broader class of illicit drug policy events. International and Russian researchers have partially attributed increased use of krokodil to decreased heroin supply after 2009. Similarly, in 2012, government policy blocked easy access to CCM. In each case, existing networks of PWID were disrupted, and patterns of illicit drug use rapidly changed [18]. Combining Internet search and offline qualitative methods would extend understanding of rapid shifts in illicit drug markets and potentially improve public health responses to emerging synthetic drugs. This qualitative research may include existing drug use prevalence, strength of

the heroin market, and Internet access among PWID in Russia. In particular, this research would assist researchers with discerning rapid shifts in drug use patterns in response to policy changes and other external shocks to existing illicit drug markets. The relationship of searches for "desomorphine" to other illicit drugs appears in Multimedia Appendices 4 and 5.

Third, media censorship is increasing in contemporary Russia. However, our analysis of online information relied on measures of unobserved population level demand for online information only. By contrast, censorship may be expected to influence the supply of illicit drug-related information. Russian government actions restricting the supply of illicit drug information are well documented in international literature (eg, [51,99]). Future research should investigate the relationship between searches for information and the censored supply of information.

Finally, search methods do not estimate actual drug user population size. However, our research suggests search methods can complement existing drug-using population estimation methods. For example, the Yandex keywords feature potentially provides a novel data source with which to track monthly shifts in keywords for illicit drug–related terms. Keywords measures provide a low-cost method for identifying spatial shifts in the relative scale of public interest in terms that are consistent with an interest in the production and use of novel and emerging illicit drugs in an increasingly complex environment where the opportunities for conventional field work and surveys in Russia for international researchers are decreasing.

Conclusions

Illicit drug use data are generally regarded as difficult to obtain through traditional survey methods. We used complementary methods to explain observed regularities in patterns of Internet search behavior before and after the imposition of Russian federal restrictions on CCM sales in 2012. Our analysis suggests it is plausible that Yandex search behavior served as a proxy for patterns of krokodil production and use during the date range we investigated. More generally, this study demonstrates the application of novel methods recently used by policy makers to both monitor illicit drug use and influence drug policy decision making.

Acknowledgments

The work of Peter Meylakhs was funded by the Basic Research Program of the National Research University Higher School of Economics, Russia. The authors would also like to thank Ms Svetlana Chernova in Ukraine for her assistance with data collection and coding, Associate Professor James Gillespie at the Menzies Centre for Health Policy, University of Sydney for his advice and support, and Ms Anya Sarang from the Andrey Rylkov Foundation in Moscow for facilitating contacts within Russia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Yandex Wordstat Keywords: combinations of words searched for with "desomorphine", November 2013.

[PDF File (Adobe PDF File), 53KB - jmir_v16i9e212_app1.pdf]

Multimedia Appendix 2

Per-capita mean searches for "desomorphine" in all Russian federal regions & subregions (**= region or subregion above national mean).

[PDF File (Adobe PDF File), 55KB - jmir_v16i9e212_app2.pdf]

Multimedia Appendix 3

Analysis of regional Internet search patterns.

[PDF File (Adobe PDF File), 34KB - jmir_v16i9e212_app3.pdf]

Multimedia Appendix 4

Illicit drug search popularity of related terms in Google Trends 2009-2014.

[JPG File, 282KB - jmir_v16i9e212_app4.jpg]

Multimedia Appendix 5

Google searches for popular illicit drugs in the Russian Federation 2009-2014.

[JPG File, 313KB - jmir_v16i9e212_app5.jpg]

References

1. Gahr M, Freudenmann RW, Hiemke C, Gunst IM, Connemann BJ, Schönfeldt-Lecuona C. Desomorphine goes "crocodile". J Addict Dis 2012;31(4):407-412. [doi: 10.1080/10550887.2012.735570] [Medline: 23244560]
2. Thekkemuriyi DV, John SG, Pillai U. 'Krokodil'--a designer drug from across the Atlantic, with serious consequences. Am J Med 2014 Mar;127(3):e1-e2. [doi: 10.1016/j.amjmed.2013.09.030] [Medline: 24274952]
3. TNS Web Index December 2011. Moscow: TNS URL: http://tns-global.ru/media/content/B7525726-B5E1-4C12-BE25-4C543F42F3EE/!Web%20Index%20Report%20201112.zip [accessed 2013-12-13] [WebCite Cache ID 6LqGJZ8Yq]
4. Shuster S. Time. 2011 Jun 20. The curse of the crocodile: Russia's deadly designer drug URL: http://content.time.com/time/world/article/0,8599,2078355,00.html [accessed 2013-12-13] [WebCite Cache ID 6LqGQbbYy]
5. Regnum.ru. A new illicit drug, never before seen in Russia has been identified in Komi URL: http://www.regnum.ru/news/accidents/260274.html [accessed 2013-12-13] [WebCite Cache ID 6Sa3CGzpW]
6. FSKN. A film about desomorphine addicts. 2010. URL: http://www.youtube.com/watch?feature=player_embedded&v=vKHBF_hXits
7. RIAN. Experts: Russia is ready for the ban on unrestricted sales of codeine containing medications URL: http://ria.ru/beznarko_law/20120601/662265774.html [accessed 2013-12-10] [WebCite Cache ID 6LkgyetiC]
8. ARF. Results of the Internet survey on the heroin situation in Russia. Drug scene in Russia 2010 URL: http://en.rylkov-fond.org/blog/drug-policy-and-russia/drug-policy-in-russia/drugscene-in-russia/ [accessed 2013-12-14] [WebCite Cache ID 6LqGvz5YC]
9. 1tv. More than 20 Russian oblasts will introduce restrictions on sales of codeine containing medications in advance of a federal ban. URL: http://www.1tv.ru/news/social/190424 [accessed 2013-12-13] [WebCite Cache ID 6LqH085hO]
10. World Drug Report. Vienna: UNODC; 2010. URL: http://www.unodc.org/unodc/en/data-and-analysis/WDR-2010.html [accessed 2014-05-31] [WebCite Cache ID 6PxSXtot1]
11. World Drug Report. Vienna: UNODC; 2012. URL: http://www.unodc.org/unodc/en/data-and-analysis/WDR-2012.html [accessed 2014-05-30] [WebCite Cache ID 6PxSV7gI9]
12. RIAN. FSKN states that desomorphine seizures in the Urals have increased by 5 times since 2008 URL: http://ria.ru/beznarko_danger/20111226/527639443.html [accessed 2013-12-10] [WebCite Cache ID 6Lkfx21Mh]
13. Grund JP, Latypov A, Harris M. Breaking worse: the emergence of krokodil and excessive injuries among people who inject drugs in Eurasia. Int J Drug Policy 2013 Jul;24(4):265-274. [doi: 10.1016/j.drugpo.2013.04.007] [Medline: 23726898]
14. FSKN. Analysis of desomoprhine use prevalence in the Urals Federal Region including volumes of codeine containing medications. 2011. URL: http://www.ufo.fskn.gov.ru/?articleID=826 [accessed 2013-12-13] [WebCite Cache ID 6LqEjdnYF]
15. Booth RE. 'Krokodil' and other home-produced drugs for injection: a perspective from Ukraine. Int J Drug Policy 2013 Jul;24(4):277-278. [doi: 10.1016/j.drugpo.2013.05.009] [Medline: 23764188]
16. Rospravosudie. Angel dust on the map of Russia. URL: https://rospravosudie.com/society/narko [accessed 2013-12-13] [WebCite Cache ID 6LqH8XjHZ]
17. Borzunova K. RIAN. The victims of Krokodil URL: http://ria.ru/columns/20111220/521480375.html [accessed 2013-12-14] [WebCite Cache ID 6LqGpbYVM]
18. Heimer R. Patterns of new drug emergence: a comment in light of 'krokodil'. Int J Drug Policy 2013 Jul;24(4):275-277 [FREE Full text] [doi: 10.1016/j.drugpo.2013.06.003] [Medline: 23827763]
19. Eddy NB. Synthetic substances with morphine-like effect: clinical experience; potency, side-effects, addiction liability. Bull World Health Organ 1957;17(4-5):569-863 [FREE Full text] [Medline: 13511135]
20. Priymak A. Desomorphine, drug for the poor, kills all of its victims. 2011 Jun 23. URL: http://english.pravda.ru/hotspots/crimes/23-06-2011/118296-desomorphine-0/ [accessed 2013-12-14] [WebCite Cache ID 6LqHBtEH2]
21. Skowronek R, Celiński R, Chowaniec C. "Crocodile"--new dangerous designer drug of abuse from the East. Clin Toxicol (Phila) 2012 Apr;50(4):269. [doi: 10.3109/15563650.2012.660574] [Medline: 22385107]
22. RIAN. 2011 Mar 31. Explosion in a house in Rostov on Don while drugs being prepared URL: http://ria.ru/antidrugs/20110331/359573834.html [accessed 2013-12-10] [WebCite Cache ID 6LkfNdYs7]
23. Walker S. The Independent. 2011 Jun 22. Krokodil: the drug that eats junkies URL: http://www.independent.co.uk/news/world/europe/krokodil-the-drug-that-eats-junkies-2300787.html [accessed 2013-12-14] [WebCite Cache ID 6LqppPhCP]
24. WHO. Management of common health problems of drug users. New Delhi: WHO SEARO; 2009. URL: http://www.who.int/hiv/topics/idu/drug_dependence/hiv_primary_care_guidelines_searo.pdf [accessed 2013-12-16] [WebCite Cache ID 6SP269edK]
25. Degenhardt L, Mathers B, Vickerman P, Rhodes T, Latkin C, Hickman M. Prevention of HIV infection for people who inject drugs: why individual, structural, and combination approaches are needed. Lancet 2010 Jul 24;376(9737):285-301. [doi: 10.1016/S0140-6736(10)60742-8] [Medline: 20650522]
26. Islam MM, Topp L, Day CA, Dawson A, Conigrave KM. The accessibility, acceptability, health impact and cost implications of primary healthcare outlets that target injecting drug users: a narrative synthesis of literature. Int J Drug Policy 2012 Mar;23(2):94-102. [doi: 10.1016/j.drugpo.2011.08.005] [Medline: 21996165]

27. Niccolai LM, Verevochkin SV, Toussova OV, White E, Barbour R, Kozlov AP, et al. Estimates of HIV incidence among drug users in St. Petersburg, Russia: continued growth of a rapidly expanding epidemic. Eur J Public Health 2011 Oct;21(5):613-619 [FREE Full text] [doi: 10.1093/eurpub/ckq115] [Medline: 20798184]
28. Amirkhanian YA, Kelly JA, Kuznetsova AV, DiFranceisco WJ, Musatov VB, Pirogov DG. People with HIV in HAART-era Russia: transmission risk behavior prevalence, antiretroviral medication-taking, and psychosocial distress. AIDS Behav 2011 May;15(4):767-777 [FREE Full text] [doi: 10.1007/s10461-010-9793-x] [Medline: 20803063]
29. Wolfe D, Carrieri M, Shepard D. Treatment and care for injecting drug users with HIV infection: a review of barriers and ways forward. Lancet 2010 Jul 31;376(9738):355-366. [doi: 10.1016/S0140-6736(10)60832-X] [Medline: 20650513]
30. Skobeyda Y. Komsomolskaya Pravda. 2011 Jun 03. Pharmacies freely sell the medications from which illicit drugs are made URL: http://www.kp.md/daily/25696/899094/?cp=0 [accessed 2013-12-14] [WebCite Cache ID 6LqpgwYv6]
31. kp.ru. Drug police check pharmacies. 2011 Jun 20. URL: http://www.kp.ru/online/news/916804 [accessed 2013-12-16] [WebCite Cache ID 6LtPjFYxJ]
32. izvmor.ru. Drug police to investigate sales of medications in pharmacies. 2007 Jan 19. URL: http://old.izvmor.ru/index.php?q=article_981.html [accessed 2014-05-30] [WebCite Cache ID 6PxSGHLFo]
33. Yermakov O. Ministerial kickbacks or a corrupt tumour in ministerial families?. 2009 Oct 05. URL: http://www.rospres.com/finance/5240 [accessed 2013-12-14] [WebCite Cache ID 6Lqq2ZW7v]
34. Sterligov I. slon.ru. 2010 May 05. The end of codeine innovations URL: http://slon.ru/business/konec_kodeinovyh_innovaciy-369152.xhtml [accessed 2013-12-14] [WebCite Cache ID 6Lqq4ecpc]
35. RIAN. 2011 Apr 18. Medvedev calls for closure of websites promoting illicit drugs URL: http://old.rian.ru/society/20110418/365642510.html [accessed 2013-12-10] [WebCite Cache ID 6Lkg40dVc]
36. RIAN. 2011 May 16. Bryun: codeine-containing medications should not be demonised in Russia URL: http://ria.ru/beznarko_danger/20110516/377488902.html [accessed 2013-12-10] [WebCite Cache ID 6LkfVqOI2]
37. WCIOM. Russian Public Opinion Research Centre. 2011 Jun 07. Codeine containing medication by prescription: for and against URL: http://wciom.ru/index.php?id=459&uid=111674
38. RIAN. 2011 May 19. FSKN calls for ban on sale of codeine containing medication without a medical prescription URL: http://ria.ru/society/20110519/376834039.html [accessed 2013-12-10] [WebCite Cache ID 6LkfrWjwJ]
39. Malinina E. kp.ru. 2011 Dec 12. Astrakhan becomes 10th Russian region to restrict sales of codeine-containing medications URL: http://www.kp.ru/online/news/1039373 [accessed 2013-12-14] [WebCite Cache ID 6Lqqclf4r]
40. RIAN. 2012 Jan 01. Codeine containing medication will be sold strictly with a medical prescription in Orenburg oblast URL: http://volga.ria.ru/society/20120101/82047502.html [accessed 2013-12-17] [WebCite Cache ID 6LvnI8r33]
41. RIAN. 2011 Sep 28. Drug den numbers fall in regions where codeine sales restricted URL: http://ria.ru/beznarko_danger/20110928/445240710.html [accessed 2013-12-10] [WebCite Cache ID 6Lkfkgwn8]
42. aktualno.ru. 2102 Jul 23. Sales of codeine containing medications in Chelyabinsk oblast have fallen by 5 times in 6 months URL: http://aktualno.ru/view2/16359 [accessed 2013-12-14] [WebCite Cache ID 6Lqun8OVQ]
43. Chernaya E. RIAN. 2012 Jul 24. Sales of codeine containing medications in Tomsk oblast have fallen - FSKN URL: http://ria.ru/beznarko_news/20120724/710654402.html [accessed 2013-12-14] [WebCite Cache ID 6LqudVudY]
44. Kravtsova K. prm.ru. 2013 Nov 11. Perm locals organised a network of desomorphine laboratories URL: http://prm.ru/accident/292096/ [accessed 2013-12-14] [WebCite Cache ID 6Lquj0aNx]
45. Yanyshev A. dvnovosti.ru. 2013 Nov 13. Thieves, addicts and hooligans - a day in the life of police on the beat URL: http://www.dvnovosti.ru/khab/2013/11/13/13491/ [accessed 2013-12-13] [WebCite Cache ID 6LqHORc3s]
46. pravda.ru. 2011 Apr 13. Internet users have become more interested in preparation of drugs URL: http://www.pravda.ru/news/society/13-04-2011/1073604-drugs-0/ [accessed 2013-12-14] [WebCite Cache ID 6LquTVPvc]
47. Zvonkaya A. aifudm.net. 2013 Dec 14. Krokodil eats addicts in Udmutriya oblast URL: http://aifudm.net/news/news31191.html [accessed 2013-12-14] [WebCite Cache ID 6LquOzMly]
48. Bogdanov V. rg.ru. 2011 Aug 05. Pain without a prescription URL: http://www.rg.ru/2011/08/05/kodein.html [accessed 2013-12-14] [WebCite Cache ID 6LquL9Zz0]
49. medvestnik.ru. 2012 Mar 11. FSKN: In 20 regions sales of codeine containing medications have decreased up to thirtyfold URL: http://www.medvestnik.ru/news/fskn_v_20_regionah_rf_prodazhi_kodeinsoderzhaschih_lekarstv_snizilis_do_30_raz/ [accessed 2014-09-07] [WebCite Cache ID 6SPJ58GY7]
50. Levada Center. 3013 Sep 03. Which social problems most disturb you, and you consider to be the most important? URL: http://www.levada.ru/03-09-2013/problemy-i-trevogi-rossiyan [accessed 2013-12-14] [WebCite Cache ID 6LquCahRY]
51. openinform.ru. 2012 May 16. Court case against website closure order by Federal Drug Control Agency begins in Moscow URL: http://openinform.ru/news/unfreedom/16.05.2012/26828/ [accessed 2013-12-14] [WebCite Cache ID 6Lqu8P3dN]
52. Kolesnikov V. narkotiki.ru. 2009. Weapons of genocide destroying the genofund of the nation URL: http://www.narkotiki.ru/oinfo_6657.html [accessed 2013-12-14] [WebCite Cache ID 6Lqu5B07i]
53. Meylakhs P. The Discourse of the Press and the Press of Discourse: The Construction of the Drug Problem in Russian Media. In Critical Readings: Moral Panics and the Media. In: Critcher C, editor. Critical readings: moral panics and the media. Maidenhead, Berkshire: Open University Press; 2006.
54. Cohen S. Folk Devils and Moral Panics. New York, NY: Routledge; 2011.

55. Jenkins P. Synthetic panics: the symbolic politics of designer drugs. New York, NY: New York University Press; 1999.
56. Lancaster AR. What Does Curiosity Really Do to the Cat? A Look at How Message Exposure Can Lead to Illicit Drug Trial Among College Students. Mass Communication and Society 2004 Mar;7(1):77-95. [doi: 10.1207/s15327825mcs0701_6]
57. Lancaster K, Hughes CE, Spicer B, Matthew-Simmons F, Dillon P. Illicit drugs and the media: models of media effects for use in drug policy research. Drug Alcohol Rev 2011 Jul;30(4):397-402. [doi: 10.1111/j.1465-3362.2010.00239.x] [Medline: 21355898]
58. Dasgupta N, Mandl KD, Brownstein JS. Breaking the news or fueling the epidemic? Temporal association between news media report volume and opioid-related mortality. PLoS One 2009;4(11):e7758 [FREE Full text] [doi: 10.1371/journal.pone.0007758] [Medline: 19924221]
59. Davey Z, Corazza O, Schifano F, Deluca P. Mass-information: mephedrone, myths, and the new generation of legal highs. Drugs and Alcohol Today 2010;10(3):24-28. [doi: 10.5042/daat.2010.0467]
60. Gallagher CT, Assi S, Stair JL, Fergus S, Corazza O, Corkery JM, et al. 5,6-Methylenedioxy-2-aminoindane: from laboratory curiosity to 'legal high'. Hum Psychopharmacol 2012 Mar;27(2):106-112. [doi: 10.1002/hup.1255] [Medline: 22389075]
61. Report of the International Narcotics Control Board. Vienna: INCB; 2011. URL: http://www.incb.org/documents/Publications/AnnualReports/AR2010/AR_2010_English.pdf [accessed 2013-12-15] [WebCite Cache ID 6LtPpLydz]
62. EC. Eurobarometer: young people and drugs among 15–24 year olds. Analytical report. Dublin: European Commission; 2008. URL: http://ec.europa.eu/public_opinion/flash/fl_233_en.pdf [accessed 2013-12-14] [WebCite Cache ID 6LqtvF546]
63. Hillebrand J, Olszewski D, Sedefov R. Legal highs on the Internet. Subst Use Misuse 2010 Feb;45(3):330-340. [doi: 10.3109/10826080903443628] [Medline: 20141450]
64. Corazza O, Schifano F, Farre M, Deluca P, Davey Z, Torrens M, et al. Designer drugs on the internet: a phenomenon out-of-control? the emergence of hallucinogenic drug Bromo-Dragonfly. Curr Clin Pharmacol 2011 May;6(2):125-129. [Medline: 21592070]
65. Corazza O, Assi S, Schifano F. From "Special K" to "Special M": the evolution of the recreational use of ketamine and methoxetamine. CNS Neurosci Ther 2013 Jun;19(6):454-460. [doi: 10.1111/cns.12063] [Medline: 23421859]
66. Forsyth AJ. Virtually a drug scare: mephedrone and the impact of the internet on drug news transmission. Int J Drug Policy 2012 May;23(3):198-209. [doi: 10.1016/j.drugpo.2011.12.003] [Medline: 22342603]
67. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
68. Granka L. Inferring the Public Agenda from Implicit Query Data. Boston, MA: ACM SIGIR Conference; 2009. URL: http://ceur-ws.org/Vol-512/paper07.pdf [accessed 2013-11-05] [WebCite Cache ID 6KtuAa2j2]
69. Zheluk A, Quinn C, Hercz D, Gillespie JA. Internet search patterns of human immunodeficiency virus and the digital divide in the Russian Federation: infoveillance study. J Med Internet Res 2013;15(11):e256 [FREE Full text] [doi: 10.2196/jmir.2936] [Medline: 24220250]
70. Ayers JW, Ribisl K, Brownstein JS. Using search query surveillance to monitor tax avoidance and smoking cessation following the United States' 2009 "SCHIP" cigarette tax increase. PLoS One 2011;6(3):e16777 [FREE Full text] [doi: 10.1371/journal.pone.0016777] [Medline: 21436883]
71. Stephens-Davidowitz S. SSRN Journal. 2012. The Effects of Racial Animus on a Black Presidential Candidate: Using Google Search Data to Find What Surveys Miss URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2050673 [accessed 2014-09-05] [WebCite Cache ID 6SN6RafdW]
72. Reis BY, Brownstein JS. Measuring the impact of health policies using Internet search patterns: the case of abortion. BMC Public Health 2010;10:514 [FREE Full text] [doi: 10.1186/1471-2458-10-514] [Medline: 20738850]
73. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]
74. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
75. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. PLoS Negl Trop Dis 2011 May;5(5):e1206 [FREE Full text] [doi: 10.1371/journal.pntd.0001206] [Medline: 21647308]
76. Jena AB, Karaca-Mandic P, Weaver L, Seabury SA. Predicting new diagnoses of HIV infection using internet search engine data. Clin Infect Dis 2013 May;56(9):1352-1353 [FREE Full text] [doi: 10.1093/cid/cit022] [Medline: 23334812]
77. Netmarketshare.com. 2011. Search Engine Market Share URL: http://www.netmarketshare.com/report.aspx?qprid=4&qptimeframe=M&qpsp=148
78. Crisp T. ComScore Voices. 2011 Jun 27. Yandex: From Russia with Love URL: http://www.comscore.com/Insights/Blog/Yandex-From-Russia-with-Love [accessed 2014-09-07] [WebCite Cache ID 6SPKDfhN6]
79. WCIOM. Russian Public Opinion Research Center. 2102 Feb 13. Russians networking: rating of social media URL: http://wciom.ru/index.php?id=459&uid=112476

80. TNS Web Index: The audience as a whole. Moscow: TNS; 2013. URL: http://www.tns-global.ru/media/content/B7525726-B5E1-4C12-BE25-4C543F42F3EE/!Web%20Index%20Report%20201212_EN.zip [accessed 2013-12-14] [WebCite Cache ID 6LqtZvkK1]
81. Degenhardt L, Bucello C, Calabria B, Nelson P, Roberts A, Hall W, GBD illicit drug use writing group, et al. What data are available on the extent of illicit drug use and dependence globally? Results of four systematic reviews. Drug Alcohol Depend 2011 Sep 1;117(2-3):85-101. [doi: 10.1016/j.drugalcdep.2010.11.032] [Medline: 21377813]
82. Reuter P, Trautmann F. Utrecht Eur Comm. 2009. A report on global illicit drugs markets 1998-2007 URL: http://ec.europa.eu/justice/anti-drugs/files/report-drug-markets-short_en.pdf [accessed 2013-12-15] [WebCite Cache ID 6Ls76Pcsa]
83. Hickman M, Taylor C. Sci Drug Abuse Epidemiol. 2002. Estimating the prevalence of problematic drug use: a review of methods and their application URL: https://www.unodc.org/pdf/bulletin/bulletin_2002_01_01_Art2.pdf [accessed 2013-12-15] [WebCite Cache ID 6Ls7DWeS3]
84. Law MG, Lynskey M, Ross J, Hall W. Back-projection estimates of the number of dependent heroin users in Australia. Addiction 2001 Mar;96(3):433-443. [doi: 10.1080/0965214002005400] [Medline: 11255583]
85. Hall WD, Ross JE, Lynskey MT, Law MG, Degenhardt LJ. How many dependent heroin users are there in Australia? Med J Aust 2000 Nov 20;173(10):528-531. [Medline: 11194736]
86. Degenhardt L, Hall W. Extent of illicit drug use and dependence, and their contribution to the global burden of disease. Lancet 2012 Jan 7;379(9810):55-70. [doi: 10.1016/S0140-6736(11)61138-0] [Medline: 22225671]
87. Yang AC, Tsai SJ, Huang NE, Peng CK. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004-2009. J Affect Disord 2011 Jul;132(1-2):179-184. [doi: 10.1016/j.jad.2011.01.019] [Medline: 21371755]
88. gks.ru. All-Russian Census. Moscow, Russia: Russ Fed State Stat Serv; 2012. URL: http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm [accessed 2013-12-15] [WebCite Cache ID 6LtPmtOoj]
89. Zheluk A, Gillespie JA, Quinn C. Searching for truth: internet search patterns as a method of investigating online responses to a Russian illicit drug policy debate. J Med Internet Res 2012;14(6):e165 [FREE Full text] [doi: 10.2196/jmir.2270] [Medline: 23238600]
90. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. Clin Infect Dis 2008 Dec 1;47(11):1443-1448 [FREE Full text] [doi: 10.1086/593098] [Medline: 18954267]
91. Mellon J. Where and When Can We Use Google Trends to Measure Issue Salience? APSC 2013 Mar 28;46(02):280-290. [doi: 10.1017/S1049096513000279]
92. Aleksandrov M. Vechernyaya Moskva. 2011 Feb 23. The codeine itch URL: http://moskva.bezformata.ru/listnews/kodeinovij-zud/522253/ [accessed 2014-05-30] [WebCite Cache ID 6PxSJpeja]
93. Google.com. 2011. What are top searches? URL: https://support.google.com/trends/answer/4355000?hl=en&rd=1 [accessed 2014-09-05] [WebCite Cache ID 6SN6tefBk]
94. Rhodes T, Platt L, Sarang A, Vlasov A, Mikhailova L, Monaghan G. Street policing, injecting drug use and harm reduction in a Russian city: a qualitative study of police perspectives. J Urban Health 2006 Sep;83(5):911-925 [FREE Full text] [doi: 10.1007/s11524-006-9085-y] [Medline: 16855880]
95. Gerber TP, Mendelson SE. Public Experiences of Police Violence and Corruption in Contemporary Russia: A Case of Predatory Policing? Law Soc Rev 2008 Mar;42(1):1-44. [doi: 10.1111/j.1540-5893.2008.00333.x]
96. Kurkchiyan M. The illegitimacy of law in post-Soviet societies. Law and Informal Practices: The Post-Communist Experience 2003:25-46. [doi: 10.1093/acprof:oso/9780199259366.003.0002]
97. Hendley K. Cardozo Journal of International and Comparative Law. 2006. Assessing the rule of law in Russia URL: https://media.law.wisc.edu/m/ytdxn/assessing_rol.pdf [accessed 2014-09-06] [WebCite Cache ID 6SMX74U7p]
98. ARF. rylkov-fond.org. 2012. Additional information to the Report to the ICESCR on implementation by the RF of article 12 of the International Covenant on ESCR as it relates to access of people who inject drugs to drug treatment and HIV prevention, care and treatment programs URL: http://en.rylkov-fond.org/wp-content/uploads/2012/01/AndreyRylkovFoundation_RussianFederation_CESCR46-eng.pdf [accessed 2013-12-15] [WebCite Cache ID 6LtPkunYP]
99. Parfitt T. Vladimir Mendelevich: fighting for drug substitution treatment. The Lancet 2006 Jul;368(9532):279. [doi: 10.1016/S0140-6736(06)69060-0]

Abbreviations

CCM: codeine-containing medication
HIV: human immunodeficiency virus
PWID: people who inject drugs
FSKN: Russian Federal Drug Control Service

Edited by G Eysenbach; submitted 21.12.13; peer-reviewed by R Heimer, G Woody; comments to author 28.04.14; revised version received 03.06.14; accepted 22.08.14; published 18.09.14.

Please cite as:
Zheluk A, Quinn C, Meylakhs P
Internet Search and Krokodil in the Russian Federation: An Infoveillance Study
J Med Internet Res 2014;16(9):e212
URL: http://www.jmir.org/2014/9/e212/
doi:10.2196/jmir.3203
PMID:25236385

Original Paper

An Internet-Based Epidemiological Investigation of the Outbreak of H7N9 Avian Influenza A in China Since Early 2013

Chen Mao[1,2*], PhD; Xin-Yin Wu[1*], PhD; Xiao-Hong Fu[1], MPH, PhD; Meng-Yang Di[1], MD, PhD; Yuan-Yuan Yu[1], MD, PhD; Jin-Qiu Yuan[1], MPH, PhD; Zu-Yao Yang[1,2], PhD; Jin-Ling Tang[1,2], PhD

[1]School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, Hong Kong, China (Hong Kong)

[2]The Shenzhen Municipal Key Laboratory for Health Risk Analysis, Shenzhen Research Institute of The Chinese University of Hong Kong, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Jin-Ling Tang, PhD
School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong
4/F, School of Public Health and Primary Care, Prince of Wales Hospital, Shatin, New Territories
Hong Kong,
China (Hong Kong)
Phone: 852 22528779
Fax: 852 26453098
Email: jltang@cuhk.edu.hk

Abstract

Background: In early 2013, a new type of avian influenza, H7N9, emerged in China. It quickly became an issue of great public concern and a widely discussed topic on the Internet. A considerable volume of relevant information was made publicly available on the Internet through various sources.

Objective: This study aimed to describe the outbreak of H7N9 in China based on data openly available on the Internet and to validate our investigation by comparing our findings with a well-conducted conventional field epidemiologic study.

Methods: We searched publicly accessible Internet data on the H7N9 outbreak primarily from government and major mass media websites in China up to February 10, 2014. Two researchers independently extracted, compared, and confirmed the information of each confirmed H7N9 case using a self-designed data extraction form. We summarized the epidemiological and clinical characteristics of confirmed H7N9 cases and compared them with those from the field study.

Results: According to our data updated until February 10, 2014, 334 confirmed H7N9 cases were identified. The median age was 58 years and 67.0% (219/327) were males. Cases were reported in 15 regions in China. Five family clusters were found. Of the 16.8% (56/334) of the cases with relevant data, 69.6% (39/56) reported a history of exposure to animals. Of the 1751 persons with a close contact with a confirmed case, 0.6% (11/1751) of them developed respiratory symptoms during the 7-day surveillance period. In the 97.9% (327/334) of the cases with relevant data, 21.7% (71/327) died, 20.8% (68/327) were discharged from a hospital, and 57.5% (188/327) were of uncertain status. We compared our findings before February 10, 2014 and those before December 1, 2013 with those from the conventional field study, which had the latter cutoff date of ours in data collection. Our study showed most epidemiological and clinical characteristics were similar to those in the field study, except for case fatality (71/327, 21.7% for our data before February 10; 45/138, 32.6% for our data before December 1; 47/139, 33.8% for the field study), time from illness onset to first medical care (4 days, 3 days, and 1 day), and time from illness onset to death (16.5 days, 17 days, and 21 days).

Conclusions: Findings from our Internet-based investigation were similar to those from the conventional field study in most epidemiological and clinical aspects of the outbreak. Importantly, publicly available Internet data are open to any interested researchers and can thus greatly facilitate the investigation and control of such outbreaks. With improved efforts for Internet data provision, Internet-based investigation has a great potential to become a quick, economical, novel approach to investigating sudden issues of great public concern that involve a relatively small number of cases like this H7N9 outbreak.

(J Med Internet Res 2014;16(9):e221) doi:10.2196/jmir.3763

KEYWORDS

influenza A virus, H7N9 subtype; Internet; big data; disease outbreaks; epidemiology

Introduction

The modern world is becoming increasingly connected by the Internet, which comprises a large system of various private, public, business, academic, and government networks that facilitate rapid and open data exchange among billions of people worldwide [1]. The Internet has removed a major constraint to accessing and sharing data, information, and knowledge. Unlike traditional media, it provides an open platform that allows various people to report, confirm, and correct details regarding issues of public concern. For example, "Internet mass hunting," which literally means to uncover the true identity of a particular person or matter through the coordinated efforts of all "netizens," is an approach to using the Internet for conducting investigations [2]. In China, this approach has been increasingly employed by the public to elucidate issues of public concern, such as cases of corruption [3].

The Internet also provides novel opportunities for health research. Eysenbach developed a method for analyzing Google queries to track cases of influenza-type illnesses in a given population [4,5]. That method was employed to accurately predict weekly influenza activity in each region in the United States with a reporting lag of approximately 1 day [6]. Health surveys and clinical trials have also been conducted through the Internet [7-10]. Other examples of Internet-based research in the health sciences include Infovigil, HealthMap, MedISys, BioCaster, and EpiSPIDER [11-15]. Compared with traditional field research methods, such as face-to-face interviews and paper-based questionnaires, Internet-based methods involve considerably fewer resources, including money, time, and human resources. The Internet is particularly useful for obtaining information and influencing behavioral responses during public crises [16], such as during outbreaks of new infectious diseases.

In early 2013, a new type of avian influenza, H7N9, emerged in China, becoming a matter of strong public concern. After the first case was reported, H7N9 rapidly became a widely discussed topic on the Internet. A considerable volume of relevant information was made publicly available on the Internet through various channels, including news reports, discussion forums, personal blogs, and reports from hospitals and government authorities. In this study, we aimed to describe the outbreak of H7N9 in China based on data available on the Internet by February 10, 2014, and to explore the methods, feasibility, validity, advantages, and limitations of employing the Internet-based approach to investigating public health issues by comparing our findings with those presented in a well-conducted conventional field epidemiologic study on the H7N9 outbreak [17].

Methods

Study Structure and Definitions

We collected publicly available Internet data related to the H7N9 outbreak from reliable websites. The field epidemiologic study by Li and colleagues [17] was based on the data they reported to the National Center for Disease Control and Prevention of the country (China CDC) and was employed as a reference for validating our study. In our study, we first summarized the epidemiological and clinical characteristics of H7N9 cases based on the data from the Internet updated until February 10, 2014. To validate our study, we compared our findings with those in Li's study [17]. Li's results were compared with ours based on data with the same cutoff date of investigation (December 1, 2013) and with our results updated to February 10, 2014.

In Li's study, suspected and confirmed cases of H7N9 virus infection were defined according to the definition of H5N1 cases recommended by the World Health Organization (WHO) in 2006. Suspected cases were identified through China's surveillance system for cases of pneumonia of unexplained origin. If laboratory tests, such as real-time reverse transcriptase–polymerase chain reaction assay, viral isolation, and serological testing, indicated the presence of the H7N9 virus, the case was considered confirmed. Once a suspected case was identified, initial field investigations were conducted and respiratory specimens were obtained by local CDCs. Information on each confirmed case was collected through field investigation until December 1, 2013.

In our study, confirmed H7N9 cases were those that claimed to be newly identified H7N9 cases reported on either government or nongovernment websites, and verified either by laboratory tests or nationally or provincially organized specialists.

Sources of Internet Data

Various H7N9-related data were reported through numerous Internet-based resources. The data in this study were obtained from only 2 website categories: government and major media websites. These categories are detailed as follows, in the order of their trustworthiness:

Websites of government organizations providing H7N9-related information in China were the primary source of data. The 5 most representative government organizations that provided our needed information were (1) municipal health bureaus; (2) national, provincial, and prefectural CDCs; (3) the National Health and Family Planning Commission; (4) the Ministry of Agriculture; and (5) the WHO.

Social media websites were used as supplementary sources of data. Websites that set up a special column or discussion forum for the outbreak of H7N9 Avian Influenza A were given a higher priority than others. In fact, most of our supplementary data were obtained from the 2 most popular public websites in the country: Sina and Sohu. Both had a special section for H7N9 on their websites [18,19]. They provided numerous information, including live reports of new H7N9 cases, individual data for confirmed H7N9 cases, simple summary data of the outbreak, prevention and treatment of H7N9 Avian Influenza A, and comments from netizens.

Governmental data were considered most trustworthy and contributed most of the data in this study. Information was

sought from social media websites when some data were missing from government websites.

We included 2 cases reported on Health Authority websites in Hong Kong, but did not search Taiwan-based websites and thus the few cases in Taiwan were not included in this study. No case was reported in Macau.

Case Identification and Data Collection

To avoid relying on or repeating the summary or aggregate results of official reports and other epidemiological studies by those who had access to official individual data on H7N9 cases, we searched and extracted only those data on individual cases from websites with free public access. These data were mainly included in daily reports published on relevant government authority websites. Websites were selected according to usefulness, accessibility, and credibility. Usefulness was determined by the comprehensiveness and up-to-datedness of the information. To obtain data for each confirmed H7N9 case, we searched daily reports published on the CDC and other health bureau websites at the national, provincial, and prefectural level from March 31, 2013 (the date of the first reported case) to February 10, 2014. Complementary searches of social media websites, predominantly Sina and Sohu, were also performed to supplement data obtained from the government websites. Cases were identified based on the date of illness onset, family name, demographic data (gender, age, and region), and exposure history. Duplicate cases were defined as those with an identical date of illness onset, family name, gender, age, and region. When cases were reported on multiple websites, we used the data from government websites of the highest level from national to provincial to prefectural.

Data Extraction and Quality Control

Two researchers (MYD and YYY) extracted the data for each confirmed case independently by using a self-designed data extraction form. Discrepancies were resolved by double-checking the websites and discussions if deemed necessary. Extracted data included (1) demographic data, such as age, sex, rural/urban residency, and occupation; (2) epidemiological data, such as potential exposure history to the H7N9 virus, the number of close contacts, secondary cases, familial aggregation cases, and confirming method for diagnosis; and (3) clinical data, such as hospitalization, intensive care unit (ICU) admission, development of acute respiratory distress syndrome (ARDS), death, and dates for these clinical outcomes (see Multimedia Appendix 1).

Statistical Analysis

We employed descriptive and analytic statistics to summarize the epidemiological and clinical characteristics of confirmed H7N9 cases. SPSS version 18.0 (SPSS Inc, Chicago, IL, USA) was used to perform all the statistical analyses.

Results

Epidemiologic Characteristics of Confirmed Cases by February 10, 2014

From February 17, 2013, to February 10, 2014, a total of 334 cases of H7N9 infection were identified. Cases occurred in the following 15 regions (Figure 1): (1) Zhejiang (130 cases); (2) Guangdong (62 cases); (3) Shanghai (42 cases); (4) Jiangsu (42 cases); (5) Fujian (20 cases); (6) Hunan (10 cases); (7i) Jiangxi (6 cases); (8) Anhui (6 cases); (9) Henan (4 cases); (10) Beijing (3 cases); (11) Guangxi (3 cases); (12) Shandong (2 cases); (13) Hong Kong (2 cases); (14) Hebei (1 case); and (15) Guizhou (1 case).

The majority of cases were obtained from official sources, with 89.5% (299/334) from government websites, 4.5% (15/334) from nongovernment websites, and 6.0% (20/334) from both government and nongovernment websites. Table 1 summarizes the epidemiologic characteristics of the confirmed cases. The median age was 58 years, with 33.9% (111/327) older than 65 years and 1.8% (6/327) younger than 5 years. More men (219/327, 67.0%) than women were reported as confirmed cases of H7N9. Most of the cases (186/278, 66.9%) lived in urban areas. Seven cases (7/201, 3.5%) were poultry workers. Five cases had comorbidities including hypertension, heart disease, diabetes mellitus, and chronic bronchitis. Among the 91 confirmed H7N9 cases with information on the method of diagnosis, 78.0% (71/91) were confirmed using nucleic acid detection and the remaining 20 cases were confirmed by a group of infectious disease specialists.

Information on recent exposure to animals was available for 56 of 334 confirmed cases of H7N9, 39 of which were patients with a history of exposure to animals. Cases with an animal exposure are detailed as follows: (1) 3 patients reported 1 instance of exposure to poultry; (2) 10 patients reported multiple instances of exposure to poultry; (3) 3 patients were exposed to pigeons, quails, and pet birds; and (4) the remaining 23 cases did not report the type of animals to which patients were exposed.

Table 1. Epidemiologic characteristics of patients with confirmed H7N9 infection in China from early 2013 to February 10, 2014.

Characteristics	Confirmed H7N9 cases[a]		
	Internet-based data (up to Dec 1, 2013)	Li's report (up to Dec 1, 2013)	Most updated Internet-based data (up to Feb 10, 2014)
Total number of cases	138	139	334
Age (years)			
Number of cases	132	139	327
Median (IQR[b])	61 (48-73)	61 (46-73)	58 (41-69)
<5 years, n/N (%)	3/132 (2.3)	4/139 (2.9)	6/327 (1.8)
≥ 65 years, n/N (%)	55/132 (41.7)	58/139 (41.7)	111/327 (33.9)
Male sex, n/N (%)	93/132 (70.5)	98/139(70.5)	219/327 (67.0)
Area of residence, n/N (%)			
Urban	83/110 (75.5)	101/139 (72.6)	186/278 (66.9)
Rural	27/110 (24.5)	38/139 (27.3)	92/278 (33.1)
Poultry worker, n/N (%)	6/99 (6.1)	9/139 (6.6)	7/201 (3.5)
Presence of underlying medical conditions, n/N (%)	3/3 (100.0)	79/108 (73.1)	5/5 (100.0)
Exposure to a symptomatic case within 2 week before illness onset, n/N (%)	5/NE[c] (NE)	5/120 (4.2)	7[d]/NE (NE)
Exposure to poultry, n/N (%)	12/26 (46.0)	107/131 (81.7)	39/56 (69.6)
Chickens	7/12 (58.3)	88/107 (82.2)	12/39 (30.8)
Ducks	3/12 (25.0)	24/107 (22.4)	4/39 (10.3)
Pigeons	1/12 (8.3)	13/107 (12.1)	1/39 (2.7)
Quails	1/12 (8.3)	2/107 (1.9)	1/39 (2.7)
Pet birds	1/12 (8.3)	3/107 (2.8)	1/39 (2.7)
Direct contact with poultry	9/12 (75.0)	63/107 (58.9)	17/39 (43.6)
Nucleic acid detection confirming, n/N (%)	58/78 (74.4)	89/139 (64.0)	71/91 (78.0)

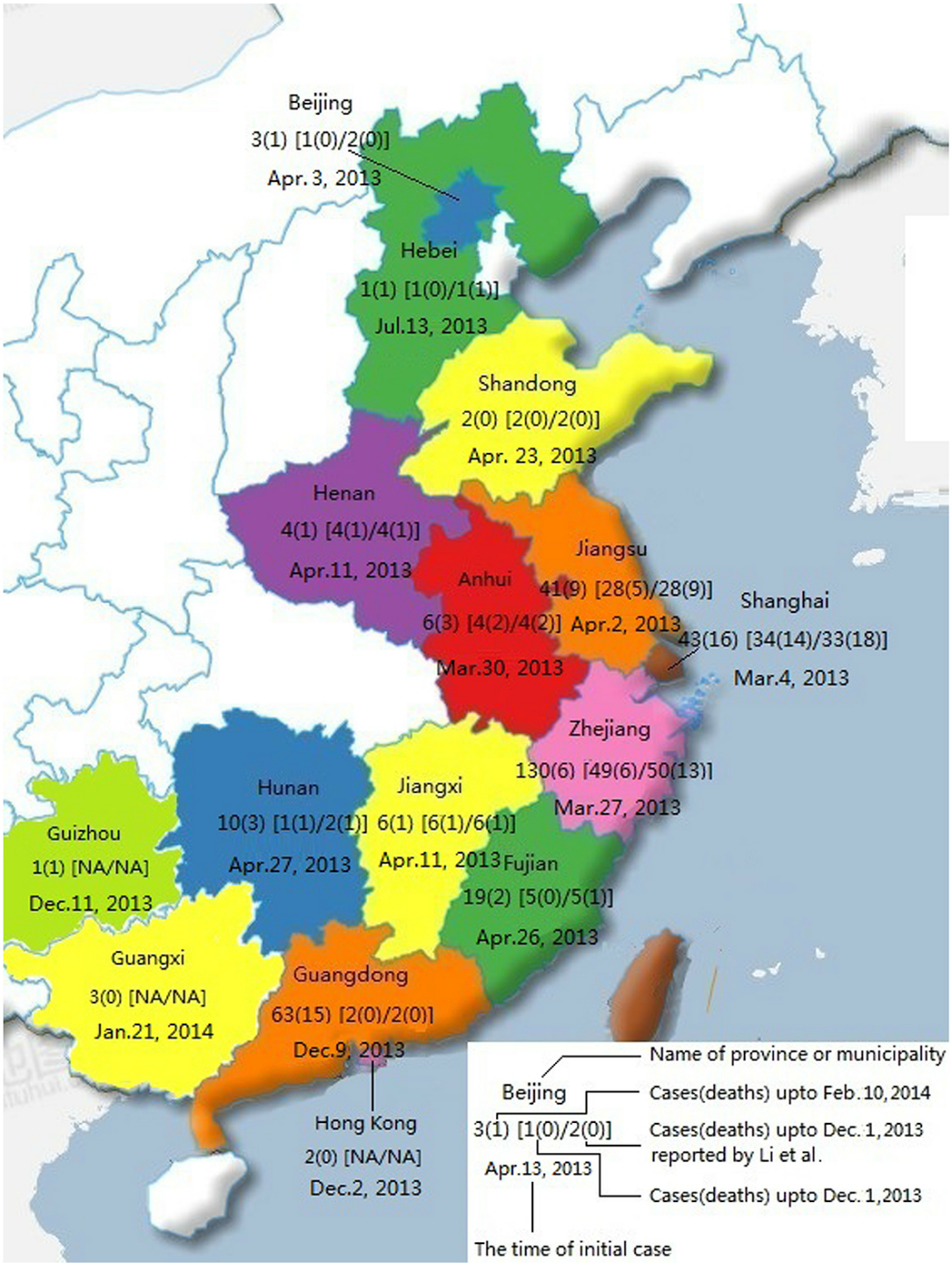
[a]Unless stated otherwise, data in the table provided are the number of patients with a certain characteristics, the total number of patients having data on that characteristic and the corresponding percentage.

[b]IQR: interquartile range

[c]NE: not estimable

[d]Two cases in Zhejiang were not confirmed by February 10, 2014.

Figure 1. Geographic distribution of patients with confirmed H7N9 infection in China from early 2013 to February 10, 2014.



Family Clusters

Among the 334 identified H7N9 cases, 5 family clusters were reported in 3 provinces: (1) a father (confirmed) and his 2 sons (1 confirmed and the other unconfirmed) in Shanghai; (2) a married couple, both of whom were confirmed cases in Shanghai; (3) a father and son, both of whom were confirmed in Shandong; (4) a father and son-in-law, both of whom were confirmed in Zhejiang; and (5) 3 confirmed cases in the same family in Zhejiang. In the last family cluster, the man was sick first and his daughter and wife became sick after giving care to him. All 3 family members also reported an identical history of exposure to poultry. The investigation for these cases was still going on as of February 10, 2014.

Clinical Characteristics

Table 2 summarizes the clinical characteristics and medical care timelines for confirmed cases. A total of 209 (209/223, 93.7%) cases were hospitalized, 33 (33/61, 54.1%) were admitted to ICUs, and 16 (16/23, 69.7%) developed ARDS. A total of 68 patients (68/327, 20.8%) were discharged from the hospital, 71 (71/327, 21.7%) died, and 188 (188/327, 57.5%) were of uncertain status. The median time from illness onset until their first medical visit, time from illness onset to hospitalization, time from illness onset to ICU admission, time from illness

onset to the development of ARDS, and time from illness onset to death was 4.0 days (71 cases), 5.0 days (55 cases), and 6.5 days (6 cases), 6.0 days (3 cases), and 16.5 days (34 cases), respectively.

We conducted a subgroup analysis to compare the case fatality rate of confirmed H7N9 cases before and after the first reported cases on March 31, 2013. Of the 29 cases reported before March 31, 2013, 44.8% (13/29) died, and the case fatality was significantly higher (χ^2_1= 15.1, P<.001) than that in patients reported after March 31, 2013 (58/305, 19.0%).

Table 2. Clinical characteristics and medical care timelines for patients with confirmed H7N9 infection in China from early 2013 to February 10, 2014.

Variables	Confirmed H7N9 cases[a]		
	Internet-based data (up to Dec 1, 2013)	Li's report (up to Dec 1, 2013)	Most updated Internet-based data (up to Feb 10, 2014)
Clinical outcome, n/N (%)			
Hospitalization	118/132 (89.4)	137/139 (98.7)	209/223 (93.7)
ICU admission	30/56 (53.6)	65/103 (63.1)	33/61 (54.1)
ARDS[b]	14/18 (77.8)	48/83 (57.8)	16/23 (69.7)
Death	45[d]/138 (32.6)	47/139 (33.8)	71/327 (21.7)
Time from onset to first medical care-days			
Number of cases	53	137	71
Median (IQR[c])	3.0 (1.0-5.0)	1.0 (0-3.0)	4.0 (1.0-6.0)
Time from onset to hospitalization-days			
Number of cases	41	137	55
Median (IQR)	5.0 (3.0-6.5)	4.0 (3.0-6.0)	5.0 (3.0-6.0)
Time from onset to ICU admission-days			
Number of cases	4	103	6
Median (IQR)	7.5 (5.5-10.3)	7.0 (5.0-9.0)	6.5 (4.8-8.8)
Time from onset to development of ARDS-days			
Number of cases	3	83	3
Median (IQR)	6.0 (NE)[e]	7.0 (5.0-9.0)	6.0 (NE)[e]
Time from onset to death-days			
Number of cases	25	47	34
Median (IQR)	17.0 (10.5-22.5)	21.0 (12.5-36.0)	16.5 (9.5-22.0)

[a]Unless stated otherwise, data in the table provided are the number of patients with a certain characteristic, the total number of patients having data on that characteristic, and the corresponding percentage.

[b]ARDS: respiratory distress syndrome

[c]IQR: interquartile range

[d]This number from aggregate data on the website of National Health and Family Planning Commission. If we use individual data, only 33 deaths were found, with a case fatality rate of 23.9%.

[e]NE: not estimable

Close Contact

By February 10, 2014, 81 confirmed H7N9 cases provided data on people who had close contact with them, with a total number of 1751 people involved. Among them, 859 occurred in Jiangsu, 310 in Shanghai, 113 in Zhejiang, 72 in Anhui, 45 in Henan, 42 in Fujian, 54 in Jiangxi, 9 in Shandong, 37 in Beijing, and 210 in Guangdong. Only 11 of these people developed respiratory symptoms during the 7-day surveillance period. Family clusters were not included in these analyses.

Comparison With Li's Study

By February 10, 2014, the number of cases collected from the Internet was twice more than that reported in Li's study. By using the same cutoff date of December 1, 2013, we obtained 138 cases, which was only 1 fewer than the number of cases (139 cases) included in Li's study.

Based on the data by February 10, 2014, and by December 1, 2013, our findings agreed with Li's report on most epidemiological characteristics, including age, sex ratio, area

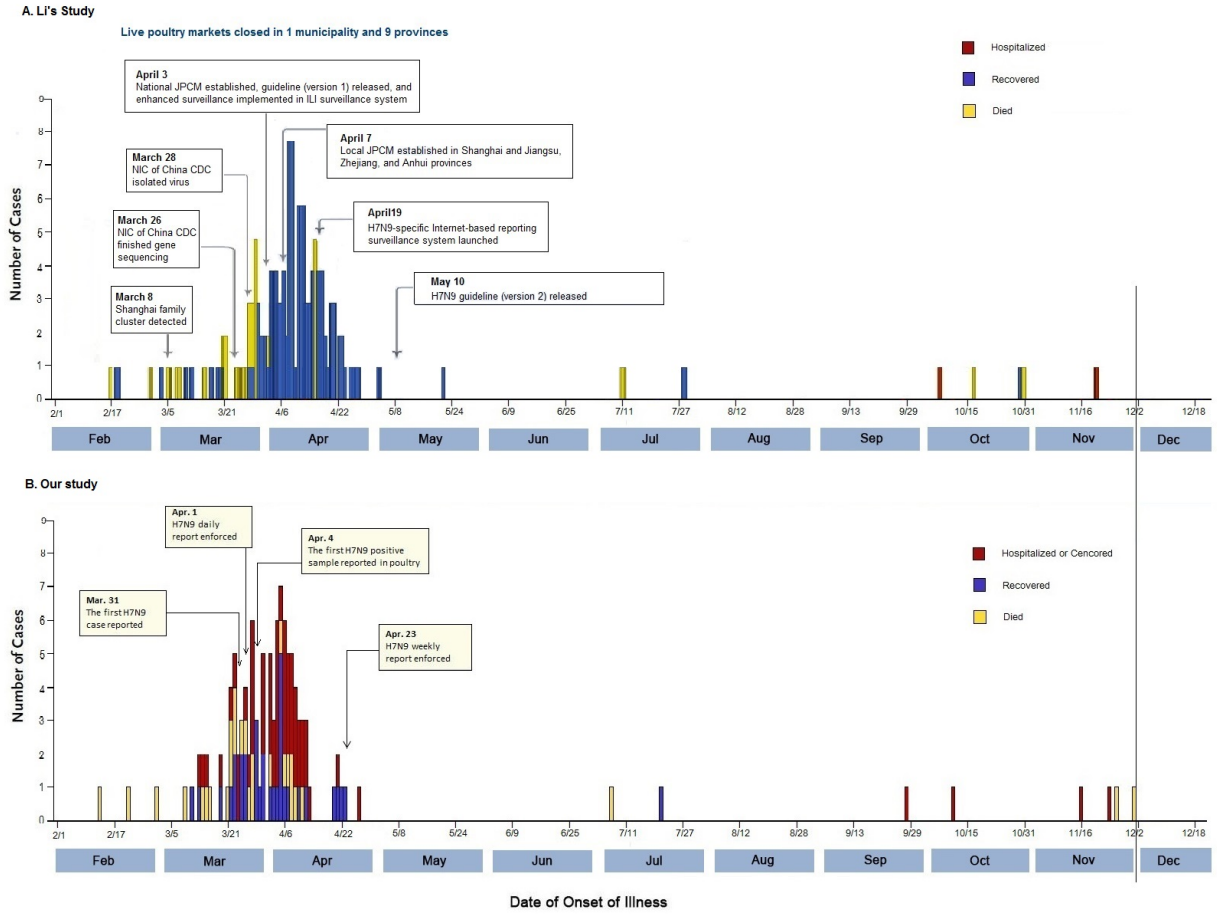
of residence, and the most commonly used method for confirming cases of H7N9, although we had less detailed information on patient exposure history and clinical characteristics. Despite the problem of missing data, our results were similar to those reported in Li's study in most clinical characteristics, such as median time from illness onset to hospitalization, ICU admission, and development of ARDS.

Based on the most updated data up to February 10, 2014, we obtained a case fatality (71/327, 21.7%) lower than that reported in Li's study (47/139, 33.8%). If using individual data from the Internet with the same cutoff date of December 1, 2013, the case fatality in our study was 23.9% (33/138), which was still lower than 33.8% (47/139) reported in Li's study. However, the website of the National Health and Family Planning Commission reported a total of 45 deaths by October 31, 2013. Based on this publically available aggregate data, a case fatality was very similar to that in Li's study. Our results on both time from illness onset to first medical care and that from illness onset to death differed from those reported by Li for both cutoff dates used. We were unable to collect detailed information on the identified family clusters (Tables 1 and 2).

Based on the Internet data by December 1, 2013, the epidemic curve shown in Figure 2 is very similar to that in Li's report, although in our study only 102 cases had the information on the date of onset of illness and there were more missing cases at end of the peak outbreak period. The first case was reported on February 19, 2013, and subsequent cases were reported sporadically between February 19 and March 27, 2013. An epidemic peak was observed between March 28 and April 18, 2013. Since then, the number of cases started to decline, with only 13 new cases reported from April 19 to December 1, 2013.

The epidemic re-emerged in 2014. Because the date of illness onset was reported for few cases in 2014, we could not show the epidemic curve for the H7N9 outbreak in 2014.

Figure 2. Dates of onset of illness in patients with confirmed H7N9 infection in China from early 2013 to February 10, 2014.



Comparison of the Plateaus by Using the Internet-Based Data

Based on the Internet data, the plateau stages in 2013 and 2014 were from March 31 to April 28, 2013 and from January 10 to February 10, 2014, with 118 case and 176 cases reported, respectively. During the 2 plateau periods, cases in 2014 were younger and more likely to live in rural areas than those in 2013. The fatality rate in 2014 was lower than that in 2013 and the sex ratio was similar in the 2 years (Table 3).

Table 3. Comparison of the characteristics of patients with confirmed H7N9 infection in China between different plateau periods

Characteristics	Confirmed H7N9 cases[a]		*P*
	The first plateau (Mar 31 to Apr 28, 2013)	The second plateau (Jan 10 to Feb 10, 2014)	
Total number of cases	118	176	NE[b]
Age (years)			
Number of cases	118	175	NE[b]
Median age (IQR[c])	61.5 (48-74)	57.0 (41-67)	.01
% ≥ 65 years	51 (43.2)	49 (28.0)	.01
%<64 years	67 (56.8)	126 (72.0)	.01
Sex, n/N (%)			
Male	83/118 (70.3)	115/175 (65.7)	.41
Female	35/118 (29.7)	60/175 (34.3)	
Area of residence, n/N (%)			
Urban	82/104 (78.8)	72/130 (55.4)	<.001
Rural	22/104 (21.2)	58/130 (44.6)	
Poultry workers, n/N (%)	5/88 (5.7)	2/95 (2.1)	.27
Clinical outcomes, n/N (%)			
Hospitalization	107/118 (90.7)	93/93 (100.0)	NE[b]
ICU admission	30/56 (53.6)	2/4 (50.0)	NE[b]
ARDS[d]	12/14 (85.7)	1/5 (20.0)	NE[b]
Death	27/118 (22.9)	23/176 (13.1)	.03

[a]Unless stated otherwise, data in the table provided are the number of patients with a certain characteristics, the total number of patients having data on that characteristic, and the corresponding percentage.

[b]NE: not estimable

[c]IQR: interquartile range

[d]ARDS: respiratory distress syndrome

Discussion

Principal Findings

The findings of this study show that the epidemiological and clinical characteristics for a disease outbreak like H7N9 can be investigated by using information publicly available on the Internet. We updated the outbreak up to February 10, 2014, and found a total of 334 confirmed cases distributed in 15 regions in China. This is probably the most updated report of this outbreak so far. Patients aged between 2.5 and 91 years, but the majority were old people (median age 58 years). A total of 67.0% were males, 66.9% lived in urban areas, and 93.7% of the reported cases were hospitalized. The overall case fatality was 21.7%. We found no evidence for human-to-human transmission of the infection.

Comparison With Prior Work

As compared with a well-conducted conventional epidemiologic study [17], the findings of our study are consistent in most of the epidemiological characteristics examined but show some discrepancies primarily in clinical characteristics of patients. In particular, the discrepancies were observed in the percentage of patients hospitalized, admitted to ICUs, and diagnosed with ARDS. In addition, our study also provided fewer details than Li's study on the history of exposure to animals and on family clusters. Importantly, the case fatality in our study was lower than that reported in Li's study. However, this problem can be easily corrected by the summary data available on multiple Internet sources. The discrepancies may be mainly a result of incompleteness of data on the Internet. For example, by December 1, 2013, we only acquired 53 cases with adequate information on the time from illness onset to first medical care, whereas Li reported 137 cases with relevant information.

The similarity of the results between our study and Li's report [17] indicates that Internet-based data sources can provide useful information on basic epidemiological characteristics of an outbreak like H7N9. The features and progress of an outbreak can be studied in a timely manner and reported by those who have special analytical skills, but who would not have access to necessary data conventionally. Such research can assist the general population and health workers in acquiring a timely overview of an epidemic even before official data are published or updated.

Understandably, Internet-based data may not be complete in some particular details, which are exposure history and clinical characteristics in our case. However, the incompleteness of data may not necessarily have a major impact on understanding and controlling the outbreak because they may still be able to give a restively accurate picture about some or all these factors like many other studies that are based on samples. For example, our study showed identical results on patient history of exposure to animals. On the other hand, 196 new cases were reported after Li's study until February 10, 2014, and our study included twice as many patients as in Li's report. Based on these data, we updated the report and showed a new attack of the infection in early 2014.

As of February 10, 2014, Internet-based data indicated that the case fatality rate of H7N9-infected patients (21.7%) was considerably lower than that of H5N1-infected patients in both China (70%) and worldwide (59%) [20,21], but higher than that (10%) of patients with (SARS) [22]. Our study and Li's report also indicate that the H7N9 virus tends to infect people who are much older than those infected by the H5N1 virus [20].

In addition, a subgroup analysis suggested that the case fatality rate before the first H7N9 case was reported (ie, March 31, 2013) was substantially higher than that after that. It could be partially explained by the incomplete reporting of patients with mild or no symptoms before the first case was reported. Because of the lack of awareness and limitations in the surveillance and detection process at the beginning of the outbreak, it is highly probable that the actual number of H7N9-infected cases was underestimated at the early stage of the outbreak [17]. Another possible explanation could be that the surveillance, prevention, and treatment of the disease improved after more cases were reported, which helped doctors and other health workers more effectively deal with the patients than during the early stage of the outbreak.

Strengths and Limitations

The results of this study indicate that the Internet could be used for investigating epidemiological and clinical characteristics of such infectious disease outbreaks as the H7N9 outbreak. Different from many Web-based biosecurity intelligence systems such as HealthMap, MedISys, and BioCaster, our investigation is aimed to provide a comprehensive and more detailed investigation of a disease outbreak. Most of these intelligence systems are established merely for alerting of health threats by providing only a summary number of cases over time without detailed information important and necessary for understanding and controlling an infectious disease outbreak.

Compared with traditional epidemiological investigations of disease outbreaks, Internet-based investigations have a few important advantages. First, Internet-based data can be quickly analyzed and reported and initial results can be useful in facilitating efficient contingency planning in particular in the early stage of an epidemic. Second, the coverage of the data has no boundary and can be national or international and thus can provide an overall picture of the problem any local and regional data cannot. Third, as publically available Internet data are open to any potentially interested people worldwide with the skills to do the analyses. This can greatly facilitate the timely analysis and update of an outbreak so that rapid and appropriate action can be taken for its control. Fourth, early initial Internet data and their analyses may greatly help alert and pressure governments and authorities to facilitate timely publication of all the data to the general public, ideally through the Internet, which can facilitate timely action against the outbreak by preventing intents of withholding data for other reasons, such as motivations resulting from political parallelism or the benefits from publishing a high impact factor journal paper. Fifth, Internet data can be easily complemented, counterchecked, and corrected by others who have access to all or some of the original data from different sources. Finally, if an epidemic persists, in particular when it involves only a few sporadic cases, data on newly confirmed cases could be collected through the Internet for analysis and updating.

The major problem of Internet-based investigations is incomplete data. For example, we had individual data on only 33 death cases by December 1, 2013, which resulted in a lower case fatality than that reported by the field study [17]. Exhaustive search of various data sources on the Internet may help reduce missing data. For summary results, such as case fatality, incomplete individual data can often be overcome by using the aggregate data that are normally more conveniently available on the Internet.

Incomplete data may become particularly severe if the epidemic persists for a lengthy period of time or at the late stage as the event is dying out and gradually moving out of the public's attention. In our case, for example, even aggregate data on the number of deaths from December 1, 2013 to February 10, 2014 as the early part of the second outbreak were not traceable on the Internet. We estimated the case fatality only based on limited data on individual cases identified. The estimate was lower than that reported in Li's study [17] for the outbreak period before December 1, 2014, but similar to that reported in Wikipedia [23], which showed that as of March 1, 2014, there were 375 confirmed cases of H7N9 and 80 people had died resulting in a case fatality rate of 21.3%. Maybe the case fatality rate for H7N9 was indeed lower in the second outbreak between the late 2013 and the early 2014 than that in the early 2013. Official aggregate data are yet needed to confirm this result.

Incomplete data are more likely to occur on details of individual cases than the cases themselves. We found clinical characteristics were often missing for individual cases. Furthermore, our epidemic curve was similar in the shape, time period, peak date, and total number of cases to those reported in Li's study before December 1, 2013. However, between December 1, 2013 and February 10, 2014, 196 cases were identified for the second outbreak but the date of illness onset was available only for 30 patients, which prevented us from describing reliably the epidemic curve for the second outbreak. This indicates that Internet-based methods may be, for the time being, more suitable for investigating the early stage of sudden and short-lasting events of public concern.

Conclusions

This study presented a new outbreak of H7N9 infection in China by using publicly accessible data on the Internet. The results of the study agreed in most epidemiological and clinical

characteristics with those from an authoritative conventional epidemiologic study [17]. Internet-based investigations appear particularly useful for sudden, emergent, relatively short-lasting issues or events of great public concern that involve a relatively small number of cases and for which relevant data can be conveniently uploaded and updated on the Internet. Such investigations can be conducted by any interested people who have the skills to do the analyses so that the issue can be publicized more quickly and updated for quicker and better understanding and controlling of the problem. It will greatly facilitate such investigations if governments, authorities, organizations, or individuals can consciously and systematically put relevant data on the Internet and make them publicly accessible for the greater good of the public.

Acknowledgments

This study was funded by a grant for establishing Shenzhen Municipal Key Laboratory for Health Risk Analysis from Shenzhen Science and Technology Innovation Commission (Grant number: SZRI-C.02.12.007).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Individual data for patients with confirmed H7N9 infection from the Internet.

[XLSX File (Microsoft Excel File), 83KB - jmir_v16i9e221_app1.xlsx]

References

1. Jacobs BP, Bent S, Tice JA, Blackwell T, Cummings SR. An internet-based randomized, placebo-controlled trial of kava and valerian for anxiety and insomnia. Medicine (Baltimore) 2005 Jul;84(4):197-207. [Medline: 16010204]
2. Hatton C. BBC. 2014 Jan 28. China's internet vigilantes and the 'human flesh search engine' URL: http://www.bbc.com/news/magazine-25913472 [accessed 2014-09-16] [WebCite Cache ID 6SbcK78AS]
3. Kaiman J. The Guardian. 2013 Sep 05. China's 'Brother Wristwatch' Yang Dacai jailed for 14 years for corruption URL: http://www.theguardian.com/world/2013/sep/05/china-brother-wristwatch-yang-dacai-sentenced [accessed 2014-08-06] [WebCite Cache ID 6RcDjGsNG]
4. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]
5. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
6. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
7. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
8. Troullos E, Baird L, Jayawardena S. Common cold symptoms in children: results of an Internet-based surveillance program. J Med Internet Res 2014 Jun;16(6):e144 [FREE Full text] [doi: 10.2196/jmir.2868] [Medline: 24945090]
9. Yom-Tov E, Borsa D, Cox IJ, McKendry RA. Detecting disease outbreaks in mass gatherings using Internet data. J Med Internet Res 2014 Jun;16(6):e154 [FREE Full text] [doi: 10.2196/jmir.3156] [Medline: 24943128]
10. van Beelen ME, Beirens TM, den Hertog P, van Beeck EF, Raat H. Effectiveness of web-based tailored advice on parents' child safety behaviors: randomized controlled trial. J Med Internet Res 2014 Jan;16(1):e17 [FREE Full text] [doi: 10.2196/jmir.2521] [Medline: 24463421]
11. Eysenbach G. Infovigil Project. 2006. URL: http://www.infodemiology.org/ [accessed 2014-09-02] [WebCite Cache ID 6SHU7EDQs]
12. Nelson R. HealthMap: the future of infectious diseases surveillance? The Lancet Infectious Diseases 2008 Oct;8(10):596. [doi: 10.1016/S1473-3099(08)70219-6]
13. Rortais A, Belyaeva J, Gemo M, Van DGE, Linge J. MedISys: An early-warning system for the detection of (re-)emerging food- and feed-borne hazards. Food Res Int 2010 Jun;43(5):1553-1556. [doi: 10.1016/j.foodres.2010.04.009]
14. Collier N, Doan S, Kawazoe A, Goodwin RM, Conway M, Tateno Y, et al. BioCaster: detecting public health rumors with a Web-based text mining system. Bioinformatics 2008 Dec 15;24(24):2940-2941 [FREE Full text] [doi: 10.1093/bioinformatics/btn534] [Medline: 18922806]
15. Lyon A, Nunn M, Grossel G, Burgman M. Comparison of web-based biosecurity intelligence systems: BioCaster, EpiSPIDER and HealthMap. Transbound Emerg Dis 2012 Jun;59(3):223-232. [doi: 10.1111/j.1865-1682.2011.01258.x] [Medline: 22182229]

16. Kittler AF, Hobbs J, Volk LA, Kreps GL, Bates DW. The Internet as a vehicle to communicate health information during a public health emergency: a survey analysis involving the anthrax scare of 2001. J Med Internet Res 2004 Mar 3;6(1):e8 [FREE Full text] [doi: 10.2196/jmir.6.1.e8] [Medline: 15111274]
17. Li Q, Zhou L, Zhou M, Chen Z, Li F, Wu H, et al. Epidemiology of human infections with avian influenza A(H7N9) virus in China. N Engl J Med 2014 Feb 6;370(6):520-532. [doi: 10.1056/NEJMoa1304617] [Medline: 23614499]
18. Patients with confirmed H7N9 infection occurred in several regions in China. 2013. URL: http://news.sina.com.cn/z/qlg2013/ [accessed 2014-09-14] [WebCite Cache ID 6SaStVJmX]
19. The battle against the outbreak of H7N9 avian influenza A in China. 2013. URL: http://health.sohu.com/s2013/qlg/ [accessed 2014-09-14] [WebCite Cache ID 6SaSwegJE]
20. Cowling BJ, Jin L, Lau EH, Liao Q, Wu P, Jiang H, et al. Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases. Lancet 2013 Jul 13;382(9887):129-137 [FREE Full text] [doi: 10.1016/S0140-6736(13)61171-X] [Medline: 23803488]
21. World Health Organization. 2013. Cumulative number of confirmed human cases for avian influenza A (H5N1) reported to WHO 2003-2013 URL: http://www.who.int/influenza/human_animal_interface/EN_GIP_20130116CumulativeNumberH5N1cases.pdf [accessed 2014-08-06] [WebCite Cache ID 6RcQOKLvi]
22. World Health Organization. 2003. Summary of probable SARS cases with onset of illness from 1 November to 31 July 2002 URL: http://wwwwhoint/csr/sars/country/table2004_04_21/en/indexhtml [accessed 2014-08-06] [WebCite Cache ID 6RcQWVAti]
23. Wikipedia. 2014. Secondary avian influenza A H7N9 subtype URL: http://zh.wikipedia.org/wiki/%E7%94%B2%E5%9E%8B%E6%B5%81%E6%84%9F%E7%97%85%E6%AF%92H7N9%E4%BA%9E%E5%9E%8B [accessed 2014-08-06] [WebCite Cache ID 6RcR7DPMf]

Abbreviations

ARDS: acute respiratory distress syndrome
CDC: Center for Disease Control and Prevention
ICU: intensive care unit
WHO: World Health Organization
SARS: severe acute respiratory syndrome

Edited by G Eysenbach; submitted 08.08.14; peer-reviewed by SG Yang, J Hawkins; comments to author 25.08.14; accepted 06.09.14; published 25.09.14.

Please cite as:
Mao C, Wu XY, Fu XH, Di MY, Yu YY, Yuan JQ, Yang ZY, Tang JL
An Internet-Based Epidemiological Investigation of the Outbreak of H7N9 Avian Influenza A in China Since Early 2013
J Med Internet Res 2014;16(9):e221
URL: http://www.jmir.org/2014/9/e221/
doi:10.2196/jmir.3763
PMID:25257217

Original Paper

A Medical Consultation Service on Facebook: Descriptive Analysis of Questions Answered

Otto Helve[1], MD, PhD

Children's Hospital, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland

Corresponding Author:
Otto Helve, MD, PhD
Children's Hospital
Helsinki University Central Hospital
University of Helsinki
Stenbäckinkatu 11
PB 281
Helsinki, 00029 HUS
Finland
Phone: 358 505824426
Fax: 358 947175299
Email: otto.helve@helsinki.fi

Abstract

Background: Social media is used increasingly by the general public to access health information. However, a lack of models for health information distribution limits the presence of publicly funded services on social media sites.

Objective: The goal of the study was to present a model for delivering child health information to parents through a social media site.

Methods: A Facebook site was launched for 11 months based on a question-and-answer service produced by a pediatrician and open to Facebook users over 18 years old. If the answer did not include a further referral to a health care service provider, the question was considered comprehensively answered. The site was funded by a pharmaceutical company, and it included an advertisement of a pharmaceutical product for children's fever and pain.

Results: During the study, 768 questions were submitted: an average of 69.8 (SD 31.7) per month. There were 245,533 independent Facebook users on the site, with an average of 727.0 (SD 2280.6) per day. Infections were the most common theme in questions (355/768, 46.2%). Questions were more likely to be comprehensively answered if they were related to infections (279/355, 78.6%) than questions related to non-infectious symptoms (265/423, 64.2%, *P*=.003).

Conclusions: On this site aimed at parents of small children, personalized answers were an effective way of delivering information. The service is likely to have reduced the need for further contacts with a health care service provider in more than half of the cases. The site could serve as a model for publicly funded health information distribution.

(J Med Internet Res 2014;16(9):e202) doi:10.2196/jmir.3194

KEYWORDS

social media; Internet; health information; health promotion

Introduction

Social media has changed the way the general public accesses health information [1]. With the increasing use of handheld mobile devices, health information may be readily available to an even larger public, as seen in the United States [2]. Facebook, the social media tool with the widest active user base, has 1.28 billion users as of March 31, 2014 [3]. Health information is distributed and discussed on many social media tools, and one in five Americans uses social media as a source of health care information [4]. Of Facebook users in the United States, 94% have used Facebook to gather information on their health [4].

However, user-generated health content on social media is generally inconsistent with clinical guidelines and professional knowledge, and non-biased information can be difficult to obtain [5]. Publicly funded services and institutes have a limited role in distributing health information on social media [6], possibly due to a lack of models for using social media in health

information promotion [7,8]. In Finland, information on child health is traditionally delivered by "well-baby" clinics, which offer both general information on health care and address individual needs. At the moment, most of the information is delivered during patient visits and by telephone consultation. The clinics are understaffed with regard to the national recommendations, and this is reflected in their capacity to deliver health information [9]. It is evident that methods to reach a larger population are needed to meet the demand. As in the United States, although no such data exist, it is likely that the use of social media for seeking health information is increasing also in Finland. Stroever et al [10] found that social media was an effective way to communicate child health information to low-income parents. Furthermore, the information was considered reliable if distributed by perceived experts.

While social media can be used to distribute specific information on issues addressed by individual Internet users, these discussions can be accessed by countless others on an open site, making the information available to a larger audience. Depicted in our study is a social media site for distributing child health information to parents. The information was distributed by a pediatrician. The aim was to determine the number of questions generated by the site, the profile of parents' concerns, and the number of visitors on the site. Also, the parental contact was assessed by evaluating whether or not the answers to questions included further referral to a health care provider.

Methods

Material

A Facebook site, Kysy lastenlääkäriltä ("Ask a pediatrician"; Figure 1) based on a question-and-answer service produced by a pediatrician (the author) was established in September 2010. The site was funded by a Finnish pharmaceutical company and included an advertisement for acetaminophen (paracetamol), a pharmaceutical product for children's fever and pain. The company had no control over issues discussed on the site or answers given by the pediatrician. The site was run in Finnish and advertised in national advertisement campaigns in magazines aimed at parents of small children. The service was open to all Facebook users aged over 18 years.

The local ethics committee approved the study.

Figure 1. Screenshot of Kysylastenlaakarilta Facebook page.

Methodology

A pilot service was initiated at 2- to 3-week irregular intervals depending on demand between September 2010 and March 2011. This experience showed that the demand for a more continuous service was substantial. Therefore, the service was launched for a continuous 7-month period from September 2011, including a 2-week Christmas closure. During the closure, the Facebook site remained open so that previously written questions and answers could be seen by visitors, but no new questions were answered. After a summer closure from April to the end of August 2012, the service was opened again in September until December 14, 2012 (4 months). The results from these 11 months are used in this analysis.

The service was intended to answer questions on children's fever and pain, but parental postings included questions on all subjects. A non-medically trained moderator evaluated all postings and messages left on the site twice daily, and postings that included product labels were removed according to Finnish law on pharmaceutical product promotion. However, the sender was given the opportunity to re-send the posting if the product label was changed to a generic name. All questions were answered by the pediatrician (the author), even if they were outside the focus of the service. The pediatrician was employed on an hourly basis to the site, and at the time was also a research coordinator at a clinical postgraduate school at the University of Helsinki, Finland. The pediatrician answered questions once daily on weekdays. In addition, the moderator contacted the pediatrician regarding possible emergency-related questions potentially requiring an urgent response, in which case the parent received a notification to seek immediate treatment.

Analysis

Typically, each posting included several questions. Postings were first analyzed for the number of questions, and then the questions were categorized according to the symptoms presented (allergies, nutrition, neurology, etc). Most of the questions did not include sufficient information to set a specific diagnosis. Therefore, categorization was performed in relation to the clinical entity the question belonged to. Infections were further categorized into subgroups. Otitis media (a middle ear infection) was the largest subgroup and as such was included in the analysis. Each of the other infection subgroups included only 1-30 questions, and thus these subgroups were omitted from further analyses, although included in the infection category.

Answers were evaluated by the author by reviewing the questions in each posting separately. If the answer did not include a referral to a health care service provider, the question was considered comprehensively answered. If the answer included a referral at a later time should certain criteria be met (such as in the case of new symptoms), the question was also considered comprehensively answered.

Analyses were performed using GraphPad Prism 6.0. Question data are presented as mean (SD). Proportions were compared using the chi-square test. User statistics were gathered from information in the Facebook user database, Page insights.

Results

Respondents

During the study, 245,533 independent Facebook users visited the site, an average of 727.0 (SD 2280.6) per day. During closure of the service (April to August 2012), there were an average of 17.4 (SD 48.8) daily visitors. In the average month, 75.0% of the visitors were women (SD 13.7). The largest visitor age group was 25-34 years in both women and men: 39.6% (SD 2.9) and 11.5% (SD 2.4) of all monthly visitors, respectively. Most of the visitors were inhabitants of the five largest cities in Finland (data not shown). By December 31, 2012, there were 12,328 Facebook likes since the beginning of the service.

Postings

There were an average of 42.6 (SD 27.6) monthly postings from parents, 478 in total. Men sent a total of 9 postings. Each posting contained a mean 1.6 (SD 0.6) questions (N=768), 69.8 (SD 31.7) questions per month. Of all postings, 63.4% (303/478) were comprehensively answered.

Questions

Infections were the most common theme in questions (355/768, 46.2%), with otitis media–related questions being the most common subgroup of infections (69/768, 9.0% of all questions). There were 413 non–infection-related questions during the study period. The other symptom groups with more than 30 questions in the study period were non-infectious rash (64/768, 8.3%), allergies (55/768, 7.2%), fever (eg, help in choosing the right thermometer, 39/768, 5.1%), neurological symptoms (34/768, 4.4%), and nutrition (30/768, 3.9%, Figure 2). Questions that could not be categorized into the above entities were grouped into miscellaneous questions (54/768, 7.0%).

In general, only one exchange of postings occurred for each question. However, the need for referral was evaluated from each posting by the pediatrician. Questions relating to fever were most likely to be comprehensively answered (36/39, 92.3%) while questions relating to neurological symptoms were least likely (19/34, 55.9%; Figure 3). Questions were more likely to be comprehensively answered if they were related to infections (279/355, 78.6%) than questions related to non-infectious symptoms (265/413, 64.2%, chi-square=15.29, P=.003). Within the otitis media category, 85.5% of questions (59/69) were considered comprehensively answered.

Large seasonal variation in question topics was present. Infection-related questions dominated in the fall (Figure 4), tracing a similar pattern to reasons for ambulatory pediatric doctor's visits.

Figure 2. Number of questions categorized into topics (topics with less than 30 questions not shown).

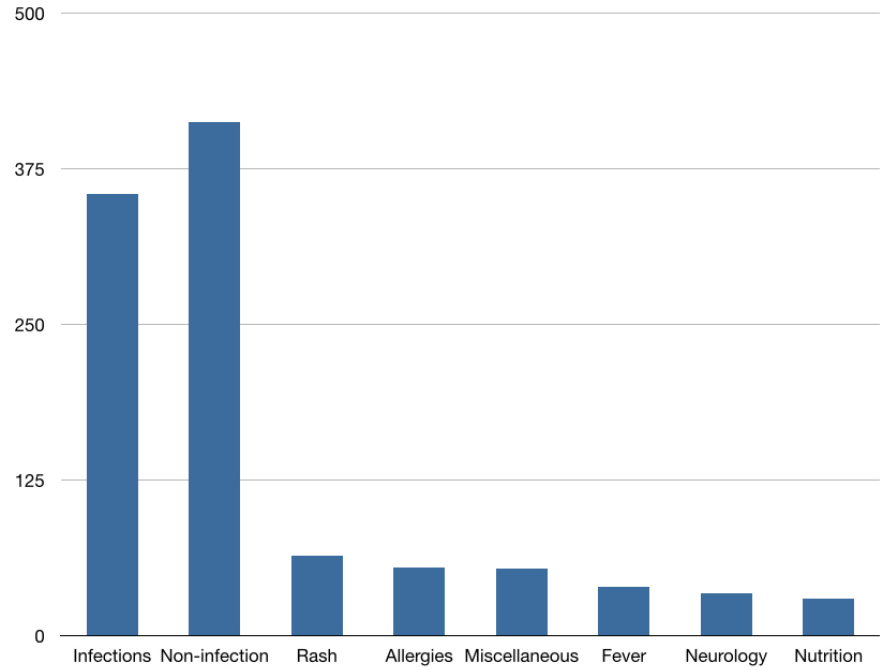


Figure 3. The proportion (%) of questions comprehensively answered (in blue) and the proportion requiring further contact with a health care worker (in green); P=.003 for the difference between questions comprehensively answered in the infections category and the non-infection category.

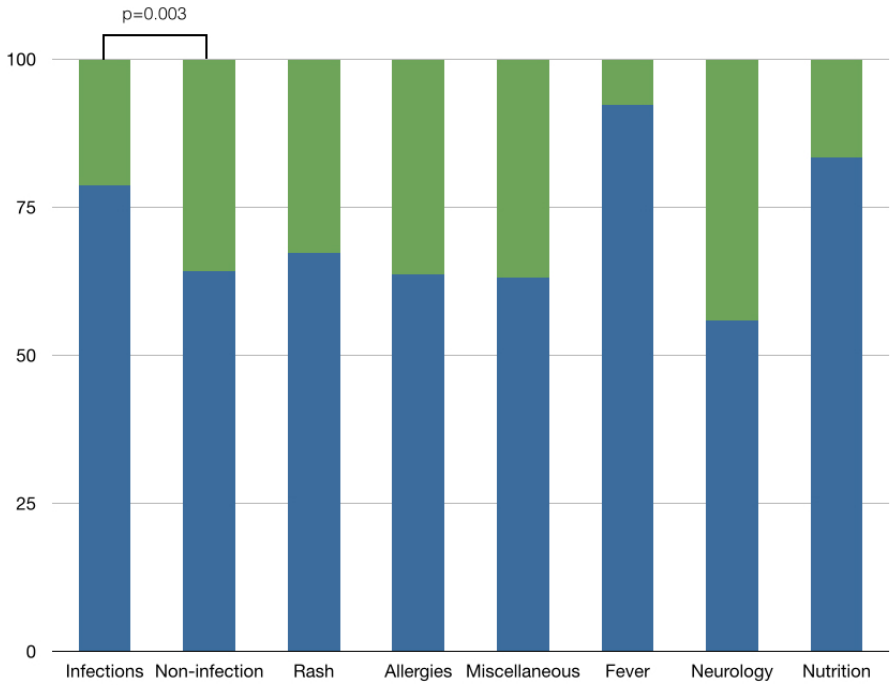
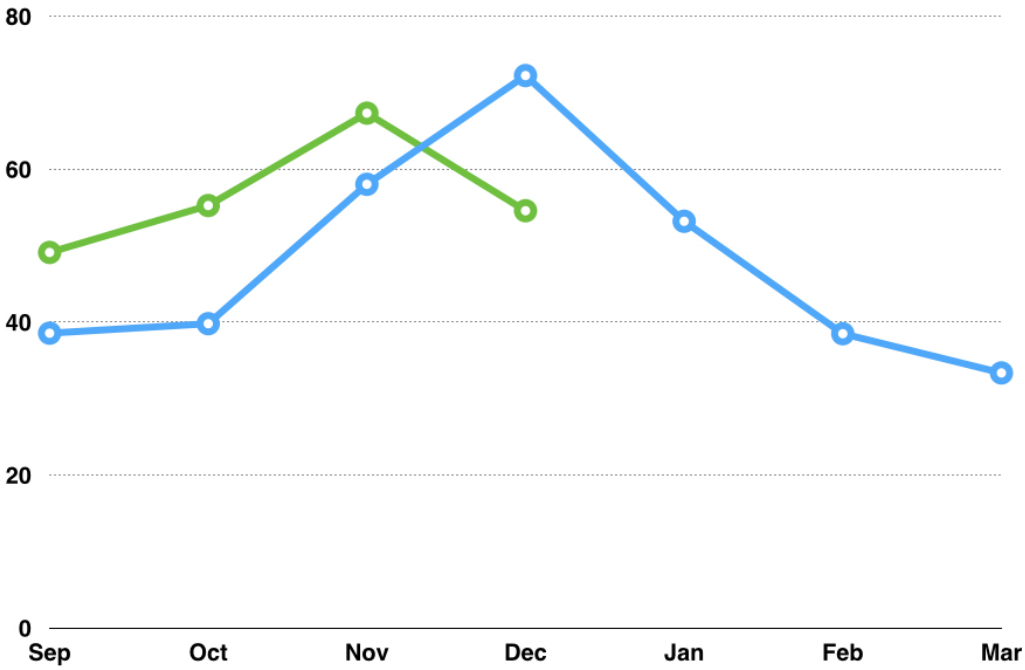


Figure 4. The percentage of questions relating to infections in 2011-2012 (blue) and 2012 (green).



Discussion

Principal Findings

This study provides an overview of a social media site for distributing child health information, a profile of the questions asked by parents, the proportions of questions comprehensively answered, and the volume of visitors on the site. On this site, which was aimed at parents of small children, personalized answers were an effective way of delivering information. Based on the number of questions considered comprehensively answered, the service may have reduced the need for further contacts with a health care service provider in more than half of the cases.

The service was directed to questions on child health. According to the site visitor profile, mostly women aged 29-34 years, it is likely that the audience was reached. The number of individual visits was almost 250,000. There was a great variation in the amount of daily visitors, but no trend could be seen to explain this. The number of postings generated by this audience was 478, so it is likely that the answer given to one posting was read by multiple users and information spread was greater than just one person, which is typical for this kind of media [11]. Evaluating the efficiency of social media is difficult [12], and choosing the correct key performance indicator is crucial [13]. In this study, the number of individual daily visitors was used alongside the number of actual postings. From these data, it is obvious that a need exists for a social media-based child health distribution site in Finland. In this study, it was not possible to assess the effect of the service on in-person visits to the health care service provider. This should be the focus of further studies.

Infections were the most common topic for parental questions. These are also the most typical cause for ambulatory pediatric doctor's visit. Pediatric infectious diseases is a subspecialty within pediatrics, and this finding supports the fact that a specialist of this field would be the optimal person to work on a site with such a focus. In addition, the seasonal variation in questions on infections followed the general trend of pediatric infections in Finland. This underlines the fact that the questions presented were not just sporadically written, but followed general trends.

Social media is an interactive and sometimes face-to-face information distribution channel, and it is therefore considered a more reliable media than passive Internet pages [14]. It is unlikely that users evaluate critically the motives behind the information given [14]. Because the reliability of health information distributed through social media tools may be contradictory to medical guidelines [5] and because commercial interests may play a large role in delivering health information on social media [15], it is clear that information in line with current medical guidelines should also be easily accessible through social media channels. In specific fields of medicine, a site such as the one presented here could be used to effectively deliver guideline-derived health information to patients, for example by publicly funded institutes. For parents of small children, the trustworthiness of health information has been shown to be dependent on the expertise of the person answering the question. If the information is delivered by someone that the parents consider an expert, it is accepted as credible [10]. It is possible that the commercially funded service depicted on the site received a wide audience, because the questions were answered by a pediatrician.

The proportion of questions comprehensively answered was over 70%. Without this social media contact, many of these questions would likely have been presented to a health care service provider, possibly by making an appointment to see a doctor. Although it was not possible to evaluate in the study,

the service may well have reduced the pressure on "well-baby" clinic personnel. Questions about infections were the most likely to be comprehensively answered, while questions concerning neurological symptoms were least likely to be comprehensively answered, demonstrating the difficulty in assessing neurological symptoms in general. In health information technology, effectiveness of the service is rarely evaluated [16]. On the current site, answers were evaluated by the pediatrician. In future studies, a patient-based evaluation should be included in the service.

When planning a site for health information, the scope has to be defined. If the scope is comprehensive and aimed at, for example, a wide array of child health issues, the expertise also needs to be more comprehensive. This is of course more costly. On the other hand, a narrow scope can also be recommended. The experience of this site suggests that even a narrow focus on infection-related questions only would have been functional especially because these questions were more likely to be comprehensively answered than non–infection-related questions. Here, the focus of the questions changed over time and the number of infection-related questions increased towards winter. Therefore, the expertise needed to answer the questions may vary over time. Another issue to be decided on is the lapse between questions and answers. It is recommended here that the lapse be fairly short. On this site, the lapse was 24 hours on weekdays and more than 2 days at weekends. This may have resulted in a lower number of visits and postings. The method here would serve its purpose best if it were focused narrowly and manned in such a way that the delay in answers was shorter.

Limitations

One of the crucial concerns when launching a site is the possible discrepancy between the demand and the number of personnel answering to the demand, that is, the possibility of an overflow of questions. According to the data from this site, the number of questions is likely to be rather limited in comparison with the number of visitors. Hence, one question and answer may be viewed a multitude of times and the information delivered efficiently. While active social media participants are usually a minority [16], it is possible that the low activity on this site could have resulted from the site also being an advertisement. However, on average, one daily visitor would visit the site twice during the day (data not shown), which more likely suggests an active interest on the question-and-answer forum rather than on the static advertisement on the site.

Another possible limitation to the study is the fact that having a pharmaceutical product advertisement on the site may have skewed the focus of the questions presented on the site. As the product was an antipyretic drug, it is likely the possible effect was an increase in the number of questions on infections.

Conclusions

There is a constant demand for social media–based child health information services. A social media health information site is likely to be well accepted if the consultant is a specialist and, when delivering child health information, possibly a pediatrician. On this site aimed at parents of children in Finland, the greatest number of questions was presented between September and January, and most of these concerned infections. Over half of the questions were estimated to be comprehensively answered, and most likely did not require further contact with a health care service provider. These factors should be taken into account when planning such a service.

Acknowledgments

The author thanks Leiras Takeda Finland for the use of Facebook data in this study. Professors Tytti Solantaus and Sture Andersson are thanked for valuable comments on the manuscript and Carol Ann Pelli for linguistic editing.

Conflicts of Interest

The author was paid a consultation fee by Leiras Takeda Finland for running a question-and-answer service on a commercial site on Facebook aimed at promoting a pharmaceutical product.

References

1. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
2. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
3. Facebook company website. URL: http://newsroom.fb.com/company-info/ [accessed 2014-06-16] [WebCite Cache ID 6QMoy05EW]
4. Corporation TNR. 2010. 1 in 5 Americans use social media for health information URL: http://hcmg.nationalresearch.com/public/News.aspx?ID=9 [accessed 2013-11-24] [WebCite Cache ID 6LNM1u2rQ]
5. Chou WY, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. Am J Public Health 2013 Jan;103(1):e9-18. [doi: 10.2105/AJPH.2012.301071] [Medline: 23153164]
6. Syed-Abdul S, Fernandez-Luque L, Jian WS, Li YC, Crain S, Hsu MH, et al. Misleading health-related information promoted through video-based social media: anorexia on YouTube. J Med Internet Res 2013;15(2):e30 [FREE Full text] [doi: 10.2196/jmir.2237] [Medline: 23406655]

7. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]
8. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
9. Hakulinen-Viitanen T, Pelkonen M, Haapakorva A. Maternity and child health care in Finland. In: Reports of the Ministry of Social Affairs and Health. Helsinki, Finland: Ministry of Social Affairs and Health; 2005.
10. Stroever SJ, Mackert MS, McAlister AL, Hoelscher DM. Using social media to communicate child health information to low-income parents. Prev Chronic Dis 2011 Nov;8(6):A148 [FREE Full text] [Medline: 22005641]
11. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
12. Hussain AA. Meaningful use of information technology: a local perspective. Ann Intern Med 2011 May 17;154(10):690-692. [doi: 10.7326/0003-4819-154-10-201105170-00010] [Medline: 21576537]
13. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
14. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. J Health Commun 2011;16 Suppl 1:45-58. [doi: 10.1080/10810730.2011.589881] [Medline: 21843095]
15. Gregerson J. Truth, lies, and rumors in the media: consider the source. J Acad Nutr Diet 2012 May;112(5):602-609. [doi: 10.1016/j.jand.2012.03.007] [Medline: 22709761]
16. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

Edited by G Eysenbach; submitted 17.12.13; peer-reviewed by G Huang, A Holton; comments to author 28.04.14; revised version received 17.06.14; accepted 17.07.14; published 04.09.14.

Please cite as:
Helve O
A Medical Consultation Service on Facebook: Descriptive Analysis of Questions Answered
J Med Internet Res 2014;16(9):e202
URL: http://www.jmir.org/2014/9/e202/
doi:10.2196/jmir.3194
PMID:25189182

Original Paper

Web-Based Self-Assessment Health Tools: Who Are the Users and What Is the Impact of Missing Input Information?

Nicole Neufingerl[1], MSc; Mark R Cobain[2], PhD; Rachel S Newson[1], PhD

[1]Nutrition & Health Department, Unilever Research & Development, Vlaardingen, Netherlands

[2]New Business Unit, Unilever Research & Development, London, United Kingdom

Corresponding Author:
Nicole Neufingerl, MSc
Nutrition & Health Department
Unilever Research & Development
Olivier van Noortlaan 120
Vlaardingen, 3133 AT
Netherlands
Phone: 31 10460 ext 5155
Fax: 31 104605993
Email: nicole.neufingerl@unilever.com

Abstract

Background: Web-based health applications, such as self-assessment tools, can aid in the early detection and prevention of diseases. However, there are concerns as to whether such tools actually reach users with elevated disease risk (where prevention efforts are still viable), and whether inaccurate or missing information on risk factors may lead to incorrect evaluations.

Objective: This study aimed to evaluate (1) evaluate whether a Web-based cardiovascular disease (CVD) risk communication tool (Heart Age tool) was reaching users at risk of developing CVD, (2) the impact of awareness of total cholesterol (TC), HDL-cholesterol (HDL-C), and systolic blood pressure (SBP) values on the risk estimates, and (3) the key predictors of awareness and reporting of physiological risk factors.

Methods: Heart Age is a tool available via a free open access website. Data from 2,744,091 first-time users aged 21-80 years with no prior heart disease were collected from 13 countries in 2009-2011. Users self-reported demographic and CVD risk factor information. Based on these data, an individual's 10-year CVD risk was calculated according to Framingham CVD risk models and translated into a Heart Age. This is the age for which the individual's reported CVD risk would be considered "normal". Depending on the availability of known TC, HDL-C, and SBP values, different algorithms were applied. The impact of awareness of TC, HDL-C, and SBP values on Heart Age was determined using a subsample that had complete risk factor information.

Results: Heart Age users (N=2,744,091) were mostly in their 20s (22.76%) and 40s (23.99%), female (56.03%), had multiple (mean 2.9, SD 1.4) risk factors, and a Heart Age exceeding their chronological age (mean 4.00, SD 6.43 years). The proportion of users unaware of their TC, HDL-C, or SBP values was high (77.47%, 93.03%, and 46.55% respectively). Lacking awareness of physiological risk factor values led to overestimation of Heart Age by an average 2.1-4.5 years depending on the (combination of) unknown risk factors (P<.001). Overestimation was greater in women than in men, increased with age, and decreased with increasing CVD risk. Awareness of physiological risk factor values was higher among diabetics (OR 1.47, 95% CI 1.46-1.50 and OR 1.74, 95% CI 1.71-1.77), those with family history of CVD (OR 1.22, 95% CI 1.22-1.23 and OR 1.43, 95% CI 1.42-1.44), and increased with age (OR 1.05, 95% CI 1.05-1.05 and OR 1.07, 95% CI 1.07-1.07). It was lower in smokers (OR 0.52, 95% CI 0.52-0.53 and OR 0.71, 95% CI 0.71-0.72) and decreased with increasing Heart Age (OR 0.92, 95% CI 0.92-0.92 and OR 0.97, 95% CI 0.96-0.97) (all P<.001).

Conclusions: The Heart Age tool reached users with low-moderate CVD risk, but with multiple elevated CVD risk factors, and a heart age higher than their real age. This highlights that Web-based self-assessment health tools can be a useful means to interact with people who are at risk of developing disease, but where interventions are still viable. Missing information in the self-assessment health tools was shown to result in inaccurate self-health assessments. Subgroups at risk of not knowing their risk factors are identifiable and should be specifically targeted in health awareness programs.

(J Med Internet Res 2014;16(9):e215) doi:10.2196/jmir.3146

KEYWORDS

cardiovascular disease; risk assessment; Web applications; consumer health information; preventive health services; cholesterol; blood pressure

Introduction

During the last decade, access to the Internet and the proportion of the population using it to seek health information have markedly increased [1,2]. Online health information seekers are typically searching for information regarding a specific disease or medical problem, including the potential to diagnose their own health status [3,4]. Health professionals, public health and governmental organizations, and private health providers are therefore using the Internet as a medium to disseminate health information and preventative educational programs. This makes the Internet a valuable instrument for increasing consumer awareness, promotion of healthy behaviors, and disease prevention [5,6]. Web-based health applications, such as self-assessment health tests or behavior change programs, that combine high-quality health information with interactive components can therefore play a role to benefit prevention, early detection, or treatment of non-communicable diseases [7-9].

The impact that Web-based health assessment tools can have on public health depends on the audience they reach, as well as on the quality and reliability of the information they provide. Criticism of Web-based health applications has been raised suggesting that they may not be reaching the people who need these tools the most: those at risk of developing disease, and where prevention efforts are still viable [10,11]. Furthermore, there is concern about the potential harm that can be caused by inaccurate health assessments provided on the Internet that can deliver incorrect diagnoses and/or cause delays in seeking appropriate medical care [12,13]. This can arise due to poorly designed Web applications that are not based on scientific evidence. Alternatively, inaccurate health assessment results may arise from users providing inaccurate information due to inputting the information incorrectly, not understanding questions, or they may not know the answer to the questions [14]. This may be particularly pertinent to tools that require information on physiological risk factors, which are unknown to a large part of the population [5,15].

The potential impact of Web-based health assessment tools on disease prevention is large, but they need to reach users with an elevated disease risk and provide accurate health assessments. Therefore, the aim of the current study was to (1) evaluate whether a Web-based health assessment tool communicating cardiovascular disease (CVD) risk was reaching users at risk of developing CVD, (2) evaluate the impact of awareness of physiological risk factors on the health assessment provided by the CVD risk communication tool, in particular total cholesterol (TC), HDL-cholesterol (HDL-C), and systolic blood pressure (SBP), and (3) evaluate the key predictors of awareness and reporting of physiological risk factors in order to understand who in the general population is more likely to provide physiological measures to Web-based tools.

Methods

Study Design

The current study was based on a global database from users of the Heart Age tool. Heart Age was developed to help users better understand their risk of CVD, which when given as a traditional percentage can be a difficult concept to understand (eg, “your risk of a CVD event in the next 10 years is 12%”) [16,17]. Heart Age is a Web-based tool that obtains an individual’s CVD risk factor information through a series of questions, calculates percentage risk of developing CVD within the next 10 years according to the Framingham risk score, and then translates this risk into a “Heart Age” [18]. An individual’s Heart Age corresponds to the age of a person with the same predicted CVD risk but with all other risk factors considered as normal. This means that depending on their CVD risk, a person’s Heart Age could be younger or older than their chronological age. For example, a 55-year-old man with normal risk factors would have a 10% risk of CVD; accordingly, a 40-year-old man with a 10% risk of CVD (due to unhealthy risk factors) would have a Heart Age of 55 years. The Heart Age tool aims to make users aware of their risk of developing a CVD event and which risk factors are contributing to this risk, in order to motivate them to make lifestyle changes. Therefore the target audience is anyone at risk of developing a CVD event, but where prevention efforts are still viable, as these users can benefit the most from awareness of their current status and potential interventions. After completion of the Heart Age test, users can sign up for a free Web-based health program providing personalized diet and lifestyle advice to lower one’s Heart Age. The Heart Age tool is directed at the general population, but in particular it intends to reach users with unhealthy risk factors. Heart Age was launched globally in 2009 as part of a brand marketing campaign by Flora/Becel, a margarine brand, via on-pack messages, advertisements on television, newspapers, in store applications, etc. The campaign focused on raising awareness of heart health; it was not targeted at specific at risk groups. The tool is available via free open access websites available in different languages [19].

Users

Data from 3,374,769 users of the Heart Age tool was collected between July 2009 and December 2011. The tool was launched in 14 countries, including United Kingdom, Germany, the Netherlands, Belgium, Finland, Austria, Poland, Turkey, Ireland, Portugal, Slovenia, Greece, Australia, and Brazil. By using the tool, users consented to the privacy policy, which could be accessed via a link on the website, and which stated that personal information provided could be collected and used in an aggregated way to evaluate the use of the site and services provided. Visitors to the website did not receive any incentives for using the tool. All of the information users entered was stored in a format where individuals could not be identified. The use of Heart Age was restricted to users aged between 20

and 80 years old with no history of heart disease. This group was selected because this was the population on which the original CVD algorithms were validated, and people with existing heart disease have different levels of risk of a future CVD event. A disproportionate number of users of the Heart Age tool reported to be 20 years old (9.59%, 323,547/3,374,769), possibly because this was the default setting. Therefore, to prevent artificial overrepresentation of 20 year olds, analysis was restricted to users aged 21-80 years.

To identify repeat users, prior to starting the self-assessment users were asked to indicate whether they had used the tool before. All return users were excluded from analysis to prevent duplication (n=143,682). Another 32,700 users whose TC, HDL-C, SBP, and body mass index (BMI) values were all missing or beyond defined valid ranges based on clinical judgment (ie, 80 mmHg≤ SBP ≤220 mmHg; 77 mg/dl≤ TC ≤423 mg/dl; 25 mg/dl≤ HDL ≤90 mg/dl; 15≤ BMI ≤45) were excluded from analysis as it was not possible to calculate a Heart Age for these users. As the tool was never promoted in Brazil and the number of users from Brazil was very low (n=102), it is likely that most entries were from a select group of internal users; therefore, all Brazilian users were excluded from the analysis. Finally, 130,647 users were excluded based on identical data of date of using the tool, BMI, TC, HDL, SPB, Heart Age, and 10-year CVD risk, which was mainly due to testing the functionality of the tool during the developmental phase. This resulted in a final dataset of 2,744,091 users. Among those users who were excluded due to reasons other than being 20 years old (n=307,137), the mean age was 40.98 (SD 14.29) years and 50.78% (155,969/307,137) were women.

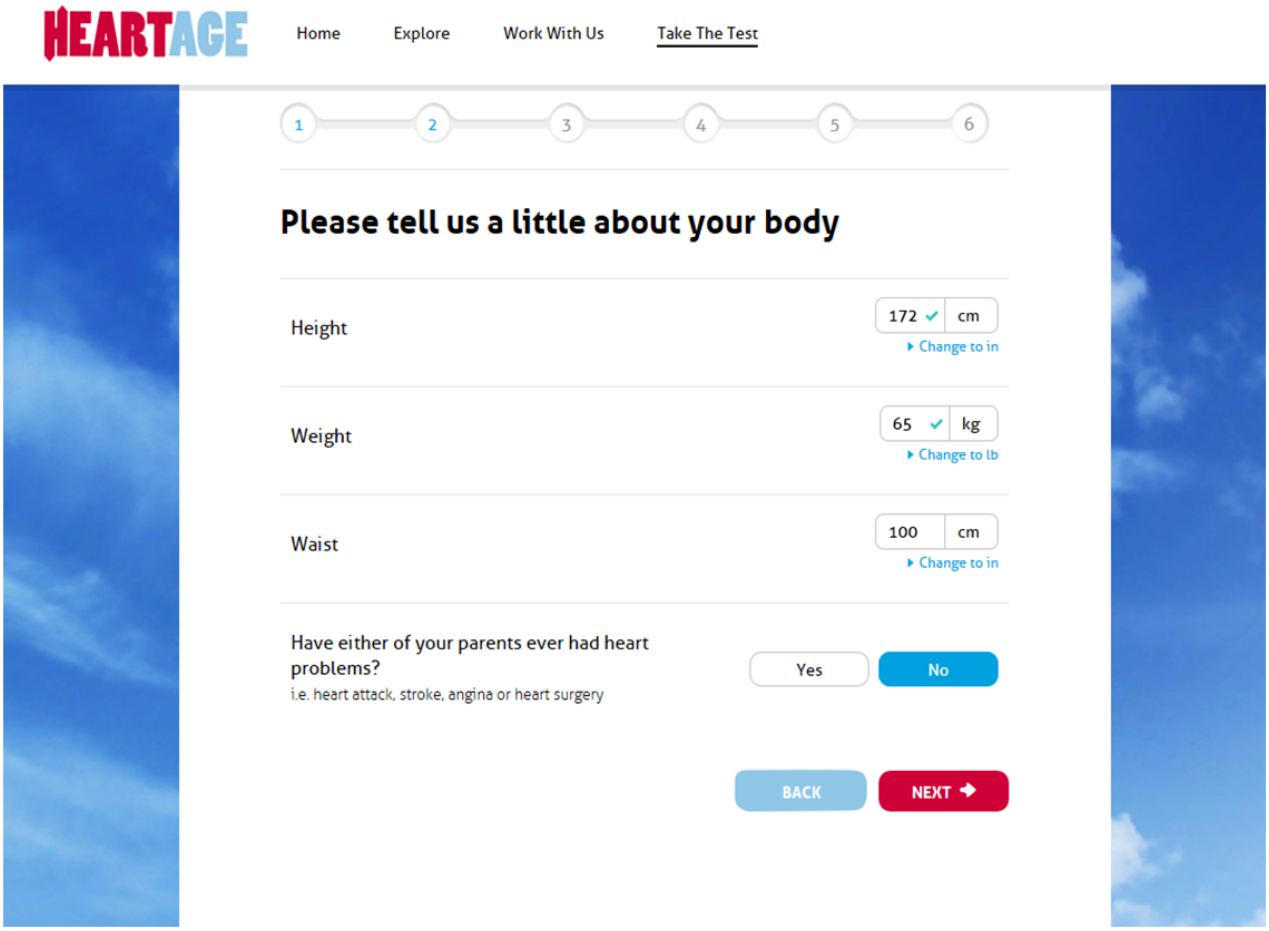
Materials

The Heart Age tool assessed CVD risk factors by self-report. Users completed a Web-based questionnaire that asked for information on demographic data (ie, age, gender, height, and weight; family history of CVD), physiological measures (ie, TC, HDL-C, and SBP values), other CVD risk factors (ie, smoking status and diabetes prevalence), and use of antihypertensive or cholesterol lowering medication. Further information was also obtained relating to CVD disease but was not relevant to the aims of the current study.

Questions were mostly presented as closed questions with locked answer options (eg, “yes/no” or “male/female”). Family history of CVD was assessed by asking “Have either of your parents ever had heart problems (ie, heart attack, stroke, angina, or heart surgery?)”. Smoking status and diabetes prevalence were determined by asking “Do you smoke?” and “Do you have diabetes?” Use of antihypertensive and blood cholesterol lowering medication was assessed with the questions “Are you taking or have you ever taken medication to lower your blood pressure?” and “Are you taking or have you ever taken medication to lower your cholesterol?” Questions on age, height, weight, and physiological measures were asked using an open answer format. Users could choose in which unit they wanted to enter their values for height (cm or inch) and weight (kg or lb); units for cholesterol were pre-determined and differed per country (mmol/l or mg/dl). Physiological measures were assessed by asking “Do you know your cholesterol level?” and “Do you know your blood pressure?” If users indicated “yes”, a sub-dialogue box opened and they could enter a value for TC and HDL-C or for SBP, respectively. The sub-dialogue boxes also contained explanations stating that “HDL cholesterol is sometimes known as ‘good’ cholesterol” and that “Systolic blood pressure is the higher number (eg, where blood pressure is 120/80, 120 is the systolic reading)”. All subject data were automatically captured in a database. See Figure 1 for a screenshot of the Heart Age tool. Screenshots of the complete user journey through the Heart Age tool are provided in Multimedia Appendix 1.

Depending on the availability of the physiological measures per individual, 10-year CVD risk and Heart Age could be calculated using different algorithms. There was one full algorithm when all physiological measures were known, and 5 alternatives for when one or more of the physiological measures were missing. The following alternative algorithms were available: (1) TC & HDL, (2) TC, SBP, BMI, (3) TC, BMI, (4) SBP, BMI, and (5) BMI. The CVD algorithms were developed using Framingham cohort data as sex-specific multivariable risk functions based on Cox proportional-hazards regression. The full CVD risk algorithm included data on age, gender, TC, HDL-C, SBP, antihypertensive medication use, smoking status, and diabetes [18]. Detailed information about the development and validation of the algorithms to calculate 10-year CVD risk can be found elsewhere [20,21]. Once calculated, CVD risk was translated into a Heart Age [20]. An individual’s Heart Age corresponds to the age of a person with the same predicted CVD risk but with all other risk factors considered as “normal”. The reference values for “normal” risk factors in the full Heart Age algorithm, were defined as not smoking, not diabetic, SBP=125 mmHg (130 mmHg if 60 or above), TC=180 mg/dl, and HDL-C=45 mg/dl [18]. For the alternative models where total cholesterol was present without HDL-C, the reference value for a person with “normal cholesterol” was increased to 200 mg/dl to increase the sensitivity of these models. To keep Heart Ages within a reasonable range, it was decided that Heart Age should be capped if it was 15 years lower or higher than chronological age, or if it fell below 18 or exceeded 80 years of age, in order to alert people of the need for change and medical advice without creating alarm. An overview of all CVD and Heart Age algorithms is provided Multimedia Appendix 2.

Figure 1. Screenshot of the Heart Age tool.



Data Analysis

Physiologic measures were analyzed both as continuous variables and dichotomous variables (normal vs high or normal vs low). Users were classified as having high TC if TC≥240 mg/dl, low HDL-C if HDL-C≤40 mg/dl [22], and high SBP was defined as SBP≥140 mmHg [23]. BMI was calculated from self-reported data on height and weight, and overweight and obesity were defined as BMI≥25 kg/m^2 and BMI≥30 kg/m^2 respectively [24]. All values beyond valid ranges were considered missing and treated as case-wise missing in analyses. To evaluate Heart Age data independent of chronological age, the difference between Heart Age and chronological age was calculated as "relative Heart Age".

To evaluate the effect of awareness of TC, HDL-C, and/or SBP values on Heart Age, a subsample was used consisting of users with valid values available on all physiological measures (173,397/2,744,091; 6.32% of the total population). For the users in this subsample, we calculated Heart Age based on the full algorithm and each of the five alternative algorithms. Repeated measures analysis of covariance (ANCOVA) were then conducted to test for significant differences in relative Heart Age between the full and alternative algorithms, adjusting for the effects of age, gender, and country. To evaluate if there were differential associations based on CVD risk category, age, and gender (ie, factors strongly associated with CVD risk), univariate ANCOVAs were conducted stratifying for CVD risk category, age, and gender. To account for multiple comparisons, *P* values in post-hoc tests were adjusted using a Bonferroni correction. Results are presented per gender, 10-year age categories, and 10-year CVD risk categories based on the Framingham risk score [18].

To evaluate the associations between subject characteristics and awareness of the three physiological cardiovascular risk factor values, logistic regression analyses were conducted based on the total sample. A separate regression model for each characteristic was conducted (ie, demographics, risk factors, CVD risk estimates, physiological measures (continuous and categorical, and awareness thereof), controlling for the effect of age, gender, and country.

Results

Characteristics of Users of the Web-Based Tool

Of the 2,744,091 Heart Age users included in the data analysis, the majority were from United Kingdom (31.19%, n=855,822), Germany (18.35%, n=503,502), the Netherlands (18.33%, n=503,063) Belgium (15.01%, n=411,948), and Finland (9.18%, n=251,952). The other countries contributed with 0.04% (n=995) to 2.33% (n=64,050) to the studied sample. The age distribution of Heart Age users covered the complete spectrum from 21-80 years of age. The number of users in their early 20s was high and peaked again between 40 and 50 years of age. Of the total sample of 2,744,091 users, 22.76% (n=624,639) were 20-29

years old, and another 23.99% (n=658,357) were 40-49 years old. After age 60, the number of Heart Age users decreased sharply, with 10.92% of users (n=299,665) aged 60-69 years and only 2.94% of users (n=80,753) aged 70-80 years.

The demographic characteristics and risk factor profile of the sample are presented in Table 1. On average, users had a Heart Age that exceeded their chronological age and reported multiple unhealthy CVD risk factors. The physiologic measures were normally distributed across the sample with mean values for TC and HDL in the normal range and for SBP and BMI above the normal range. Almost one third of the 2,744,091 users reported having a family history of CVD (32.84%, n=901,134) and prevalence of diabetes was low (3.39%, n=93,020). Heart Age was used more often by women (56.03%, n=1,537,490) than by men. Female users of the tool had on average a lower CVD risk and a lower relative Heart Age than men, a lower prevalence of modifiable risk factors, and more favorable physiological risk factor values, with the exception of TC. Interestingly, family history of CVD was more frequent among female users (35.70%, n=548,949) than male users (29.19%, n=352,185). Compared to the total sample, users with valid values for all physiologic measures and BMI were older. There were also more diabetic people, people with a family history of CVD, and fewer smokers in this subsample of users. Mean TC, HDL-C, and SBP values were similar between the total and subsample. Users in the subsample had fewer unhealthy CVD risk factors than the total sample. While the 10-year CVD risk of users with valid values for all physiologic measures and BMI was higher compared to that of the average user (due to older chronological age), the average user had a higher relative Heart Age than the users in the subsample.

Table 1. Characteristics of users of the Heart Age tool, per gender and for a subsample of users with complete valid values for TC, HDL-C, and SBP.

Characteristics	Total sample (N=2,744,091)	Men (n=1,206,601)	Women (n=1,537,490)	Subsample (n=173,397)
Demographics				
Age in yrs, mean (SD)	42.9 (14.0)	42.9 (14.3)	42.9 (13.8)	52.5 (12.0)
Male gender, n (%)	1,206,601 (43.97)	1,206,601 (100.00)	0 (0.00)	85,107 (49.08)
BMI (m/kg^2), mean (SD)[a]	25.9 (4.8)	26.3 (4.3)	25.6 (5.2)	26.0 (4.3)
CVD risk estimate, mean (SD)				
Heart Age, yrs[b]	46.9 (16.3)	47.8 (16.2)	46.2 (16.4)	53.6 (15.3)
Heart Age minus age, yrs[b]	4.0 (6.4)	4.9 (5.9)	3.3 (6.7)	1.2 (8.5)
10-yr CVD risk, %	7.15 (9.03)	10.30 (11.26)	4.69 (5.66)	9.90 (9.59)
Total unhealthy CVD risk factors, n	2.9 (1.4)	3.1 (1.4)	2.8 (1.4)	1.9 (1.4)
Risk factor prevalence, n (%)				
Current smoker	641,338 (23.37)	314,452 (26.06)	326,886 (21.26)	20,253 (11.68)
Overweight, BMI≥25[a]	1,391,691 (50.74)	693,071 (57.46)	698,620 (45.46)	91,721 (52.90)
Obese, BMI≥30[a]	475,715 (17.34)	201,322 (16.69)	274,393 (17.86)	26,525 (15.29)
Diabetic	93,020 (3.39)	49,577 (4.11)	43,443 (2.83)	12,448 (7.18)
High TC (≥240 mg/dL)[c]	87,257 (14.11)	35,624 (13.00)	51,633 (15.00)	23,432 (13.51)
Low HDL-C (≤40 mg/dL)[d]	31,067 (16.25)	20,616 (22.07)	10,451 (10.68)	27,881 (16.08)
High SBP (≥140 mmHg)[e]	116,836 (7.97)	60,506 (9.75)	56,330 (6.66)	13,716 (7.91)
Family history of CVD	901,134 (32.84)	352,185 (29.19)	548,949 (35.70)	76,919 (44.36)
Physiologic measures, mean (SD)				
TC (mg/dL)[c]	195.9 (47.0)	193.5 (44.0)	197.7 (46.6)	195.5 (44.6)
HDL-C (mg/dL)[d]	58.5 (15.6)	61.6 (16.4)	62.5 (15.1)	58.6 (15.6)
SBP (mmHg)[e]	124.6 (14.4)	127.1 (13.3)	122.4 (14.6)	126.1 (13.0)

[a]Users with valid BMI values (n=2,742,818 in total sample; 1,206,092 men and 1,536,726 women).

[b]Heart Age was calculated from different algorithms depending on available valid values of TC, HDL-C, and SBP.

[c]Users with valid TC values (n=618,210 in total sample; 274,102 men and 344,108 women).

[d]Users with valid HDL-C values (n=191,219 in total sample; 93,380 men and 97,839 women).

[e]Users with valid SBP values (n=1,466,836 in total sample; 620,684 men and 846,152 women).

Awareness of Physiological Risk Factors

The majority of Heart Age users were not aware of their TC values (77.47%; n=2,125,881) or HDL-C values (93.03%; n=2,552,872). Awareness of SBP values was relatively the highest, with 53.45% (n=1,466,836) of the users reporting valid values. Of all users, 6.37% (n=174,670) reported valid values for all three physiological risk factors.

Estimated Impact of Awareness of Physiological Risk Factors on Heart Age (in Subsample With Complete CVD Risk Factor Values)

Relative Heart Age, as calculated according to the five alternative algorithms, was compared to the relative Heart Age calculated with the full algorithm. This was done among the subsample of users with complete valid values for all physiologic measures and BMI to enhance comparability across algorithms. Alternative algorithms generally overestimated Heart Age, resulting in a significantly higher relative Heart Age than the full algorithm (P<.001). Alternative algorithms that were based on two physiological measures were closer to the full algorithm than alternative algorithms with only one, or no, physiological measure. The mean differences in relative Heart Age between the alternative and full models, adjusted for age, gender, and country, was 2.1 years (SD 0.1) for the alternative algorithm #1 (TC & HDL-C), 2.2 years (SD 0.1) for alternative algorithm #2 (TC, SBP, & BMI), 3.0 years (SD 0.1) for alternative algorithm #4 (SBP & BMI), and 3.9 years (SD 0.1) for alternative algorithm #3 (TC & BMI). The alternative algorithm #5, which included BMI as the only physiological measure, deviated most from the full algorithm, with on average 4.5 (SD 0.1) years higher than the full algorithm.

ANCOVAs comparing the difference in relative Heart Age between the full and all alternative algorithms revealed significant effects of age, gender, and CVD risk category (P<.001). These analyses were also restricted to the subsample of users with complete valid values for all physiologic measures and BMI. Figure 2 shows the adjusted differences in relative Heart Age between the alternative and full algorithms per gender and then per 10-year age categories. For all alternative algorithms, the overestimation of the relative Heart Age increased with age. For women, alternative algorithms overestimated relative Heart Age significantly more than for men, except for alternative algorithm #4, for which no gender effect was seen (P=.84). As shown in Figure 3, for all alternative algorithms, the overestimation of the relative Heart Age lessened with increasing CVD risk.

Figure 2. Mean difference in relative Heart Age (i.e. the difference between users' Heart Age and chronological age) between the full and alternative algorithms per gender and age category. The analysis was restricted to a sub-sample of users with complete valid values for total cholesterol (TC), HDL-cholesterol (HDL-C), systolic blood pressure (SBP) and BMI (n=173,397), and adjusted for age, gender and country where appropriate. Means are estimated marginal means. * significantly different, p < 0.001.

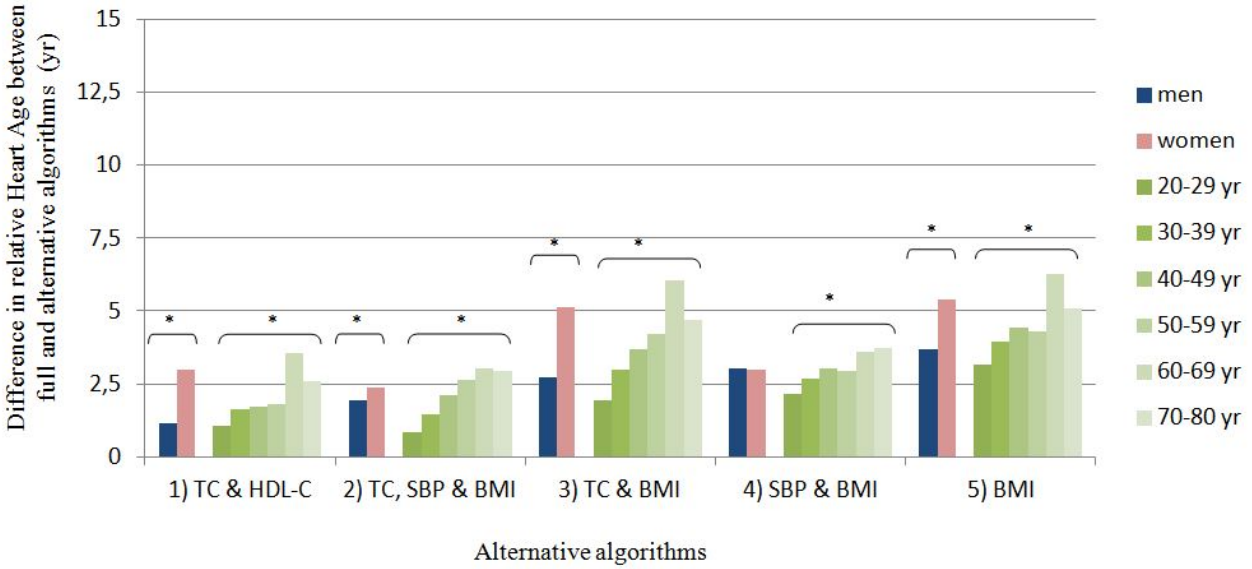
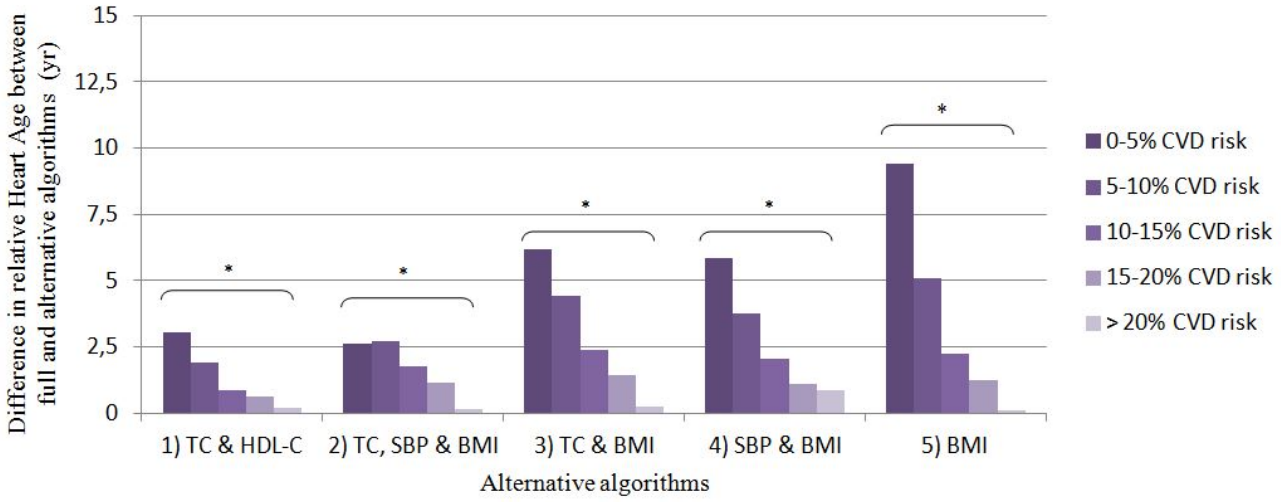


Figure 3. Mean difference in relative Heart Age (i.e. the difference between users' Heart Age and chronological age) between the full and alternative algorithms per category of absolute 10 year CVD risk based on Framingham risk score [18]. The analysis was restricted to a sub-sample of users with complete valid values for total cholesterol (TC), HDL-cholesterol (HDL-C), systolic blood pressure (SBP) and BMI (n=173,397), and adjusted for age, gender and country where appropriate. Means are estimated marginal means. * significantly different, $p < 0.001$.



Association Between Subject Characteristics and Awareness of Physiological Risk Factor Values

Adjusted odds ratios for the relations between users' characteristics and their awareness of TC, HDL-C, and SBP are presented in Table 2. Relatively, age, total number of risk factors, and awareness of other physiological risk factors were the strongest predictors. Among men, awareness of HDL-C values was 26% (95% CI 25-27) higher, while awareness of SBP values was 20% (95% CI 79-80) lower, compared to women. For all three physiological risk factors, an increase in awareness was associated with increasing age (ie, 5-7% per additional year). Diabetes and family history of CVD were consistently associated with higher awareness of physiological risk factor values. Diabetics were 1.47-1.74 times more likely to know their values than non-diabetics, and awareness among users with a family history of CVD was 22-43% higher compared to users without a family history of CVD. Users with overweight or obesity were significantly more aware of their SBP values (OR 1.13, 95% CI 1.12-1.14 and OR 1.20, 95% CI 1.19-1.21) compared to normal weight users, but significantly less aware of their HDL-C values (OR 0.89, 95% CI .89-.91 and OR 0.91, 95% CI .73-.75).

Awareness of the three physiological risk factors decreased with increasing number of unhealthy CVD risk factors (by 41-47% per additional risk factor) and with increasing Heart Age (by 3-8% per additional Heart Age year). Among current smokers, awareness of TC, HDL-C, and SBP was 29-48% lower compared to non-smokers. Users with high SBP values were significantly less likely to be aware of their TC and HDL-C values compared to those with normal SBP, with OR 0.94 (95% CI 0.93-0.95) and OR 0.46 (95% CI 0.34-0.63) respectively. Using the exact value of a physiological measure as predicting variable, instead of classification into unhealthy or normal values, generally did not lead to a better fit of the logistic regression models. Awareness of the three physiological risk factor values was strongly interrelated (ie, users who knew one value were very likely to also know the other values). This was particularly the case for awareness of TC and HDL-C values; users who were aware of their TC were over 4000-fold more likely to also know their HDL-C and vice versa.

Table 2. Adjusted logistic regression examining the relations between users' characteristics and their awareness of TC, HDL-C, and SBP values (N=2,744,091)[a].

	Awareness TC[b]		Awareness HDL-C[b]		Awareness SBP[b]	
Characteristics	OR (95% CI)	R^2	OR (95% CI)	R^2	OR (95% CI)	R^2
Demographics						
Age, yrs	1.07 (1.07-1.07)	.164	1.06 (1.06-1.06)	.085	1.05 (1.05-1.05)	.123
Gender, male	1.08 (1.08-1.09)	.000	1.26 (1.25-1.27)	.002	0.80 (0.79-0.80)	.003
CVD risk estimate						
Heart Age, yrs[c]	0.97 (0.96-0.97)	.011	0.92 (0.92-0.92)	.042	0.97 (0.97-0.97)	.009
10-yr CVD risk, %	0.96 (0.96-0.97)	.011	0.94 (0.94-0.94)	.020	1.02 (1.02-1.02)	.002
Total unhealthy CVD risk factors, n	0.54 (0.54-0.54)	.110	0.59 (0.59-0.60)	.065	0.53 (0.53-0.53)	.141
Risk factor prevalence						
Current smoker	0.67 (0.66-0.68)	.004	0.52 (0.52-0.53)	.007	0.71 (0.71-0.72)	.005
Overweight, BMI≥25[d]	1.05 (1.04-1.06)	.000	0.90(0.89-0.91)	.001	1.13 (1.12-1.14)	.002
Obese, BMI≥30[d]	1.03 (1.02-1.04)	.000	0.74 (0.73-0.75)	.001	1.20 (1.19-1.21)	.002
Diabetic	1.74 (1.71-1.77)	.003	1.48 (1.45-1.51)	.001	1.47 (1.46-1.50)	.001
High TC (≥240 mg/dl)[e]	—	—	1.07 (1.06-1.09)	.000	0.67 (0.66-0.68)	.004
Low HDL-C (≤40 mg/dl)[f]	0.46 (0.34-0.63)	.007	—	—	1.05 (1.01-1.10)	.000
High SBP (≥140 mmHg)[g]	0.94 (0.93-0.95)	.000	0.79 (0.77-0.80)	.001	—	—
Family history of CVD	1.43 (1.42-1.44)	.006	1.26 (1.24-1.27)	.001	1.22 (1.22-1.23)	.002
Awareness of physiological measures						
Awareness TC values	—	—	4364.73 (3,797.62-5,016.52)	.379	4.68 (4.65-4.72)	.069
Awareness HDL-C values	4,125.41 (3,589.46-4,741.37)	.208	—	—	6.24 (6.13-3.34)	.027
Awareness SBP values	4.85 (4.81-4.89)	.082	6.18 (6.08-6.29)	.056	—	—

[a]All logistic regressions were adjusted for age, country, and gender where appropriate. All ORs were significant (P<.001) with the exception of low HDL-C and awareness of SBP (P=.024).

[b]R^2 shows the Nagelkerke R^2 unique variance attributable to this variable after controlling for age, gender, and country where appropriate.

[c]Heart Age was calculated from different algorithms depending on available valid values of TC, HDL-C, and SBP.

[d]Users with valid BMI values (n=2,742,818), reference category is normal weight (BMI<25).

[e]Users with valid TC values (n=618,210).

[f]Users with valid HDL-C values (n=191,219).

[g]Users with valid SBP values (n=1,466,836).

Discussion

Principal Results and Implications

The potential positive impact of Web-based health assessment tools on public health is highly dependent upon whether the tools reach the right audience and the quality of health assessment that is provided. This study evaluated whether a health assessment tool communicating CVD risk was reaching an audience with elevated risk, evaluated the impact of having correct input information (eg, awareness of key physiological risk factors), and examined whether certain subgroups were less likely to provide complete input information. Overall, the Heart Age tool was shown to reach the intended audience of people who are at risk of developing a CVD event, and it was demonstrated that missing information on physiological risk factors resulted in overestimation of (relative) Heart Age. Specific subgroups of people were indeed at higher risk for not providing complete risk factor information. The Heart Age tool reached on average users who were at moderate risk (men) and low risk (women); few users had a high risk (>20% CVD risk). Of note, the average Heart Age user had multiple CVD risk factors. Prevalence of smoking, overweight, and obesity among Heart Age users was within the usual ranges commonly observed in the studied countries [25-27]. However, rates of high TC, high SBP, and diabetes were lower [15,26,28,29]. This may be explained by the underrepresentation of people aged 60 years and older and to high rates of undiagnosed diabetes commonly seen in these populations [30-32]. Information bias may also play a role in the relatively low prevalence of high TC and SBP, as TC and SBP levels of those who were aware of their values may systematically differ from those who were not aware [33].

Others previously reported that specific population groups that are more likely to have high CVD risk, like older people, and those with preventable health problems or chronic conditions, and people of lower socioeconomic status are generally less likely to search for health information, possibly due to limited Internet access and usage [3,10,34,35] and relatively low eHealth literacy (ie, ability to seek, find, understand, and apply health information from electronic sources) [36]. The accessibility of Web-based health applications to these population groups at particularly high (absolute) CVD risk should therefore still be improved (eg, through increasing Internet access in public places, developing health apps for mobile use, using user friendly language, and targeted promotion).

The Heart Age tool was designed to reach people at risk of developing CVD, where interventions are still viable. The absolute risk of the current users was lower than expected, suggesting that the tool may be reaching a large proportion of "healthy" users. However, on average, these people had multiple unhealthy CVD risk factors, and their heart age was on average 4 years older than their real age. This suggests that their risk factors are elevated for their age, and they have high lifetime risk of developing CVD if nothing is done to improve their lifestyle. Therefore, the audience reached by Heart Age *was* the intended audience—a group that is a good target for disease prevention. This is of particular interest for younger people where a low absolute 10-year CVD risk can be falsely reassuring [37-40], especially in the presence of multiple raised risk factors. Heart Age was shown to be more likely perceived as a wake-up call among younger (30-45 year old) smokers and/or obese users [41]. Previous studies also showed that Heart Age motivated people to make healthy behavior changes or had an emotional impact [7,42,43], and led to greater reductions in cholesterol, blood pressure, and CVD risk than when a percentage for 10-year CVD risk was provided [42]. This highlights that health self-assessment tools, and particularly Heart Age, can be useful for health promotion at the population level.

A large proportion of Heart Age users did not know their TC, HDL-C, or SBP values and therefore received an overestimation of (relative) Heart Age that may cause inappropriate distress among users of the tool. This highlights the importance of complete risk factor information for reliable risk communication of Web-based health self-assessment tools. Therefore, Web-based health assessment tools should mention the potential impact of not knowing your risk factors and encourage users who have not recently been tested to get tested. As expected, the more physiological risk factors that were known, the more accurately (relative) Heart Age was predicted. Because awareness of TC alone did not add much to the accuracy of predicted Heart Age, cholesterol tests should generally measure both TC and HDL-C concentrations. Furthermore, the use of proxy-measures to estimate lacking physiological risk factor values in Web-based health applications should be explored and validated. These could include questions on whether people are diagnosed with hypertension or hypercholesterolemia or have been prescribed cholesterol-lowering medication. The Heart Age tool is continuously improved based on current learnings, and its algorithms are fine-tuned to increase its accuracy; this may include space for new measurements that improve the performance of the tool but also important variations that may be necessary for different countries or ethnic groups. The finding that overestimation of Heart Age diminished with increasing CVD risk is presumably attributable to the capping of Heart Age. Precautionary measures to restrict extreme outcomes of Web-based health assessment tools may generally be useful to limit large inaccuracies in predicted disease risk. Withholding people who have incomplete risk factor information from accessing Web-based self-assessment tools may be a missed opportunity as the Heart Age tool has been shown to motivate women to get their cholesterol tested [41].

Smokers and people with several unhealthy CVD risk factors were particularly at risk of not being aware of their (other) physiological CVD risk factors. Therefore, if health self-assessment tools are targeted at these groups, special attention should be paid to a thorough inventory of all relevant risk factors, a careful communication of the estimated disease risk within some confidence limits, and compelling encouragement to get to know one's risk factor values. As awareness of physiological CVD risk factors by itself has been shown to motivate people to take preventive actions [44,45], awareness of risk factor values should generally be promoted among smokers and those with other known CVD risk factors. This also aligns with the general policy in CVD risk prevention, which encourages looking at the combination of all CVD risk factors [22,46]. The finding that awareness of physiological risk factors decreased with increasing Heart Age needs to be considered with care as this may be partly explained by Heart Age being overestimated when risk factor information is missing. The number of total CVD risk factors explained 6.5-14.1% of the variation in awareness of risk factors, indicating that other characteristics beyond risk factors may be more important predictors of awareness of TC, HDL-C, and SBP.

Strengths and Limitations

The current study was based on a large sample size of almost three million Heart Age tool users, providing confidence in findings. Moreover, the sample was drawn from 13 countries, enhancing the generalizability of results. However, there are some limitations that should be taken into account when interpreting the results. First, the data are self-reported and not verified for accuracy through objective measures; this is particularly an issue for physiological data, which are not always accurately recalled or may not be from a recent measurement. Second, the impact of awareness of physiological risk factors on Heart Age was estimated based on a relatively small and healthy subsample of Heart Age users; therefore, this finding may not be generalizable. Thus, while from this study the impact of missing risk factor information is shown to be a concern, future research is required to investigate the effect of awareness of physiological CVD risk factors on Heart Age in a sample where self-reported information on awareness of physiological risk factor values is combined with objective measurements of all Heart Age tool parameters.

Conclusions

Heart Age is a Web-based CVD risk communication tool, which was found on average to reach people with multiple elevated CVD risk factors and a Heart Age higher than their real age. This highlights that Web-based self -assessment health tools can be a useful means to interact with people who are at risk of developing disease, and where interventions are still viable. Missing information in the self-assessment health tools was shown to result in inaccurate self-health assessments. This demonstrates the importance of complete input information into health self-assessment tools and the need to encourage people who do not know their values to get tested. Specific subgroups of people were found to be particularly at risk for not having complete input information required for health self-assessment tools, and as such, more care needs to be taken if health self-assessment tools are being deployed to these groups. Extra efforts need to be made to raise awareness in these subgroups.

Acknowledgments

All authors were involved in the design of the study and drafting or revising the manuscript for intellectual content, and all approved the final version to be published.

Conflicts of Interest

NN, MC, and RSN are employees of Unilever R&D. The Heart Age tool was developed by Unilever R&D and has registered trademarks and a granted patent.

Multimedia Appendix 1

Screenshots of the Heart Age tool.

[PPTX File, 1MB - jmir_v16i9e215_app1.pptx]

Multimedia Appendix 2

Full and alternative algorithms used to calculate 10-year CVD risk and Heart Age.

[PDF File (Adobe PDF File), 1MB - jmir_v16i9e215_app2.pdf]

References

1. 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]
2. 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]
3. Fox S. Health Topic: 80% of internet users look for health information online. 2011 Feb 01. URL: http://www.pewinternet.org/~/media//Files/Reports/2011/PIP_Health_Topics.pdf [accessed 2013-10-28] [WebCite Cache ID 6KhiH3ciA]
4. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. Int J Med Inform 2006 Jan;75(1):29-41. [doi: 10.1016/j.ijmedinf.2005.07.029] [Medline: 16129659]
5. Lira MT, Kunstmann S, Caballero E, Guarda E, Villarroel L, Molina JC. Cardiovascular prevention and attitude of people towards behavior changes: state of the art. Rev Med Chil 2006 Feb;134(2):223-230 [FREE Full text] [Medline: 16554932]
6. Logan RA. Evaluating consumer informatics: learning from health campaign research. Stud Health Technol Inform 2004;107(Pt 2):1147-1151. [Medline: 15360992]
7. Ayres K, Conner M, Prestwich A, Hurling R, Cobain M, Lawton R, et al. Exploring the question-behaviour effect: randomized controlled trial of motivational and question-behaviour interventions. Br J Health Psychol 2013 Feb;18(1):31-44. [doi: 10.1111/j.2044-8287.2012.02075.x] [Medline: 22519696]
8. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
9. Parekh S, Vandelanotte C, King D, Boyle FM. Improving diet, physical activity and other lifestyle behaviours using computer-tailored advice in general practice: a randomised controlled trial. Int J Behav Nutr Phys Act 2012;9:108 [FREE Full text] [doi: 10.1186/1479-5868-9-108] [Medline: 22963354]
10. Bessell TL, Silagy CA, Anderson JN, Hiller JE, Sansom LN. Prevalence of South Australia's online health seekers. Aust N Z J Public Health 2002 Apr;26(2):170-173. [Medline: 12054338]
11. Levy JA, Strombeck R. Health benefits and risks of the Internet. J Med Syst 2002 Dec;26(6):495-510. [Medline: 12385532]
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. Robinson TN, Patrick K, Eng TR, Gustafson D. An evidence-based approach to interactive health communication: a challenge to medicine in the information age. Science Panel on Interactive Communication and Health. JAMA 1998 Oct 14;280(14):1264-1269. [Medline: 9786378]
14. Nijland N, van Gemert-Pijnen J, Boer H, Steehouder MF, Seydel ER. Evaluation of internet-based technology for supporting self-care: problems encountered by patients and caregivers when using self-care applications. J Med Internet Res 2008 May;10(2):e13 [FREE Full text] [doi: 10.2196/jmir.957] [Medline: 18487137]
15. Tolonen H, Keil U, Ferrario M, Evans A, MONICA Project WHO. Prevalence, awareness and treatment of hypercholesterolaemia in 32 populations: results from the WHO MONICA Project. Int J Epidemiol 2005 Feb;34(1):181-192 [FREE Full text] [doi: 10.1093/ije/dyh298] [Medline: 15333620]
16. Angus J, Evans S, Lapum J, Rukholm E, St Onge R, Nolan R, et al. "Sneaky disease": the body and health knowledge for people at risk for coronary heart disease in Ontario, Canada. Soc Sci Med 2005 May;60(9):2117-2128. [doi: 10.1016/j.socscimed.2004.08.069] [Medline: 15743659]
17. van Steenkiste B, van der Weijden T, Stoffers HE, Kester AD, Timmermans DR, Grol R. Improving cardiovascular risk management: a randomized, controlled trial on the effect of a decision support tool for patients and physicians. Eur J Cardiovasc Prev Rehabil 2007 Feb;14(1):44-50. [doi: 10.1097/01.hjr.0000239475.71805.1e] [Medline: 17301626]
18. 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]
19. Heart UK, World Heart Foundation, Unilever. Heart Age. 2013. Heart Age Calculator URL: http://www.heartage.me/ [accessed 2013-10-28] [WebCite Cache ID 6Khrp6jdw]
20. Cobain M. Heart Age Assessment. USA patent US 8,409,104 B2. 2013 Apr 2. Conopco Inc URL: http://www.google.com/patents/US8529456 [accessed 2014-09-14] [WebCite Cache ID 6SaA1YGlv]
21. 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]
22. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation,Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002 Dec 17;106(25):3143-3421 [FREE Full text] [Medline: 12485966]
23. Mancia G, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M, et al. 2013 ESH/ESC guidelines for the management of arterial hypertension: the Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). Eur Heart J 2013 Jul;34(28):2159-2219 [FREE Full text] [doi: 10.1093/eurheartj/eht151] [Medline: 23771844]
24. WHO. Obesity and overweight. 2013. URL: http://www.who.int/mediacentre/factsheets/fs311/en/index.html [accessed 2013-10-28] [WebCite Cache ID 6KhdL6Ic1]
25. Organisation for Economic Co-operation and Development (OECD). Health at a Glance: Europe 2012.: OECD Publishing; 2012. URL: http://www.oecd.org/els/health-systems/HealthAtAGlanceEurope2012.pdf [accessed 2014-09-14] [WebCite Cache ID 6SaB9KXcI]
26. WHO. Global Status Report on Noncommunicable Diseases 2010. Geneva: World Health Organization; 2011.
27. WHO. WHO Report on the global tobacco epidemic. 2013. URL: http://www.who.int/tobacco/global_report/2013/appendix_x_graph_10_1.pdf [accessed 2014-03-20] [WebCite Cache ID 6ODBEX9sq]
28. Australian Bureau of Statistics. 2013. Health Survey: Updated Results URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/80E00969554B94C7CA257B8200179637?opendocument [accessed 2013-10-28] [WebCite Cache ID 6KhfL7yMu]
29. International Diabetes Federation. Diabetes Atlas, third edition. Brussels, Belgium: International Diabetes Federation; 2006.
30. Jousilahti P, Vartiainen E, Tuomilehto J, Puska P. Sex, age, cardiovascular risk factors, and coronary heart disease: a prospective follow-up study of 14 786 middle-aged men and women in Finland. Circulation 1999 Mar 9;99(9):1165-1172 [FREE Full text] [Medline: 10069784]
31. Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract 2011 Dec;94(3):311-321. [doi: 10.1016/j.diabres.2011.10.029] [Medline: 22079683]
32. International Diabetes Federation. IDF Diabetes Atlas, 5th edition. 2013. URL: http://www.idf.org/diabetesatlas [accessed 2013-10-28] [WebCite Cache ID 6Khnlh8yl]
33. Deutekom M, Aziz Y, van Dis I, Stronks K, Bossuyt PM. [The Dutch National Colesterol Test: participants mainly healthy]. Ned Tijdschr Geneeskd 2008 Nov 1;152(44):2425-2429. [Medline: 19055144]
34. 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]
35. 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]

36. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
37. Williams CL, Hayman LL, Daniels SR, Robinson TN, Steinberger J, Paridon S, et al. Cardiovascular health in childhood: A statement for health professionals from the Committee on Atherosclerosis, Hypertension, and Obesity in the Young (AHOY) of the Council on Cardiovascular Disease in the Young, American Heart Association. Circulation 2002 Jul 2;106(1):143-160 [FREE Full text] [Medline: 12093785]
38. Marteau TM, Lerman C. Genetic risk and behavioural change. BMJ 2001 Apr 28;322(7293):1056-1059 [FREE Full text] [Medline: 11325776]
39. Tymstra T, Bieleman B. The psychosocial impact of mass screening for cardiovascular risk factors. Fam Pract 1987 Dec;4(4):287-290. [Medline: 3692036]
40. Lloyd-Jones DM, Dyer AR, Wang R, Daviglus ML, Greenland P. Risk factor burden in middle age and lifetime risks for cardiovascular and non-cardiovascular death (Chicago Heart Association Detection Project in Industry). Am J Cardiol 2007 Feb 15;99(4):535-540 [FREE Full text] [doi: 10.1016/j.amjcard.2006.09.099] [Medline: 17293199]
41. Soureti A, Murray P, Cobain M, van Mechelen W, Hurling R. Web-based risk communication and planning in an obese population: exploratory study. J Med Internet Res 2011 Nov;13(4):e100 [FREE Full text] [doi: 10.2196/jmir.1579] [Medline: 22126827]
42. Lopez-Gonzalez AA, Aguilo A, Frontera M, Bennasar-Veny M, Campos I, Vicente-Herrero T, et al. Effectiveness of the Heart Age tool for improving modifiable cardiovascular risk factors in a Southern European population: a randomized trial. Eur J Prev Cardiol 2014;Feb 3:1-2.
43. 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]
44. Aubin M, Godin G, Vézina L, Maziade J, Desharnais R. Hypercholesterolemia screening. Does knowledge of blood cholesterol level affect dietary fat intake? Can Fam Physician 1998 Jun;44:1289-1297 [FREE Full text] [Medline: 9640523]
45. Mooney LA, Franks AM. Impact of health screening and education on knowledge of coronary heart disease risk factors. J Am Pharm Assoc (2003) 2011 Nov;51(6):713-718. [doi: 10.1331/JAPhA.2011.10127] [Medline: 22068192]
46. Perk J, De Backer G, Gohlke H, Graham I, Reiner Z, Verschuren M, European Association for Cardiovascular Prevention & Rehabilitation (EACPR), ESC Committee for Practice Guidelines (CPG). European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). Eur Heart J 2012 Jul;33(13):1635-1701 [FREE Full text] [doi: 10.1093/eurheartj/ehs092] [Medline: 22555213]

Abbreviations

ANCOVA: analysis of covariance
BMI: body mass index
CVD: cardiovascular disease
HDL-C: high density lipoprotein cholesterol
TC: total cholesterol
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 28.11.13; peer-reviewed by C Bonner, R Peters, A Lerman; comments to author 14.03.14; revised version received 24.04.14; accepted 10.07.14; published 26.09.14.

Please cite as:
Neufingerl N, Cobain MR, Newson RS
Web-Based Self-Assessment Health Tools: Who Are the Users and What Is the Impact of Missing Input Information?
J Med Internet Res 2014;16(9):e215
URL: http://www.jmir.org/2014/9/e215/
doi:10.2196/jmir.3146
PMID:25261155

Original Paper

Estimation of Geographic Variation in Human Papillomavirus Vaccine Uptake in Men and Women: An Online Survey Using Facebook Recruitment

Erik J Nelson[1], MPH, PhD; John Hughes[2], PhD; J Michael Oakes[1], PhD; James S Pankow[1], MPH, PhD; Shalini L Kulasingam[1], PhD

[1]School of Public Health, Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, United States
[2]School of Public Health, Division of Biostatistics, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Erik J Nelson, MPH, PhD
School of Public Health
Division of Epidemiology and Community Health
University of Minnesota
1300 South Second Street
Suite 300
Minneapolis, MN, 55454
United States
Phone: 1 1 612 624 1818
Fax: 1 1 612 624 0315
Email: nels6712@umn.edu

Abstract

Background: Federally funded surveys of human papillomavirus (HPV) vaccine uptake are important for pinpointing geographically based health disparities. Although national and state level data are available, local (ie, county and postal code level) data are not due to small sample sizes, confidentiality concerns, and cost. Local level HPV vaccine uptake data may be feasible to obtain by targeting specific geographic areas through social media advertising and recruitment strategies, in combination with online surveys.

Objective: Our goal was to use Facebook-based recruitment and online surveys to estimate local variation in HPV vaccine uptake among young men and women in Minnesota.

Methods: From November 2012 to January 2013, men and women were recruited via a targeted Facebook advertisement campaign to complete an online survey about HPV vaccination practices. The Facebook advertisements were targeted to recruit men and women by location (25 mile radius of Minneapolis, Minnesota, United States), age (18-30 years), and language (English).

Results: Of the 2079 men and women who responded to the Facebook advertisements and visited the study website, 1003 (48.2%) enrolled in the study and completed the survey. The average advertising cost per completed survey was US $1.36. Among those who reported their postal code, 90.6% (881/972) of the participants lived within the previously defined geographic study area. Receipt of 1 dose or more of HPV vaccine was reported by 65.6% women (351/535), and 13.0% (45/347) of men. These results differ from previously reported Minnesota state level estimates (53.8% for young women and 20.8% for young men) and from national estimates (34.5% for women and 2.3% for men).

Conclusions: This study shows that recruiting a representative sample of young men and women based on county and postal code location to complete a survey on HPV vaccination uptake via the Internet is a cost-effective and feasible strategy. This study also highlights the need for local estimates to assess the variation in HPV vaccine uptake, as these estimates differ considerably from those obtained using survey data that are aggregated to the state or federal level.

(J Med Internet Res 2014;16(9):e198) doi:10.2196/jmir.3506

KEYWORDS

online recruitment; social media; Facebook; local estimation; geographic variability; human papillomavirus; HPV

Introduction

Human papillomavirus (HPV) is the most common sexually transmitted infection in the United States [1] and is the necessary cause of cervical cancer [2]. HPV infections are also associated with other cancers (eg, anogenital and oropharyngeal) as well as genital warts [3,4]. In total, it is estimated that 5.2% of cancers in men and women worldwide are attributable to HPV [5].

Two vaccinations against HPV infection are currently licensed in the United States. The vaccinations were originally licensed for use in girls, but as of October 2011, the Advisory Committee on Immunization Practices extended their recommendation of the quadrivalent vaccine to include both boys and girls aged 11 or 12 years old [6,7]. However, vaccine uptake has been far lower than expected, with only about half of eligible young women receiving at least one dose of the vaccine [8]. Initiation of the HPV vaccine series has been shown to be higher among minority adolescent girls; however, completion of the three-dose series is substantially lower among black and Hispanic adolescent girls compared to white adolescent girls [9]. Although male vaccination data are very limited (due to a later date of approval of the HPV vaccine for boys), racial and income differences in terms of vaccine series initiation and completion have also been observed among adolescent boys [10].

Previous research on HPV vaccine coverage has used publicly available data from five national health surveys (National Survey of Family Growth, National Immunization Survey [NIS]-Teen, National Health and Nutrition Examination Survey, National Health Interview Survey [NHIS], and the Behavioral Risk Factor Surveillance System) [11-15]. These surveys are designed to gather information on a variety of health topics and ask only a few questions regarding HPV vaccination. However, none of these surveys address cervical cancer screening practices and potential barriers to screening or HPV vaccine receipt. In addition, due to the small number of responses in many geographic areas, local data from these surveys are routinely suppressed and aggregated to state boundaries in order to protect the confidentiality of survey respondents, which means that variations at a local level (ie, between counties or postal codes) cannot be adequately assessed. Further, these surveys have, to date, primarily surveyed adolescent girls; HPV vaccination practice data of adolescent boys are limited [8].

The Internet provides a unique point of contact to reach young adults for health research. Several studies have demonstrated that Internet-based research can be used to elicit high response rates at a fraction of the cost of traditional recruitment methods [16-18]. In addition, it has been shown that when compared to in-person interviews, Internet-based surveys have the potential to reach more respondents, include otherwise inaccessible populations, and reduce bias in responses as respondents may be willing to report more sensitive information online compared to in-person interviews [19-24]. A number of studies have also shown that recruitment via Facebook (the leading social media site with more than one billion active users worldwide) can be used to enroll representative samples of the general population [16,25-30]. This combination of reach, utility, and reduced cost indicates that social media networks can be a cost-effective medium for research.

The objective of this study was to estimate HPV vaccination practices among a local population of young adult men and women in the United States using an Internet-based recruitment strategy.

Methods

Participants

Men and women from Minnesota were surveyed about their HPV vaccination practices via the Internet from November 21, 2012, through January 31, 2013. Participants were English-speaking, aged 18-30 years, had a Facebook account, and resided in the greater Twin Cities Metropolitan Area (ie, within 25 miles of downtown Minneapolis, MN). This age range was used to target men and women who were eligible to receive the HPV vaccine, participate in cervical cancer screening (women), and able to provide informed consent. The Twin Cities Metropolitan Area was selected due to the variation of HPV-related cancer incidence rates exhibited in this area during the past 15 years, the high concentration of colleges and universities, and the large population of 18-30 year olds residing in this area [31]. The University of Minnesota Institutional Review Board approved this study.

Facebook Recruitment Campaign

Participants were recruited online via Facebook advertisements (Figure 1). Tailored advertisements were used to target Facebook users who had profiles that matched the study inclusion criteria. The advertisement criteria were adjusted as needed to target specific postal codes with fewer responses in order to achieve a balanced sample of participants by postal code. Facebook uses an advertisement algorithm that automatically selects the best advertisement to display based on its performance and the advertiser's bid [32]; 14 unique advertisements were created and approved by Facebook. For this study, multiple advertisements were submitted for auction simultaneously to create a continuous recruitment window in the event that a particular advertisement performed poorly. Bidding prices and advertisement availability (advertisements can be paused and released at the discretion of the advertiser) were monitored daily and adjusted as necessary until the intended number of completed questionnaires was obtained. The bidding price for advertisements ranged from US $0.75 to US $2.75, with a maximum daily budget of US $50. When a Facebook user clicked on the study advertisement, they were automatically redirected to the secure study website and invited to complete a questionnaire regarding HPV vaccination practices.

The Facebook Ads Manager was used to track the total number of impressions (each time an advertisement was displayed), the number of times an ad was clicked, the average cost-per-click, and the number of people reached (ie, the number of Facebook users that had an opportunity to view one of the study advertisements). Google Analytics software was used to tabulate the total number of visits, the unique visits, the average duration of visits, and the bounce rate (the percentage of visitors that

visit a website and leave the site without further browsing) of the study website.

Figure 1. Examples of Facebook advertisements.

Study Procedures

Participants who clicked on a Facebook advertisement were directed to a secured study website and were provided with information regarding the purpose of our study. Participants provided informed consent by clicking on a button that directed them to the study questionnaire. After providing consent, study participants were immediately asked to self-report their age and state of residence. Participants who did not meet the age criteria or who reported that they did not live in Minnesota were considered ineligible and were disqualified from answering the remainder of the questionnaire. Study participants who met the eligibility criteria were also asked to self-report their gender, postal code of their home address, race/ethnicity, highest level of education attained, attendance at religious services, political preferences, sexual orientation, their awareness of HPV, and whether or not they had received the HPV vaccine. Conditional upon participants' responses, skip logic patterns (ie, participants skip over survey questions that, based on their answers to other questions, do not need to be filled out) were implemented in order to ask applicable follow-up questions regarding the number of shots received, the vaccine type (quadrivalent/bivalent), and reason(s) for not having received the vaccination, as well as future vaccination intentions. Female participants were also asked a series of adaptive questions about past cervical cancer screening. The survey questions regarding HPV vaccination and cancer screening that we used in this study were questions used in the five national surveys mentioned above, in order to facilitate comparisons between studies. Participants were not required to answer every question and could exit the survey at any time. Computer Internet Protocol (IP) addresses were tracked, and multiple entries from the same IP address were not accepted. Survey responses that contained repeated email addresses across multiple survey attempts (n=86)

were not accepted. Additionally, 8 surveys that were only partially completed (ie, the participant withdrew) were not included in the analyses. The survey was anonymous and was administered using the online survey assessment tool SurveyMonkey. Eligible respondents who provided informed consent and completed the online survey were emailed an electronic gift card in the amount of US $20 for Target.

Results

Of the 2079 men and women who were recruited via Facebook and visited the study website, 1003 (48.24%) enrolled in the study and completed the survey. Targeted advertising within Facebook based on geographic and age criteria limited the number of ineligible participants (4.4% of all survey attempts) who attempted to access the survey. In total, 86 survey attempts (7.5% of all survey attempts) were identified as duplicate surveys, indicating that an individual attempted to complete the survey more than once (Figure 2). Facebook advertising and recruitment resulted in an average cost of US $1.36 per completed survey. In addition, 90.6% (881/972) of study participants who self-reported their postal code were located within the recruitment target area (ie, located within a 25-mile radius of downtown Minneapolis, Minnesota; Figure 3).

The recruitment target area for this study was a 25-mile radius from downtown Minneapolis, Minnesota. Of the 972 participants who reported their postal code, 881 (90.6%) lived within the recruitment study area.

A total of 1003 participants (557 women and 446 men) completed the online survey. Characteristics of the study population are presented in Table 1. With respect to race/ethnicity, the study population was broadly similar to that of 18-34 year-olds in the greater Minneapolis-St. Paul Metropolitan Area based on US Census data. However, due to the inclusion and exclusion criteria, the study population was more educated than the general population of 18-34 year-olds in the Minneapolis-St. Paul Metropolitan Area. In all, 44.2% of respondents (396/896) who knew of the HPV vaccine had been vaccinated against HPV (ie, received ≥1 dose of HPV vaccine), with 65.6% of women (351/535) having been vaccinated with ≥1 dose of HPV vaccine compared to 13.0% of men (45/347). Completion of the HPV vaccine series (ie, receipt of all 3 doses) was reported by 74.9% of women (263/351) and 22.2% of men (10/45) who had ever received an HPV vaccine (Table 2). Among the 351 women who had received ≥1 dose of HPV vaccine, 265 (75.5%) had also received at least one Pap smear in their lifetime. Of the 479 unvaccinated men and women, 403 (84.1%) were not interested or were unsure about receiving the vaccine in the future.

Table 1. Selected study participant characteristics compared to US Census estimates for Minneapolis and St. Paul, Minnesota.[a]

	Study participants			Census data
	Men, n=446	Women, n=557	Total, N=1003	Minneapolis/St. Paul, %
Mean age, years	23	23	23	18 to 34
Race, n (%)				
White	384 (86.3)	457 (82.3)	841 (84.10)	79.3
Black	17 (3.8)	33 (5.9)	50 (5.00)	9.1
Asian	30 (6.7)	30 (5.4)	60 (6.00)	8.1
American Indian or Alaska native	2 (0.4)	7 (1.3)	9 (0.90)	0.8
Native Hawaiian or Pacific Islander	1 (0.2)	3 (0.5)	4 (0.40)	0.03
Other	11 (2.5)	25 (4.5)	36 (3.60)	2.6
Ethnicity, n (%)				
Hispanic	15 (3.4)	19 (3.4)	34 (3.42)	5.2
Non-Hispanic	427 (96.6)	533 (96.6)	960 (96.58)	94.8
Education, n (%)				
<High school	2 (0.4)	0 (0.0)	2 (0.20)	1.9
Some high school	6 (1.3)	8 (1.4)	14 (1.40)	8.0
High school graduate	36 (8.1)	36 (6.5)	72 (7.19)	21.9
Some college/tech. school	190 (42.7)	209 (37.6)	399 (39.86)	36.3
College graduate	167 (37.5)	237 (42.6)	404 (40.36)	25.1
Graduate school	44 (9.9)	66 (11.9)	110 (10.99)	6.8

[a]Data are 5-year estimates for 18-34 year-olds in the Minneapolis-St. Paul Metropolitan Area as described in the 2006-2010 American Community Survey of the United States Census Bureau.

Table 2. Selected survey responses regarding vaccination against human papillomavirus.

Survey question	Men (n=446) n (%)	Women (n=557) n (%)	Total (N=1003) n (%)
Ever heard of HPV [a]			
Yes	409 (93.0)	536 (96.8)	945 (95.07)
No	31 (7.0)	18 (3.2)	49 (4.93)
Ever heard of HPV vaccine			
Yes	361 (82.4)	535 (96.6)	896 (90.32)
No	77 (17.6)	19 (3.4)	96 (9.68)
Ever had an HPV vaccination among those who had heard of the HPV vaccine			
Yes	45 (13.0)	351 (66.5)	396 (45.26)
No	302 (87.0)	177 (33.5)	479 (54.74)
Number of HPV shots received			
1 shot	11 (24.4)	31 (8.8)	42 (10.61)
2 shots	14 (3.9)	38 (10.8)	52 (13.13)
3 shots (complete vaccine series)	10 (22.2)	263 (74.9)	273 (68.94)
Don’t know	10 (22.2)	19 (5.4)	29 (7.32)
Likelihood of HPV vaccine receipt in the next 12 months among those not vaccinated [a]			
Very likely	7 (2.3)	13 (7.3)	20 (4.18)
Somewhat likely	26 (8.6)	30 (16.9)	56 (11.69)
Not too likely	75 (24.8)	47 (26.6)	122 (25.47)
Not likely at all	173 (57.3)	84 (47.5)	257 (53.65)
Not sure/don’t know	21 (7.0)	3 (1.7)	24 (5.01)
Reason stated for not receiving the HPV vaccine in the next 12 months [a]			
Not needed or necessary	140 (52.4)	40 (31.3)	180 (45.57)
Not sexually active	33 (12.4)	23 (18.0)	56 (14.18)
Knowledge[b]	25 (9.4)	10 (7.8)	35 (8.86)
Safety concerns/side effects	9 (3.4)	22 (17.2)	31 (7.85)
Costs	13 (4.9)	14 (10.9)	27 (6.84)
Already have HPV	19 (7.1)	5 (3.9)	24 (6.08)
Monogamous	8 (3.0)	6 (4.7)	14 (3.54)
Other[c]	6 (2.2)	6 (4.7)	12 (3.04)
Not for men	11 (4.1)	0 (0.0)	11 (2.78)
Too old	3 (1.1)	2 (1.6)	5 (1.27)

[a]Responses presented are for the 479 individuals who reported not having been vaccinated against HPV.

[b]Don’t know about HPV or HPV vaccine.

[c]Responses included “fear of needles”, “too busy/no time”, “don’t use vaccines”, or “already sexually active”.

Figure 2. Recruitment summary flowchart.

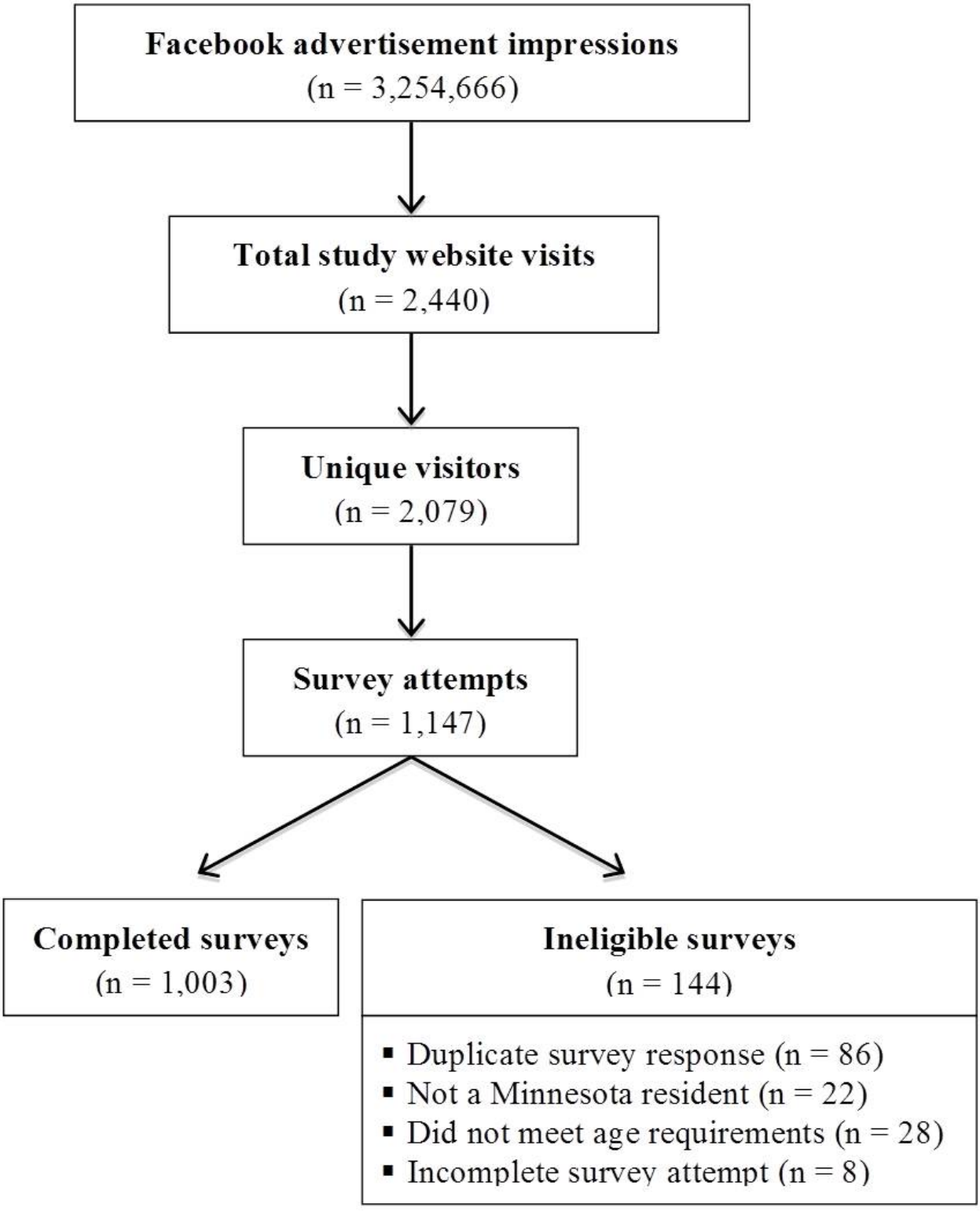
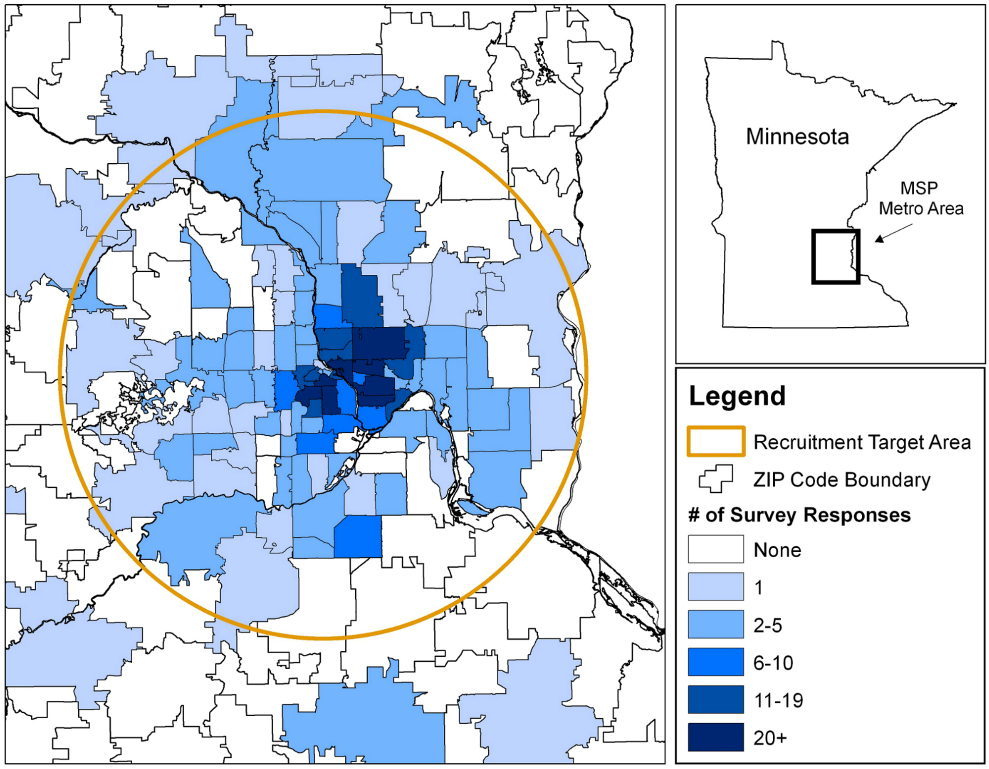


Figure 3. Map of the recruitment target area and the number of completed surveys by ZIP code.



Discussion

Principal Findings

In this study, we found that recruiting a locally representative sample of young adults via the Internet to participate in a survey about HPV vaccination was cost-effective and efficient. Approximately half of the 2079 individuals that clicked on an advertisement and visited our study website participated and completed our survey at an estimated advertising cost of US $1.36 per enrolled participant. Consistent with other studies, this study found that using the Internet, and in particular social media sites such as Facebook, is successful for recruiting and engaging young adults and hard-to-reach populations for health research [16-18]. This method of recruitment is particularly noteworthy given declining response rates from traditional recruitment techniques such as random digit dialing or mailed surveys [33-35]. In addition to higher participation rates, the targeted advertising features embedded within social media websites drastically reduce costs associated with identifying and reaching a large pool of eligible participants [25,26,28]. The targeted advertising used in this study also allowed us to collect data within an accelerated timeline (eg, pilot testing of a specific intervention) from a specific geographic location.

Notably, the characteristics of our study population were similar to those of the source population. An estimated 90% of Internet users aged 18-29 years in the United States access social media sites (71% accessed Facebook) in 2013; thus, this finding is likely attributable to the wide reach of social media recruitment [36]. However, our study population was more educated than the general population in the Minneapolis-St. Paul Metropolitan Area, which may be due to the large number of colleges and universities in this area. It cannot be ruled out that people with lower education were less likely to access Facebook and view the advertisements, although other studies have shown that lower income and less educated participants are as likely to participate in Internet-based research studies as those with higher incomes and higher levels of education [26,37,38].

In this study, we were also able to collect detailed HPV vaccination data, including participation in screening (for women) and potential barriers to receiving these services among a representative sample of men and women in a defined local geographic area. National surveys including the Behavioral Risk Factor Surveillance System, the NHIS, and the NIS-Teen do not simultaneously assess these factors within the same respondents in their populations. Additionally, these (and other) national surveys aggregate or suppress responses due to participant identification concerns and consequentially local variation and patterns may be obscured. However, HPV vaccine policies, availability, costs, financial assistance, and education materials vary widely across states or even more defined geographic regions [39]. As a result, variation at state and national levels may not reflect the variation in HPV vaccine uptake occurring at a local level.

Of note, the proportion of all adults in this study who had been vaccinated against HPV (ie, received at least one dose of an HPV vaccine) was 45.3% (66.5% for women and 13.0% for men). These estimates are much higher than the HPV vaccine coverage estimates from the 2012 NHIS for women (34.5%) and men (2.3%) aged 19-26 years (Table 3) [40]. Although the results for women are more similar to those obtained from the NIS-Teen for girls (53.8%), the estimate for men is much lower than the NIS-Teen estimate for boys (20.8%) aged 13-17 years who received at least one dose of HPV vaccine in 2012 [41]. Although the differences in the observed rates may be partially

explained by the sampling frame, response rates, or the small number of eligible respondents who received the HPV vaccine question series in the national surveys, the estimates of HPV vaccine uptake are noticeably different from the current study.

Table 3. HPV vaccine coverage estimates for men and women in the United States from three surveys.

	HPV vaccine coverage (≥1 dose)			
	Men		Women	
Survey	%	95% CI[a]	%	95% CI
SMASH[b]	13.0	9.4-16.5	65.6	61.6-69.6
NHIS	2.3	1.6-3.4	34.5	31.7-37.3
NIS–Teen	20.8	19.3-22.3	53.8	51.9-55.7

[a]95% confidence interval.

[b]Data are from the Survey of Minnesotans About Screening and HPV, 2013.

Limitations

Limitations include the fact that the survey responses were self-reported by persons over the Internet and may be subject to under or overreporting. However, other Internet-based studies have shown increased self-disclosure and reporting with online surveys, which may reduce potential response biases (eg, interviewer bias or social desirability) [19,21]. Additionally, there was no failproof method to ensure that survey responses were unique, and there remains a small probability that some participants responded more than once. We also cannot be certain that those that saw the Facebook advertisements were the same people who completed the survey. The 10% of respondents who were not located within the targeted geographical area may be due to the sharing of the study website with friends, or due to outdated user profiles (ie, Facebook thinks a user lives within the study area and displays the ad although the user has since relocated outside of the target area but has not updated their account info), or because the advertisement algorithm was misspecified by Facebook.

Conclusions

To our knowledge, this is the first study to estimate local level vaccination uptake among young men in the United States. Understanding the local variation and patterns of HPV vaccination of young men could serve to identify areas where HPV infection-related health disparities may continue if neglected. In particular, the online survey also allowed us to collect data on sexual orientation, which in turn would allow us to understand whether men who have sex with men, who are at high risk of HPV-related anal cancer, are receiving the vaccine and to also determine whether reductions in the overall risk of HPV infection will affect transmission to females [42,43].

The results from this study suggest that more detailed and local assessments of HPV vaccine uptake are necessary as estimates vary greatly from national surveys. In addition, recruiting young adults via the Internet is efficient, cost-effective, and can produce a representative sample of the target population. Future work is needed to understand the pattern of HPV vaccine uptake at local levels in order to identify areas that may be best served by vaccine programs.

Acknowledgments

This study was supported by the JB Hawley Student Research Award from the University of Minnesota School of Public Health and by the Minnesota Medical Foundation through Grant 4120-9227-12.

Authors' Contributions

EN and SK conceived and designed the study. EN conducted the data collection and drafted the manuscript. JH, JP, and JMO assisted in the survey design, supervised the statistical analysis, and assisted in reviewing/revising the manuscript. SK provided contributions to the concept and analytical approach for the article and oversaw the analysis, interpretation, and reviewing/revising of the manuscript.

Conflicts of Interest

None declared.

References

1. Weinstock H, Berman S, Cates W. Sexually transmitted diseases among American youth: incidence and prevalence estimates, 2000. Perspect Sex Reprod Health 2004;36(1):6-10. [doi: 10.1363/psrh.36.6.04] [Medline: 14982671]
2. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999 Sep;189(1):12-19. [doi: 10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F] [Medline: 10451482]

3. Centers for Disease Control and Prevention (CDC). Human papillomavirus-associated cancers - United States, 2004-2008. MMWR Morb Mortal Wkly Rep 2012 Apr 20;61:258-261 [FREE Full text] [Medline: 22513527]
4. Saraiya M, Ahmed F, White M, Lawson H, Unger ER, Eheman C. Toward using National Cancer Surveillance data for preventing and controlling cervical and other human papillomavirus-associated cancers in the US. Cancer 2008 Nov 15;113(10 Suppl):2837-2840 [FREE Full text] [doi: 10.1002/cncr.23753] [Medline: 18980202]
5. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006 Jun 15;118(12):3030-3044. [doi: 10.1002/ijc.21731] [Medline: 16404738]
6. Centers for Disease Control and Prevention (CDC). FDA licensure of bivalent human papillomavirus vaccine (HPV2, Cervarix) for use in females and updated HPV vaccination recommendations from the Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Rep 2010 May 28;59(20):626-629 [FREE Full text] [Medline: 20508593]
7. Advisory Committee on Immunization Practices. Recommended adult immunization schedule: United States, 2012. Ann Intern Med 2012 Feb 7;156(3):211-217. [doi: 10.7326/0003-4819-156-3-201202070-00388] [Medline: 22298576]
8. Centers for Disease Control and Prevention (CDC). National and state vaccination coverage among adolescents aged 13-17 years--United States, 2012. MMWR Morb Mortal Wkly Rep 2013 Aug 30;62(34):685-693 [FREE Full text] [Medline: 23985496]
9. Niccolai LM, Mehta NR, Hadler JL. Racial/Ethnic and poverty disparities in human papillomavirus vaccination completion. Am J Prev Med 2011 Oct;41(4):428-433. [doi: 10.1016/j.amepre.2011.06.032] [Medline: 21961471]
10. Jeudin P, Liveright E, Del Carmen MG, Perkins RB. Race, ethnicity, and income factors impacting human papillomavirus vaccination rates. Clin Ther 2014 Jan 1;36(1):24-37. [doi: 10.1016/j.clinthera.2013.11.001] [Medline: 24417783]
11. Laz TH, Rahman M, Berenson AB. Human papillomavirus vaccine uptake among 18- to 26-year-old women in the United States: National Health Interview Survey, 2010. Cancer 2013 Apr 1;119(7):1386-1392 [FREE Full text] [doi: 10.1002/cncr.27894] [Medline: 23508594]
12. Liddon NC, Leichliter JS, Markowitz LE. Human papillomavirus vaccine and sexual behavior among adolescent and young women. Am J Prev Med 2012 Jan;42(1):44-52. [doi: 10.1016/j.amepre.2011.09.024] [Medline: 22176845]
13. Pruitt SL, Schootman M. Geographic disparity, area poverty, and human papillomavirus vaccination. Am J Prev Med 2010 May;38(5):525-533 [FREE Full text] [doi: 10.1016/j.amepre.2010.01.018] [Medline: 20409501]
14. Wei F, Moore PC, Green AL. Geographic variability in human papillomavirus vaccination among U.S. young women. Am J Prev Med 2013 Feb;44(2):154-157 [FREE Full text] [doi: 10.1016/j.amepre.2012.09.061] [Medline: 23332332]
15. Anhang Price R, Tiro JA, Saraiya M, Meissner H, Breen N. Use of human papillomavirus vaccines among young adult women in the United States: an analysis of the 2008 National Health Interview Survey. Cancer 2011 Dec 15;117(24):5560-5568 [FREE Full text] [doi: 10.1002/cncr.26244] [Medline: 21732336]
16. 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]
17. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
18. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. J Med Internet Res 2008;10(5):e50 [FREE Full text] [doi: 10.2196/jmir.1001] [Medline: 19073542]
19. Cantrell MA, Lupinacci P. Methodological issues in online data collection. J Adv Nurs 2007 Dec;60(5):544-549. [doi: 10.1111/j.1365-2648.2007.04448.x] [Medline: 17973718]
20. McCabe SE, Boyd CJ, Couper MP, Crawford S, D'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. J Stud Alcohol 2002 Nov;63(6):755-761. [Medline: 12529076]
21. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]
22. Schonlau M, Zapert K, Simon L, Sanstad K, Marcus S, Adams J, et al. A Comparison Between Responses From a Propensity-Weighted Web Survey and an Identical RDD Survey. Soc Sci Comput Rev 2004 Feb 01;22(1):128-138. [doi: 10.1177/0894439303256551]
23. Schonlau M. Are "webographic" or attitudinal questions useful for adjusting estimates from Web surveys using propensity scoring? Surv Res Methods 2007;1(3):155-163 [FREE Full text]
24. Thériault N, Bi P, Hiller JE, Nor M. Use of web 2.0 to recruit Australian gay men to an online HIV/AIDS survey. J Med Internet Res 2012;14(6):e149 [FREE Full text] [doi: 10.2196/jmir.1819] [Medline: 23128646]
25. Im EO, Chee W. Recruitment of research participants through the Internet. Comput Inform Nurs 2004;22(5):289-297. [Medline: 15520599]
26. Jones L, Saksvig BI, Grieser M, Young DR. Recruiting adolescent girls into a follow-up study: benefits of using a social networking website. Contemp Clin Trials 2012 Mar;33(2):268-272 [FREE Full text] [doi: 10.1016/j.cct.2011.10.011] [Medline: 22101207]
27. Facebook. Community Info. 2014. URL: http://newsroom.fb.com/company-info/ [accessed 2014-04-30] [WebCite Cache ID 6PEIMHeSc]

28. Batterham PJ. Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun;23(2):184-191. [doi: 10.1002/mpr.1421] [Medline: 24615785]
29. Frandsen M, Walters J, Ferguson SG. Exploring the viability of using online social media advertising as a recruitment method for smoking cessation clinical trials. Nicotine Tob Res 2014 Feb;16(2):247-251. [doi: 10.1093/ntr/ntt157] [Medline: 24127266]
30. Ramo DE, Liu H, Prochaska JJ. Reliability and validity of young adults' anonymous online reports of marijuana use and thoughts about use. Psychol Addict Behav 2012 Dec;26(4):801-811 [FREE Full text] [doi: 10.1037/a0026201] [Medline: 22082344]
31. Nelson EJ, Hughes J, Kulasingam SL. Spatial patterns of human papillomavirus-associated cancers within the state of Minnesota, 1998-2007. Spat Spatiotemporal Epidemiol 2014 Jun;9:13-21. [doi: 10.1016/j.sste.2014.02.003] [Medline: 24889990]
32. Facebook. Campaign Cost and Budgeting. 2014. URL: https://www.facebook.com/unsupportedbrowser [accessed 2014-05-01] [WebCite Cache ID 6PFgbB9oc]
33. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
34. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 1;163(3):197-203 [FREE Full text] [doi: 10.1093/aje/kwj036] [Medline: 16339049]
35. 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]
36. Pew Research Internet Project. Social Networking Fact Sheet. 2014. URL: http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/ [accessed 2014-04-23] [WebCite Cache ID 6P3HgIRa0]
37. Lohse B. Facebook is an effective strategy to recruit low-income women to online nutrition education. J Nutr Educ Behav 2013;45(1):69-76. [doi: 10.1016/j.jneb.2012.06.006] [Medline: 23305805]
38. Kesse-Guyot E, Andreeva V, Castetbon K, Vernay M, Touvier M, Méjean C, et al. Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Santé study. J Med Internet Res 2013;15(9):e205 [FREE Full text] [doi: 10.2196/jmir.2488] [Medline: 24036068]
39. Katz ML, Reiter PL, Kluhsman BC, Kennedy S, Dwyer S, Schoenberg N, et al. Human papillomavirus (HPV) vaccine availability, recommendations, cost, and policies among health departments in seven Appalachian states. Vaccine 2009 May 21;27(24):3195-3200 [FREE Full text] [doi: 10.1016/j.vaccine.2009.03.042] [Medline: 19446191]
40. Williams WW, Lu PJ, O'Halloran A, Bridges CB, Pilishvili T, Hales CM, Centers for Disease Control and Prevention (CDC). Noninfluenza vaccination coverage among adults - United States, 2012. MMWR Morb Mortal Wkly Rep 2014 Feb 7;63(5):95-102 [FREE Full text] [Medline: 24500288]
41. Centers for Disease Control and Prevention (CDC). National and state vaccination coverage among adolescents aged 13-17 years--United States, 2012. MMWR Morb Mortal Wkly Rep 2013 Aug 30;62(34):685-693 [FREE Full text] [Medline: 23985496]
42. Palefsky JM. Human papillomavirus-related disease in men: not just a women's issue. J Adolesc Health 2010 Apr;46(4 Suppl):S12-S19 [FREE Full text] [doi: 10.1016/j.jadohealth.2010.01.010] [Medline: 20307839]
43. Watson M, Saraiya M, Ahmed F, Cardinez CJ, Reichman ME, Weir HK, et al. Using population-based cancer registry data to assess the burden of human papillomavirus-associated cancers in the United States: overview of methods. Cancer 2008 Nov 15;113(10 Suppl):2841-2854 [FREE Full text] [doi: 10.1002/cncr.23758] [Medline: 18980203]

Abbreviations

HPV: human papillomavirus
IP: Internet Protocol
NHIS: National Health Interview Survey
NIS: National Immunization Survey

Edited by G Eysenbach; submitted 01.05.14; peer-reviewed by M Frandsen, Y Jayasinghe; comments to author 16.07.14; revised version received 13.08.14; accepted 16.08.14; published 01.09.14.

Please cite as:
Nelson EJ, Hughes J, Oakes JM, Pankow JS, Kulasingam SL
Estimation of Geographic Variation in Human Papillomavirus Vaccine Uptake in Men and Women: An Online Survey Using Facebook Recruitment
J Med Internet Res 2014;16(9):e198
URL: http://www.jmir.org/2014/9/e198/
doi:10.2196/jmir.3506
PMID:25231937

Original Paper

Baby Boomers' Adoption of Consumer Health Technologies: Survey on Readiness and Barriers

Cynthia LeRouge[1*], PhD; Craig Van Slyke[2*], PhD; Deborah Seale[3*], PhD; Kevin Wright[4*], PhD

[1]College for Public Health and Social Justice, Department of Health Management & Policy, Saint Louis University, Saint Louis, MO, United States

[2]W.A. Franke College of Business, Northern Arizona University, Flagstaff, AZ, United States

[3]Health Informatics and Information Management, Saint Louis University, Saint Louis, MO, United States

[4]College of Humanities and Social Sciences, Communications Department, George Mason University, Fairfax, VA, United States

[*]all authors contributed equally

Corresponding Author:
Cynthia LeRouge, PhD
College for Public Health and Social Justice
Department of Health Management & Policy
Saint Louis University
3545 Lafayette Avenue
Room 365
Saint Louis, MO, 63104
United States
Phone: 1 314 977 8195
Fax: 1 314 977 1674
Email: lerouge@uw.edu

Abstract

Background: As they age, baby boomers (born 1946-1964) will have increasing medical needs and are likely to place large demand on health care resources. Consumer health technologies may help stem rising health care needs and costs by improving provider-to-patient communication, health monitoring, and information access and enabling self-care. Research has not explored the degree to which baby boomers are ready for, or are currently embracing, specific consumer health technologies This study explores how baby boomers' readiness to use various technologies for health purposes compares to other segments of the adult population.

Objective: The goals of the study are to (1) examine what technologies baby boomers are ready to use for health purposes, (2) investigate barriers to baby boomers' use of technology for health purposes, and (3) understand whether readiness for and barriers to baby boomers' use of consumer health technologies differ from those of other younger and older consumers.

Methods: Data were collected via a survey offered to a random sample of 3000 subscribers to a large pharmacy benefit management company. Respondents had the option to complete the survey online or by completing a paper-based version of the survey.

Results: Data from 469 respondents (response rate 15.63%) were analyzed, including 258 baby boomers (aged 46-64 years), 72 younger (aged 18-45 years), and 139 older (age >64 years) participants. Baby boomers were found to be similar to the younger age group, but significantly more likely than the older age group to be ready to use 5 technologies for health purposes (health information websites, email, automated call centers, medical video conferencing, and texting). Baby boomers were less ready than the younger age group to adopt podcasts, kiosks, smartphones, blogs, and wikis for health care purposes. However, baby boomers were more likely than older adults to use smartphones and podcasts for health care purposes. Specific adoption barriers vary according to the technology.

Conclusions: Baby boomers have commonalities with and distinctions from both younger and older adults in their readiness to adopt specific consumer health technologies and the barriers they experience to adoption. Baby boomers' nuances regarding readiness to adopt and the barriers associated with the various forms of consumer health technology should be taken into account by those interested in promoting consumer health technologies use among baby boomers when developing applications, choosing technologies, preparing users for use, and in promotional tactics.

(J Med Internet Res 2014;16(9):e200) doi:10.2196/jmir.3049

KEYWORDS

baby boomer; readiness; adoption; consumer health technology; man-machine systems; aging; health; human factors; design; user interfaces; personal computing

Introduction

Background

The United States health care system faces a coming "gray tsunami" as baby boomers (those born between 1946 and 1964) begin requiring more health care resources [1,2]. The United States is one of many countries around the world that are experiencing new demands on their health care systems due to a rise in the number of older adults and health conditions associated with aging [3-5]. According to a report by the American Hospital Association and First Consulting Group [6], baby boomers make up the largest segment of the population in the United States (approximately 78 million Americans based on the 2000 US Census) [7,8]. In 2011, the first members of the baby boom generation reached age 65. As baby boomers reach retirement age, many will have increasing medical needs and thus demand more health care resources than other segments of the population, particularly given the size of this age cohort and that their life expectancies are longer than many past generations [6]. Age is often associated with growing health problems and chronic disease [9]. Approximately 60% of baby boomers have already been diagnosed with at least 1 chronic medical condition. Arthritis, diabetes, heart disease, obesity, osteoporosis, and hypertension are common chronic conditions among baby boomers. These conditions require regular health care checkups, prescription medications, and significant dietary changes [6]. Given the size of the baby boomer cohort, responding to age-related health concerns will likely be a significant challenge for health care systems over the next decade and beyond [10].

The health care industry must be prepared to accommodate this growing segment of health care consumers. Similar to previous generations, many baby boomers will require extensive health services, training and reinforcement for medical self-management, as well as continued connection to clinicians and contact with their peers and caregivers as they grow older [11]. It is questionable whether the traditional US health care system can handle the level of health care demands imposed by the gray tsunami given the immense size of the baby boomer cohort compared to previous older generations [1]. If empowerment leads to an increase in self-care, patient empowerment efforts aimed at prevention and self-management of chronic disease could play a key role in relaxing some of the demands on the health care system [12]. Given current changes in the health care marketplace and the need to find more cost-efficient ways to manage health conditions, baby boomers will likely be the generation that leads the movement toward patient self-management of chronic disease [13].

Emerging health information and communication technologies bring the promise of transformations in the delivery of care, empowering patients to make more informed health care decisions, connecting patients directly to providers and other caregivers and personalizing services in response to patients' unique needs and preferences [13]. Information and communication technologies used by consumers for health purposes are increasingly allowing individuals to conveniently learn about, manage, and monitor their health via electronic devices. The use of consumer health technologies may help stem rising health care costs by improving provider-to-patient communication, health condition monitoring, and health information access by enabling self-care [14,15]. Baby boomers may differ from older cohorts in terms of their exposure to new technologies (e.g, being exposed to new technologies in the workplace). Because baby boomers will be the largest older adult cohort in history, it is important to assess the barriers to consumer health technologies within this population; consumer health technologies will likely continue to be a means of controlling health care costs in the face of growing demands on the health care system (due to its increased use by baby boomers as they begin to face health issues associated with aging).

But, are baby boomers currently embracing or even ready to use modern technologies to manage their health? Although baby boomers may have more technology access and experience than prior generations, it is not clear if this leads to greater readiness to adopt consumer health technologies. For example, researchers have found that although barriers exist for some technologies, older consumers are ready and able to adopt other technologies for health-related tasks [16-19]. Designing and developing consumer health technologies that effectively meet the unique requirements of baby boomers requires understanding which types of consumer health technologies baby boomer consumers are ready to adopt and what barriers exist for the specific technologies where readiness is low. Previous research has not adequately addressed these issues from the perspective of baby boomers.

The purpose of this study is to address the following research goals. First, we examine what consumer health technologies baby boomers are ready to use. Second, we investigate barriers to baby boomers' use of consumer health technologies. Finally, we seek to understand whether readiness for and barriers to baby boomers' use of consumer health technologies differ from those for other consumers (younger and older). Gaining a better understanding of these issues will help proponents of consumer health technology better understand how to build and promote systems that bring about benefits to baby boomer consumers. This is especially noteworthy because baby boomers may become a driving force behind the development of consumer health technologies given the demands on the health care system they are predicted to create.

Theoretical Framework

Consumer health technologies attempt to engage health consumers to interact with technology to promote healthy behaviors and informed decision making. The Agency for Healthcare Research and Quality notes consumer health information technology (IT) applications are a key topic and indicate that:

> *These [consumer health technology] applications have various purposes including assisting with self-management through reminders and educational prompts, delivering real-time data on a patient's health condition to both patients and providers, facilitating Web-based support groups, and compiling and storing personal health information in an easily accessible format...Moreover, consumer health IT applications that allow gathering and integrating data from various health care sources can serve as a comprehensive resource for patients and their providers. In addition to convenience, consumer health IT applications also can be important in emergency situations to provide critical health information to medical staff. [20]*

Most consumer health technologies are designed to change attitudes or behaviors and provide information. The effectiveness of consumer health technologies requires choosing a receptive audience and an appropriate technology [21-23]. Persuasive technology design can facilitate coaxing the user toward healthy action (motivating factor) and underscore the need to choose a receptive audience and befitting technology [22-24]. A model of the persuasive design process (drawn from demonstrated success in industries including health care) begins with defining the persuasion goal to match a receptive target audience with an appropriate technology (Figure 1 depicts this first stage of the persuasive design process; areas of focus for the current study are shaded) [24]. This type of alignment coincides with modern human-computer interaction design philosophy in which the needs, desires, and limitations of users are investigated and analyzed [25]. If various forms of consumer health technology are to be successful, we must have a fundamental understanding of which technology tools align with baby boomers' needs, desires, and limitations. Indeed, recent research focused on physical and psychological attributes as 1 of 4 types of patient barriers to eHealth opportunity (the other 3 types include the provision of eHealth opportunity, the support others that may have to use eHealth, and economic barriers) [26].

Unfortunately, the literature that specifically focuses on adoption of consumer health technologies and its use by baby boomers is limited. Some noted exceptions relate to home monitoring devices. For example, work by Mihailidis and colleagues [27] found a general willingness of current baby boomers to accept various forms of home monitoring technology (eg, personal emergency response systems, fall detection systems). However, it is important to note that the Mihailidis et al study was largely exploratory and it was limited by a small sample size [27]. In addition, a study of Australian baby boomers found baby boomers were generally open to use assistive technologies for a temporary period after hospital discharge [28]. In both studies, the home monitoring and assistive technologies studied were quite passive as compared to the interactive, multipurpose technologies frequently used in consumer health informatics, such as the Internet and smartphones. We know of no study that specifically explores various common interactive technology tools for consumer health informatics use by baby boomers. It is from this need that we provide our first research question:

Research question 1: What technologies are baby boomers ready to use to promote healthy behaviors and informed decision making?

As part of that foundational understanding of which consumer health technology tools align with baby boomers' needs, desires, and limitations, we must recognize that baby boomers are different from younger and older consumers in many ways, 2 of which are particularly important to note. First, baby boomers are not the "digital native" youths, who have known these technologies their entire lives, nor are they like their elders, most of whom had little exposure to interactive information technologies in their work lives. Many baby boomers experienced the transition to more of an information technology workforce. There might be some social bias that those who are not digital natives may not be as open to consumer health technologies. However, recent studies indicate that this is not the case. A recent study found baby boomers and older adults were generally open to home monitoring devices and did not have strong preferences with respect to the types and locations of the technology [27]. But, do these similarities extend to more interactive technologies? Unlike their predecessors, many baby boomers are comfortable with interactive technology [29,30]. But, unlike their successors, many baby boomers do not naturally turn to technology as their first choice when communicating, seeking information, or looking for task support for health needs [11]. For example, according to the Pew Research Center's 2013 update on smartphone ownership [31], only 39% of those surveyed aged 55-64 years owned a smartphone. This percentage was even lower (18%) for those 65 years or older. Younger consumers reported much higher ownership percentages [31]. For example, 81% of those aged 25-34 years reported owning a smartphone [31].

The second important difference relates to baby boomers' increased expectations concerning health care services [6]. Because aging baby boomers have higher levels of education, more disposable income in terms of being in their peak earning years compared to younger and older age cohorts (although overall income and savings may be affected by numerous variables, including life circumstances and the recent economic downturn), and are more active than previous generations, baby boomers are naturally more focused on health care services that ensure their long-term mobility and independence [32]. The higher expectations of baby boomers are reflected in increased demands for innovative and personalized health care services that eliminate barriers to treatment and provide timely and accurate health-related information and services. In addition, many baby boomers are now caring for elderly parents, while trying to maintain an active lifestyle, which further increases their health information and service needs [33]. To further explore how baby boomers' readiness for various consumer health technology tools compares to other segments of the adult population, we introduce the following research question:

Research question 2: How do the technology tools that baby boomers report they are ready to use for health purposes differ from the technology tools younger adults and older adults report that they are ready to use for health purposes?

Prior research indicates that baby boomers may face a number of barriers in adopting consumer health technologies [34,35]. Innovation diffusion theory [36] provides a useful theoretical framework for investigating the "internal" barriers particular to the personal decision process of adoptions (in contrast to external factors such as provision of eHealth opportunity and economic barriers). Rogers [36] posits that potential innovation adopters go through a 5-stage process when deciding whether to adopt and use an innovation: knowledge, persuasion, decision, implementation, and confirmation. The first 2 stages, knowledge and persuasion, are of primary interest to this study and require additional explication (note that the discussion of the stages is based on Rogers [36] unless otherwise indicated).

In the knowledge stage, the potential adopter becomes aware that an innovation exists. This awareness is followed by the adopter forming an understanding of how the innovation functions. Two important sets of information related to the adopter are important to this stage. First, prior conditions, such as prior experience with similar innovations, problems faced by the adopter, and social system norms, impact the knowledge stage. For example, problems faced by the adopter that may be met by the innovation's use are likely to influence how the adopter frames knowledge about the innovation. Prior experience with similar innovations may likewise influence knowledge of the innovation. For example, a consumer who has experience with a smartphone will build knowledge of a tablet computer differently than a consumer without such prior experience. Characteristics of the adopter are also important in the knowledge stage. This is particularly important for our research given that age is an essential individual characteristic related to innovation adoption [37].

In the persuasion stage, the potential adopter forms attitudes related to the innovation and its use. (It is important to note that Rogers defines *persuasion* as the formation of attitudes rather than a change agent's activities to influence those attitudes.) Perceptions regarding the innovation's attribute use are the building blocks of the adopter's innovation-related attitudes. Generally, the adopter is concerned with advantages and disadvantages of the innovation, given the adopter's particular situation.

During the decision stage, the adopter makes the choice to adopt or reject an innovation. Note that adoption is the decision to make use of the innovation, not the actual use of the innovation. Use occurs in the implementation stage. Use represents an explicit behavioral change that puts the innovation into practice. Post implementation information seeking intended to reinforce the already-made innovation decision follows the implementation stage. In Rogers' model, this is known as the confirmation stage. The confirmation stage may result in continuation or reversal of the prior innovation decision.

In this study, we are interested in 2 types of barriers to consumers' use of technology: knowledge-based barriers and motivation-based barriers. We acknowledge that we are making an implicit assumption that the consumer has material access to the technology. In other words, the consumer has the means to physically possess the technology and necessary network access [38]. Our research model (presented in Figure 1) recognizes these 2 categories of barriers to the adoption and use of consumer health technologies. Knowledge-based barriers concern a lack of knowledge of the technology's existence, purpose, and operation. Motivation-based barriers relate to beliefs about the benefits of using the technology relative to the drawbacks of using the technology. Knowledge and motivation barriers align with the first 2 stages of Rogers' [36] innovation-decision process: knowledge and persuasion. During the knowledge stage, consumers become aware of the innovation and begin to understand its uses. In the persuasion stage, consumers form beliefs about the technology and its uses. Individual characteristics, such as age, impact both knowledge of and beliefs about a technology [36]. Therefore, we believe it will be instructive to examine baby boomers' awareness and perceptions of various consumer health technologies.

We first address barriers related to knowledge. The most fundamental of these is a lack of awareness of the technology and its uses [34]. Awareness may partially explain differences between different age cohorts when it comes to the adoption of the new communication technologies. For example, baby boomers are much more likely to own smartphones, desktop computers, and laptop computers than older cohorts, but they are somewhat less likely to own newer technologies (such as iPads) or use certain applications, such as using a smartphone to send/receive emails or access the Internet, than younger cohorts [30,39,40]. Knowledge barriers beyond awareness also exist. Consumers may be aware of the technology, but lack knowledge of its purpose or its operation. For example, consumers may be aware that kiosks exist, but may not know what they can be used for or how to use them.

The second category of barriers relates to motivations to adopt the technology. These barriers concern beliefs about the costs and benefits of using a technology for a specific purpose. For example, making consumer health technologies available for baby boomers with different cognitive, perceptual, and physical abilities is challenging [35,41] because these differences change the cost-benefit calculus. Moreover, baby boomer perceptions of the usefulness and usability of various consumer health technologies, the efficiency of care delivery, cost, and improvement of quality of life stemming from the use of these technologies may be barriers that inhibit adoption and use [28,35].

In the persuasion stage of the innovation adoption decision, the consumer begins to form beliefs about the technology as it relates to health care information acquisition and use. This is the beginning of an adoption cost-benefit evaluation by the consumer that determines the outcome of the adoption decision [42]. Because our primary interest here is in barriers, we focus on the adoption cost side of the equation. Several adoption costs are of particular interest to consumer health technologies. First, the relative difficulty of using the consumer health technology may serve as a barrier to its use. The relationship between perceived complexity and the use of an innovation is well established [43-46]. In some cases, training may reduce the effort required to use the technology. However, in other cases, the user may already know how to use the technology (ie, is trained), but may feel that the effort required to put the technology into use is too high. These barriers align with the

van Dijk and Hacker's [38] concept of skills access, which indicates that a lack of digital skills causes a combination of inadequate training and high complexity.

Beliefs regarding the suitability of a technology for health information tasks are also important. When a consumer perceives a technology unsuitable for use (not compatible with particular uses), he or she may be reluctant to adopt that technology. Prior research has demonstrated that compatibility beliefs impact adoption decisions [36,47,48]. In the context of consumer health technologies, it is possible that a consumer has no issues with using a particular technology for non-health care purposes, but believes that the technology is not appropriate to use for health-related tasks. For example, impropriety beliefs may emerge if consumers are particularly concerned about whether the technology is sufficiently secure to protect sensitive health-related information.

Finally, it may be that a consumer believes that a technology is suitable for health-related tasks and that he or she has the ability to use the technology without undue effort, but simply does not enjoy using the technology. For example, many consumers may have the ability to use call centers and believe that call centers are appropriate for health-related tasks. However, these same consumers may not enjoy using call centers [49]. (In fact, our results, presented later, support this contention.) Although perceived enjoyment has not received as much attention as the other beliefs discussed here, there are studies that demonstrate a link between perceived enjoyment and adoption of a technology [50-52].

In response to the aforementioned issues, we pose the following questions:

Research question 3: What knowledge and motivation barriers exist for baby boomers using various forms of technology tools for health purposes?

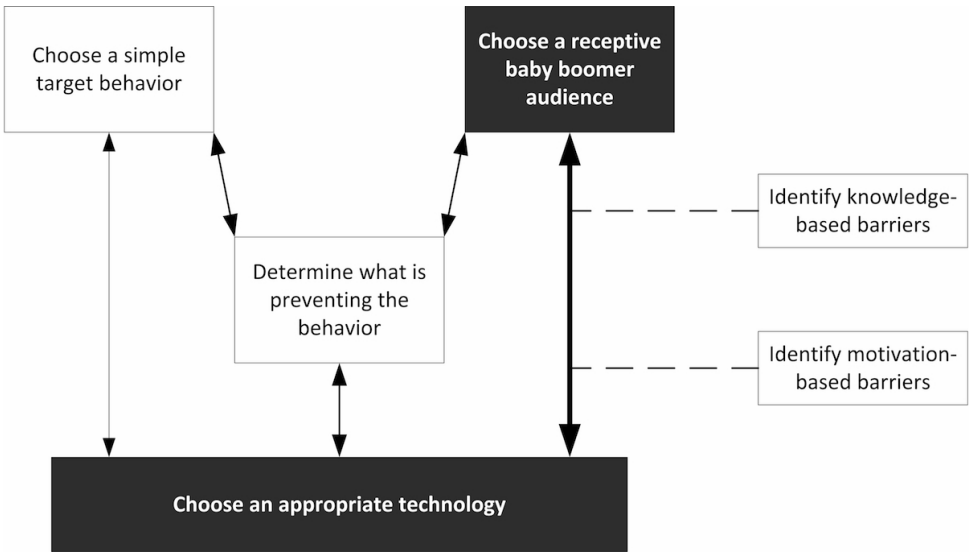
Research question 4: How do knowledge and motivation barriers that exist for baby boomers using various forms of technology tools for health purposes differ from younger adults and older adults?

Table 1 shows the specific knowledge and motivation belief-based barriers included in this study, along with citations for supporting research.

Table 1. Adoption process barriers to consumers adopting health technology.

Barrier	Supporting research
Knowledge-based barriers	
Awareness of the technology and its purpose	Rogers [36], Tak [35]
Not knowing how to use the technology	Rogers [36], Hsieh [53], Zeman [54], Katz [55], Moorman [21], Jones [41]
Motivation-based barriers	
Finding the technology too difficult to use	Rogers [36], Hsieh [53], Moorman [21], Phang [56], Emani [46], Kim [57], Jones [41]
Needing more training on the technology's use	Rogers [36], Hsieh [53]
Technology is not sufficiently secure	Matthews [58], Phang [56], Zulman [59]
Technology is not appropriate for health care use	Rogers [36], Karahanna [60], van Slyke [61], Emani [46]
Do not enjoy using the technology	Hsieh [53], van Der Heijden [62], Thong [63], Brown [64]

Figure 1. Phase 1 of steps in the persuasive design process (adapted from Fogg [24]). Areas of focus for the current study are highlighted in black and gray.



Methods

Overview

The readiness/barrier instrument used in this study (see Multimedia Appendix 1) was part of a larger consumer health technology survey. The University Institutional Review Board of the first author approved the study as an exempt study. A paper-based and an online version of the survey were developed in parallel. The online version was custom-developed using Qualtrics software (Qualtrics Corp, Provo, UT, USA). The survey was developed in both formats for convenience and to allow participation by those who were and were not regular technology users. Participants were allowed to choose their preferred method of responding.

Demographic data were collected for each respondent at the end of the survey. The consumer health technology readiness/barrier instrument was located in the middle of the survey. The instrument asks 8 questions about the use of 11 types of technology for health purposes (phone, website, email, call center, video conference, texting, podcast, kiosk, smartphone, blog, and wiki). One question asked respondents' perspectives concerning their readiness to use a specific type of technology and the remaining 7 questions addressed barriers to use of the same technology for health purposes.

To address research questions 1 and 2 (regarding respondents' readiness to use technology for health purposes) respondents were asked to check a box by each type of technology if they agreed with the statement: "Would have no problem using this technology for health care." This question stem is similar to that used in past research to ascertain that no barriers exist [41].

To address research questions 3 and 4, respondents were instructed to check a box by any issue (ie, barrier) they would face in using the specified type of technology to learn more about or manage their health (see Multimedia Appendix 1 for specific instructions). The list of 7 barriers was informed by aggregating the knowledge-based and motivation-based technology barriers reported and assessed in prior studies (see Table 1 for the list of supporting research). The barrier statements were as follows: (1) don't know what this technology is, (2) don't know how to use this technology, (3) this technology would be too difficult to use, (4) would need more training on how to use this technology, (5) don't feel the technology is secure, (6) doesn't seem appropriate for health care purposes, and (7) don't enjoy using this technology. Respondents were allowed to check multiple barriers. To show deliberate intent for those not selecting any barriers, respondents were asked to check a box indicating they would have no problem using a specific type of technology only if they checked none of the barriers.

Before the full data collection, a pilot study was conducted to assess the survey and data collection methods. A convenience sample of 14 individuals of varying ages completed 1 of the versions of the survey (7 on paper and 7 online) and then verbally answered a number of questions regarding the survey in general and specific survey items within each section of the survey. The pilot effort confirmed that the online version of the survey was functioning without error. Pilot participants of both the online and paper-based versions indicated that the parsimonious presentation (table format) of the barrier section of the consumer survey (1) was easy to understand, (2) provided a means for respondents to easily find and consider all barriers associated with a represented technology used for health care purposes as one holistic task, and (3) allowed respondents to quickly and efficiently answer the barrier section of the survey. Timed recordings indicated that the survey took between 20 and 35 minutes for respondents to complete and the method of data collection (paper vs online) did not seem to impact completion time. Feedback received from the pilot study led to minor modifications focused on enhancing readability and understandability.

Survey invitations and directions provided a layperson description of the consumer health technologies of interest to facilitate understanding of the survey context (see Multimedia Appendix 1 for the types of consumer health technologies and the definitions provided in the survey). Respondents were able to select their preferred method of responding (paper vs online) as an attempt to increase the response rate. Offering a choice accommodated personal preference for survey method and also facilitated a means to participate for those that may not have ready access to a personal computer.

Recruitment

The sample frame was composed of a sample of enrollees in a large pharmacy benefit management company. Participants who completed the survey were allowed to request that a small donation be made to a charitable organization on their behalf. The sampling frame afforded the opportunity to collect data from a national pool of respondents in the United States. The pharmacy benefit management company sent a personalized postal mail invitation to participate in the survey to a random sample of 3000 of their subscribers. The invitation included a paper-based copy of the survey as well as instructions for completing the online version as an option, if preferred. Only those who received the mail invitation were provided with the Web address for the online version of the survey. As a means to check for duplicate responses, paper-based surveys were stamped with unique identifiers and entry of this identifier was required as a log-in requirement to take the online version. One duplicate was identified and removed. Acceptable responses to the survey were provided by 506 individuals. However, age of the respondent was not available for 37 respondents. Therefore, data from 469 respondents were analyzed for this study resulting in a 15.63% response rate.

The surveys were completed from January 2010 through May 2011. Administration of the survey was timed to coincide with when the oldest boomers were reaching the retirement age of 65. The survey was initiated in a single state; thereafter, it was rolled out to other parts of the country. Reminders were sent by email at 3 months after the invitation was sent and then again at 7 months. It was not possible to send follow-up reminders specifically to nonrespondents, so only general reminders were sent that also directed respondents not to complete the survey a second time.

To understand respondents' preferences for survey method, a binomial test was conducted to compare respondents' selection of the paper versus the online version of the survey. Compared to an expected probability of 50%, respondents were significantly more likely to choose the paper version (59%, 277/469) over the online version (40.9%, 192/469; $P<.001$). Using the Kruskal-Wallis 1-way ANOVA on ranks test (H), age group differences for version of the survey completed were found to be statistically significant ($H_2=59.56$, $P<.001$). All 3 post hoc pairwise comparisons using the Mann-Whitney *U* test were also found to be statistically significant ($P<.05$). Compared to baby boomers, the younger age group was significantly more likely to choose the online version ($H_1=-39.46$, $P=.03$), whereas the older age group was significantly more likely to choose the paper version ($H_1=-76.50$, $P<.001$). In the younger age group, 65.3% (47/72) completed the online version compared to 48.1% (124/258) of baby boomers and 15.8% (22/139) of the older age group. These results reinforce our decision to give participants a choice in how to respond. Had we used only paper-based or online surveys, it is likely that we would have introduced a systematic bias in our sample. Therefore, this study appropriately followed recommendations to provide choice regarding the survey response media for purposes of inclusiveness [65].

Statistical Analysis

Frequencies and percentages were computed for the sample as a whole and by age group for relevant demographic, readiness, and barrier survey questions. Readiness and barrier survey questions used for this study were coded as yes/no responses. SPSS 20.0 (IBM Corp, Armonk, NY, USA) was used for all computations.

To address the first and second research questions regarding readiness to use the specified types of technology, the 11 types of technology were ordered from high to low based on the percentage of respondents that indicated they were ready to use each type. Kruskal-Wallis 1-way ANOVA on ranks test for independent samples was conducted for each of the 11 types of technology to test for age group differences. For these statistical tests, readiness responses were treated as a 2-point ordinal scales with unchecked responses coded as zero for "do not agree" and checked responses coded as 1 for "agree" in response to the statement "Would have no problem using this technology for health care." For significant Kruskal-Wallis tests, pairwise post hoc comparisons were conducted using the Mann-Whitney *U* test. Bonferroni corrections were made to *P* values to protect against inflated type I error rate when carrying out multiple tests. In addition, eta squared (η^2) was used to estimate the effect size for significant omnibus and post hoc comparisons. Effect sizes <0.06 were considered small, ≥0.06 and <0.14 were medium, and ≥0.14 were considered large [66,67].

The third research question assessed respondents' perceptions concerning 7 barriers to using the 11 types of technologies. A respondent could select more than 1 barrier. Although respondents were instructed to only check that they would have no problem using a specific type of technology if no barrier was checked, a few respondents (between 1 and 8 respondents for each type of technology) indicated that they had no problem using a technology and also checked a barrier to using that technology. For analysis, the number and percentage of respondents that checked at least 1 barrier were calculated for each type of technology. The types of technology were organized in a table format and ordered from low to high based on the percentage of total respondents that identified at least 1 barrier. The overall percentages for readiness and barriers for each type of technology were descriptively contrasted. Individual barriers identified for each type of technology by more than 14.9% (70/469) of the respondents were highlighted and described.

Analysis for the fourth research question regarding age group differences for barriers to technology use consisted of conducting chi-square tests on the raw nominal level data. Responses for the baby boomer age group were compared to responses of the younger and older age groups on the 7 barrier questions for each of the 11 types of technology. A total of 154 chi-square tests were computed for 2×2 contingency tables. That is, 77 chi-square tests compared baby boomers with the younger age group for each barrier and type of technology (7 barriers * 11 technologies=77 contingency tables) and 77 chi-square tests compared baby boomers to the older age group. For chi-square tests with expected cell counts less than 5 (ie, 59 contingency tables), the Fisher's exact test was used instead of Pearson chi-square test. Only 2 contingency tables with expected cell counts less than 5 were statistically significant. Notably, low expected cell counts were more prevalent for comparisons between baby boomers and younger adults (eg, 40 contingency tables) because younger adults reported fewer barriers overall and were the age group with the smallest number or respondents.

Results

Demographics

Basic demographic data are presented in Table 2 for the sample as a whole (N=469) and by age group (younger age group: 72/469, 15.4%; baby boomer age group: 258/469, 55.0%; older age group: 139/469, 29.6%). For the baby boomer age group, responses were relatively gender balanced. This was not the case with the older and younger age groups. The younger age group had a significantly higher proportion of responses from females (males: 19/72, 26.4%; females: 53/72, 73.6%; $P=.02$) whereas the older age group had a significantly higher proportion of males (males: 61/139, 43.9%; females: 78/139, 56.1%; $P<.001$). Ethnicity distributions were similar for all age groups. Marital status and household income varied across the groups. These differences were not unexpected given the different life stages of the groups. For example, a significantly higher proportion of the younger group reported being single than was the case for the baby boomers (younger: 17/72, 26.6%; baby boomer: 17/258, 7.2%, $P=.001$) and the older group (younger: 17/72, 26.6%; older: 6/139, 4.5%, $P=<.001$).

Overall, the survey sample was healthy. Of the 436 respondents that answered the question "How would you rate your overall health?" 369 (84.6%) considered themselves to be in good to excellent health, 57 (13.1%) fair, and 10 (2.3%) poor. Of respondents that reported they were in poor health, 1 was in the

older age group, 6 were in the baby boomer group, and 3 were in the younger age group.

Respondents were asked to indicate if they "would have no problems using [a specific type of] technology for health care...to learn more about or manage [their] health." No particular health conditions were specified. As noted above, only 2% (10/436) of respondents considered themselves to be in poor health.

Table 2. Demographics of respondents (N=469).

Demographics	Age group, n (%)			Total
	18-45 (n=72)	46-64 (n=258)	>64 (n=139)	
Age (years), mean (SD)	35.3 (7.9)	56.9 (4.8)	73.2 (6.2)	58.4 (13.5)
Gender				
Female	53 (73.6)	143 (55.4)	61 (43.9)	257 (54.8)
Male	19 (26.4)	115 (44.6)	78 (56.1)	212 (45.2)
Total responses	72	258	139	469
Ethnicity				
Caucasian	57 (91.9)	198 (88.4)	115 (93.5)	370 (90.5)
African-American	2 (3.2)	7 (3.1)	4 (3.3)	13 (3.2)
Hispanic	1 (1.6)	8 (3.6)	2 (1.6)	11 (2.7)
Asian	0 (0.0)	10 (4.5)	1 (0.8)	11 (2.7)
Other	2 (3.2)	1 (0.4)	1 (0.8)	4 (1.0)
Total[a]	62	224	123	409
Marital status				
Single	17 (26.6)	17 (7.2)	6 (4.5)	40 (9.2)
Married/partner	43 (67.2)	199 (84.3)	108 (81.2)	350 (80.8)
Separated/divorced/widowed	4 (6.3)	20 (8.5)	19 (14.3)	43 (9.9)
Total[a]	64	236	133	433
Household income (US $)				
Under $30,000	1 (1.9)	13 (7.1)	24 (26.7)	38 (11.7)
$30,000-$49,999	6 (11.5)	35 (19.0)	24 (26.7)	65 (19.9)
$50,000-$99,999	19 (36.5)	80 (43.5)	29 (32.2)	128 (39.3)
$100,000 or more	26 (50.0)	56 (30.4)	13 (14.4)	95 (29.1)
Total[a]	52	184	90	326
Health rating				
Excellent	8 (12.3)	21 (8.8)	8 (6.1)	37 (8.5)
Very good	29 (44.6)	81 (33.9)	36 (27.3)	146 (33.5)
Good	21 (32.3)	105 (43.9)	60 (45.5)	186 (42.6)
Fair	4 (6.2)	26 (10.9)	27 (20.5)	57 (13.1)
Poor	3 (4.6)	6 (2.5)	1 (0.8)	10 (2.3)
Total[a]	65	239	132	436
Survey version				
Paper	25 (37.7)	133 (51.6)	117 (81.5)	265 (56.5)
Online	47 (65.3)	125 (48.4)	22 (15.8)	204 (43.5)
Total	72	258	139	469

[a]Within-group differences in totals are due to missing responses.

Boomer Readiness

As shown in Table 3, the percentages of all respondents that were ready to use each type of technology for health purposes varied widely—from a high of 82.7% (388/469) for use of the standard telephone to a low of 22.3% (105/469) for wikis. A scan of the distribution of percentages for all respondents revealed that the technologies fell into 3 clusters. Most of the respondents were comfortable with using the standard telephone, a health informational website, and email for health purposes—82.7% (388/469) to 73.8% (346/469) of respondents (the first cluster). Close to half of respondents (230/469, 49.0%) were ready to use automated call centers, medical video conferencing, and short message service (SMS) text messaging for health purposes—49.0% (230/469) to 42.0% 197/469) of respondents (the second cluster). The percentage of respondents that were ready to use technologies such as podcasts, personal digital assistant (PDA)/smartphone applications, blogs, and wikis for health purposes ranged from a high of 35.2% (165/469) to a low of 22.4% (105/469) (the third cluster). Looking at the distributions for the 3 age groups in Table 3, these technology clusters were more distinct for baby boomers than the younger or older age groups. Note that no cells had expected counts less than 5.

Table 3. Readiness to use consumer health technology by type of technology and age group.

Type of technology	Age group, n (%)			Total	Statistical tests		
	18-45 (n=72)	46-64 (n=258)	>64 (n=139)	(N=469)	H_2	*P*	η^2
Phone	57 (79.2)	222 (86.0)	109 (78.4)	388 (82.7)	4.4	.11	—
Website	65 (90.3)	218 (84.5) [b]	83 (59.7)	366 (78.0)	39.7	<.001	0.09
Email	57 (79.2)	209 (81.0) [b]	80 (57.6)	346 (73.8)	26.9	<.001	0.06
Call center	43 (59.7)	135 (52.3) [b]	52 (37.4)	230 (49.0)	11.9	.003	0.03
Video conference	47 (65.3)	128 (49.6 [b]	36 (25.9)	211 (45.0)	34.6	<.001	0.07
Texting	41 (56.9)	128 (49.6) [b]	28 (20.1)	197 (42.0)	39.9	<.001	0.09
Podcast	41 (56.9)	95 (36.8 [a,b]	29 (20.9)	165 (35.2)	27.7	<.001	0.06
Kiosk	37 (51.4)	91 (35.3) [a]	33 (23.7)	161 (34.3)	16.3	<.001	0.04
Smartphone	41 (56.9)	89 (34.5) [a,b]	16 (11.5)	146 (31.1)	48.6	<.001	0.10
Blog	38 (52.8)	74 (28.7) [a]	26 (18.7)	138 (29.4)	26.6	<.001	0.06
Wiki	30 (41.7)	55 (21.3) [a]	20 (14.4)	105 (22.4)	20.7	<.001	0.04

[a]Significantly different pairwise comparison between the baby boomer (46-64) and younger (18-45) age groups (*P*<.05).

[b]Significantly different pairwise comparison between the baby boomer (46-64) and older (> 64) age groups (*P*<.05).

Readiness Comparisons Among Age Groups

In testing for age group differences, the Kruskal-Wallis H omnibus tests found the proportions for the 3 age groups were significantly different for all types of technology except the phone (see Table 3). Effect sizes were computed using eta squared to determine practical significance of these age group differences. Eta squared indicated a medium effect (0.06-0.10) for 7 types of technology (website, email, video conference, texting, podcast, kiosk, and blog) and a small effect (0.03-0.04) for 3 types (call center, kiosk, and wiki) (Table 3). The difference in percentages of respondents that indicated their readiness to use call centers for health purposes (η^2=0.03) was only 8% between the younger and baby boomer groups and 15% between the baby boomer and older age groups. Similarly, the percent differences for kiosks (η^2=0.04) was only 16% between the younger and baby boomer groups and 11% between baby boomer and older age groups. For wikis (η^2=0.04), the percentage for the younger age group was double (42% vs 21%) the percentage for baby boomers whereas the difference between the baby boomer and older age groups was only 7%.

Post hoc pairwise comparisons conducted to determine which groups were statistically significantly different found differences between the younger and older groups for all 10 types of technology. Given that the focus of this study is on baby boomers, however, only the results from pairwise comparisons with baby boomers are annotated in Table 3 (significant pairwise comparisons are annotated with an “a” for younger and baby boomer comparisons and “b” for older and baby boomer comparisons) and discussed subsequently.

Based on the statistical and practical significance of post hoc comparisons, the 10 types of technologies fell into 2 clusters of 5 each. The first cluster included websites, email, call centers, video conferencing, and texting. There were no significant differences between baby boomers and the younger age group in their readiness to use the technologies for health purposes in the first cluster. However, in comparisons with the older age group, baby boomers were significantly more likely to be ready to use these technologies for health purposes. For these differences, the *P* values were small (*P*<.001) and the effect sizes were large (η^2=0.14-0.18) for all technologies in this cluster except the call center which had a medium effect size

(P=.01, η^2=0.09). As noted earlier, the effect size was small for the omnibus test for call centers and the percent differences were 8% (younger to boomer) and 15% (boomer to older).

The second cluster (which included the remaining 5 technologies) was less distinct, but still demonstrated the role of age in readiness to adopt specific technologies for health purposes. In comparisons with the younger age group, baby boomers were significantly less likely (1.5 to 2 times less likely) to be ready to use the technologies in this cluster for health purposes (ie, podcasts, kiosks, smartphone apps, blogs, and wikis). These differences were significant (P<.005) with large effect sizes ($\eta^2 \geq 0.14$) for all technologies except the kiosk with a medium (P=.03, η^2=0.12). As with the call center, the effect size reported earlier for the omnibus test for kiosks (see Table 3) was small (0.04) and the between-group percent differences were 16.1% for younger to boomer and 11.6% for boomer to older comparisons. For the kiosk, the percent difference was smaller between the boomer and older age group (11.6% vs 16.1% between the boomer and younger) and for the call center the percent difference was smaller between the boomers and the younger age group (7.4% vs 14.9% between boomer and older age groups).

In examining comparisons with the older age group for technologies in the second cluster, baby boomers were significantly more likely to be ready to use podcasts with a medium effect size (P=.005, η^2=0.10) and smartphones with a large effect size (P<.001, η^2=0.14). On the other hand, baby boomers were no more likely than the older age group to be ready to use kiosks, blogs, and wikis. Because baby boomers were less likely than the younger age group, but more likely than the older age group to be ready to use podcasts and smartphones, these 2 technologies may be promising targets for consumer health technologies applications.

It should be noted that although statistically significant, the effect sizes for the call center and kiosk indicated that the age group differences for these 2 types of technology were of limited practical significance.

Boomer Barriers

Not surprisingly, the technologies fell into the same 3 clusters for the barrier measures as they did for the readiness measure (presented in Table 3) based on total percentages of respondents that checked at least 1 barrier. Of the 469 respondents, most indicated they were ready to use the phone (82.7%, 388/469) for health purposes (see Table 3) and few respondents identified barriers (16.6%, 78/469) to using the phone for health purposes (see Table 4). Similarly, websites (78.0%, 366/469) and email (73.8%, 346/469) had similar proportions of respondents that indicated they were ready to use these technologies for health purposes (see Table 3) and relatively few respondents selected barriers to using these technologies: 20.8% (98/469) for websites and 26.6% (125/469) for email (see Table 4).

Table 4. Respondents with barriers checked by type of technology and type of barrier.

Type of technology	Barriers to consumer health technology adoption, n (%)[a]							
	Total	Knowledge-based		Motivation-based				
	Checked >1 barrier	Don't know what it is	Don't know how to use	Too difficult to use	Need more training	Not secure	Not appropriate	Don't enjoy using
Phone	78 (16.6)	5 (1.1)	9 (1.9)	4 (0.9)	6 (1.3)	15 (3.2)	16 (3.4)	31 (6.6)
Website	98 (20.9)	22 (4.7)	30 (6.4)	4 (0.9)	26 (5.5)	16 (3.4)	9 (1.9)	23 (4.9)
Email	125 (26.7)	9 (1.9)	23 (4.9)	5 (1.1)	12 (2.6)	44 (9.4)	32 (6.8)	29 (6.2)
Call center	238 (50.7)	15 (3.2)	23 (4.9)	6 (1.3)	9 (1.9)	40 (8.5)	75 (16.0)[a]	119 (25.4)[a]
Video conference	251 (53.5)	35 (7.5)	74 (15.8)[a]	16 (3.4)	70 (14.9)	29 (6.2)	30 (6.4)	54 (11.5)
Texting	269 (57.4)	16 (3.4)	57 (12.2)	12 (2.6)	28 (6.0)	61 (13.0)	87 (18.6)[a]	91 (19.4)[a]
Podcast	302 (64.4)	74 (15.8)[a]	89 (19.0)[a]	9 (1.9)	64 (13.6)	27 (5.8)	49 (10.4)	79 (16.8)[a]
Kiosk	307 (65.5)	31 (6.6)	45 (9.6)	7 (1.5)	29 (6.2)	83 (17.7)[a]	96 (20.5)[a]	96 (20.5)[a]
Smartphone	319 (68.0)	48 (10.2)	107 (22.8)[a]	19 (4.1)	63 (13.4)	49 (10.4)	45 (9.6)	91 (19.4)[a]
Blog	331 (70.6)	52 (11.1)	88 (18.8)[a]	10 (2.1)	54 (11.5)	81 (17.3)[a]	113 (24.1)[a]	92 (19.6)[a]
Wiki	357 (76.1)	93 (19.8)[a]	92 (19.6)[a]	8 (1.7)	43 (9.2)	91 (19.4)[a]	112 (23.9)[a]	74 (15.8)[a]

[a]Percentages of respondents (N=469) above 15% are marked to highlight the highest concentration of barriers.

However, among all respondents, the percentages that indicated barriers for the other technologies were double and triple the percentages that selected barriers for the phone, website, and email: 16.6% (78/469) to 26.7% (125/469) compared to 50.7% (238/469) to 76.1% (357/469) (see Table 4). As can be inferred from Table 3, 51.0% (239/469) to 58.0% (272/469) of respondents indicated that they were not ready to use call centers, video conferencing, or texting for health purposes (the second cluster). By the same token, more than half (50.7%, 238/469 to 57.3%, 269/469) of respondents identified at least

1 barrier to using these technologies. The 2 barriers checked by the highest percentage of respondents for call center and texting were “don’t enjoy using” (25.4%, 119/469 and 19.4%, 95.8/469, respectively) and “not appropriate” (16.0%, 75/469 and 18.6%, 87/469, respectively). For video conferencing, the 2 barriers checked by the highest percentage of respondents were “don’t know how to use” (16%, 74/469) and “need more training” (14.9%, 70/469). These percentages indicate that respondents are willing to use video conferencing if they were trained on its use. The objections to call centers and texting may be harder to overcome. Examining age group differences may provide more insights to guide adoption efforts for these technologies.

The third cluster of technologies (podcasts, kiosks, smartphones, blogs, and wikis) fared poorly overall. The percentage of respondents that cited at least 1 barrier for this cluster of technologies was lowest for podcasts (64.4%, 302/469) and highest for wikis (76.1%, 357/469). In examining the top barriers to use, the motivation-based barriers stood out, especially enjoyment, which was identified as a problem by more than 14.9% (>70/469) of respondents for every technology in this cluster (see Table 4). In addition to enjoyment, appropriateness and security were among the top motivation-based barriers checked for kiosks, blogs, and wikis. The most notable knowledge-based barrier was “don’t know how to use” with 19.0% (89/469) to 22.8% (107/469) of respondents checked this barrier for every technology in this cluster except kiosks. In addition, 15.8% (74/469) of respondents indicated that they did not know what a podcast was and what a wiki was.

Barrier Comparisons Among Age Groups

Chi-square tests comparing the baby boomer age group with the younger and older age groups yielded statistically significant results for 31 of the 154 pairwise comparisons (see Figures 2 and 3 for bar charts and Table 5 for actual counts and percentages). In all, 10 tests were statistically significant for the baby boomer and younger age group comparisons and 21 tests were significant for the baby boomer and older age group comparisons. Significant age group differences were found primarily among the knowledge-based barriers, 22 knowledge-based and 9 motivation-based. Consistent with the readiness findings reported earlier, the phone was the only type of technology for which there were no significant between-group differences for any of the barriers.

Table 5. Barriers to readiness to use consumer health technology by type of barrier, type of technology, and age group.

Type of tech by age group	Barriers to consumer health technology adoption, n (%)[a]						
	Knowledge-based		Motivation-based				
	Don't know what it is	Don't know how to use	Too difficult to use	Need more training	Not secure	Not appropriate	Don't enjoy using
Phone							
18-45	0 (0.0)[b]	0 (0.0)[b]	1 (1.4)[b]	0 (0.0)[b]	3 (4.2)[c]	2 (2.8)[c]	9 (12.5)
46-64	2 (0.8)	4 (1.6)	3 (1.2)	4 (1.6)	8 (3.1)	9 (3.5)	15 (5.8)
>64	3 (2.2)[b]	5 (3.6)[c]	0 (0.0)[b]	2 (1.4)[b]	4 (2.9)[c]	5 (3.6)[c]	7 (5.0)
Website							
18-45	0 (0.0)[c]	0 (0.0)[c]	0 (0.0)[b]	1 (1.4)[c]	1 (1.4)[c]	1 (1.4)[c]	2 (2.8)[c]
46-64	8 (3.1)	12 (4.7)	2 (0.8)	13 (5.0)	11 (4.3)	6 (2.3)	10 (3.9)
>64	14 (10.1) [d]	18 (12.9) [d]	2 (1.4)[b]	12 (8.6)	4 (2.9)	2 (1.4)[c]	11 (7.9)
Email							
18-45	0 (0.0)[b]	0 (0.0)[b]	0 (0.0)[b]	0 (0.0)[b]	7 (9.7)	8 (11.1)	1 (1.4)[c]
46-64	3 (1.2)	5 (1.9)	3 (1.2)	6 (2.3)	25 (9.7)	16 (6.2)	15 (5.8)
>64	6 (4.3)[c]	18 (12.9) [d]	2 (1.4)[b]	6 (4.3)[c]	12 (8.6)	8 (5.8)	13 (9.4)
Call center							
18-45	0 (0.0)[b]	0 (0.0)[b]	0 (0.0)[b]	0 (0.0)[b]	5 (6.9)	10 (13.9)	16 (22.2)
46-64	6 (2.3)	6 (2.3)	4 (1.6)	4 (1.6)	22 (8.5)	47 (18.2)	72 (27.9)
>64	9 (6.5) [d]	17 (12.2) [d]	2 (1.4)[b]	5 (3.6)[c]	13 (9.4)	18 (12.9)	31 (22.3)
Video conference							
18-45	1 (1.4)[c]	4 (5.6)	2 (2.8)[c]	6 (8.3)	2 (2.8)[c]	2 (2.8)[c]	7 (9.7)
46-64	17 (6.6)	33 (12.8)	11 (4.3)	39 (15.1)	15 (5.8)	19 (7.4)	34 (13.2)
>64	17 (12.2)	37 (26.6) [d]	3 (2.2)[c]	25 (18.0)	12 (8.6)	9 (6.5)	13 (9.4)
Texting							
18-45	0 (0.0)[b]	3 (4.2)[c]	0 (0.0)[c]	1 (1.4)[c]	6 (8.3)	16 (22.2)	9 (12.5)
46-64	6 (2.3)	13 (5.0)	7 (2.7)	13 (5.0)	38 (14.7)	51 (19.8)	56 (21.7)
>64	10 (7.2) [d]	41 (29.5) [d]	5 (3.6)[c]	14 (10.1)	17 (12.2)	20 (14.4)	26 (18.7)
Podcast							
18-45	3 (4.2) [e]	5 (6.9)[e]	0 (0.0)[c]	6 (8.3)	3 (4.2)[c]	4 (5.6)	12 (16.7)
46-64	38 (14.7)	44 (17.1)	7 (2.7)	38 (14.7)	16 (6.2)	28 (10.9)	52 (20.2)
>64	33 (23.7) [d]	40 (28.8) [d]	2 (1.41)	20 (14.4)	8 (5.8)	17 (12.2)	15 (10.8) [d]
Kiosk							
18-45	0 (0.0) [c,e]	0 (0.0) [c,e]	0 (0.0)[b]	1 (1.4)[c]	10 (13.9)	10 (13.9) [e]	16 (22.2)
46-64	21 (8.1)	15 (5.8)	4 (1.6)	16 (6.2)	46 (17.8)	64 (24.8)	60 (23.3)
>64	10 (7.2)	30 (21.6) [d]	3 (2.2)[b]	12 (8.6)	27 (19.4)	22 (15.8)[d]	20 (14.4) [d]
Smartphone							
18-45	3 (4.2)[c]	5 (6.9) [e]	1 (1.4)[c]	3 (4.2)	5 (6.9)	2 (2.8) [e]	14 (19.4)
46-64	18 (7.0)	58 (22.5)	14 (5.4)	30 (11.6)	28 (10.9)	29 (11.2)	57 (22.1)

Type of tech by age group	Barriers to consumer health technology adoption, n (%)[a]						
	Knowledge-based		Motivation-based				
	Don't know what it is	Don't know how to use	Too difficult to use	Need more training	Not secure	Not appropriate	Don't enjoy using
>64	27 (19.4) [d]	44 (31.7) [d]	4 (2.9)	30 (21.6) [d]	16 (11.5)	14 (10.1)	20 (14.4)
Blog							
18-45	0 (0.0) [e]	6 (8.3)	0 (0.0)[b]	3 (4.2)	12 (16.7)	11 (15.3) [e]	13 (18.1)
46-64	25 (9.7)	46 (17.8)	6 (2.3)	30 (11.6)	43 (16.7)	73 (28.3)	59 (22.9)
>64	27 (19.4) [d]	36 (25.9)	4 (2.91)	21 (15.1)	26 (18.7)	29 (20.9)	20 (14.4) [d]
Wiki							
18-45	6 (8.3) [e]	6 (8.3)	0 (0.0)[b]	4 (5.6)	12 (16.7)	13 (18.1)	10 (13.9)
46-64	53 (20.5)	41 (15.9)	5 (1.9)	22 (8.5)	56 (21.7)	71 (27.5)	51 (19.8)
>64	34 (24.5)	45 (32.4) [d]	3 (2.2)[c]	17 (12.2)	23 (16.5)	28 (20.1)	13 (9.4) [d]

[a]Percentage of respondents that indicated agreement with the barrier statement within an age group. Younger age group (18-45, n=72); baby boomers (46-64, n=258); older age group (>64, n=139).

[b]Contingency tables (n=24) containing 2 cells with expected cell counts <5.

[c]Contingency tables (n=35) containing 1 cell with expected cell counts <5.

[d]Significant pairwise comparisons (*P*<.05) between the boomer (46-64) and older (>64) age groups (italicized).

[e]Significant pairwise comparisons (*P*<.05) between the boomer (46-64) and younger (18-45) age groups (italicized).

In baby boomer and younger age group comparisons, the younger age group was more favorable toward the technologies (ie, the younger group identified fewer barriers). Moreover, the comparisons between baby boomers and younger age group aligned closely with the readiness results reported earlier. As with the readiness measure, the technologies fell into 2 clusters. Consistent with the readiness findings, there were no differences between the baby boomer and younger age groups on the barrier measures for the first technology cluster (ie, website, email, call center, video conferencing, or texting). All 10 statistically significant findings between baby boomers and the younger age group were in the second technology cluster (podcasts, kiosks, smartphones, blogs, and wikis). Moreover, 7 of the 10 significant findings were for knowledge-based barriers. Baby boomers were significantly more likely to check "don't know what it is" for all these technologies with the exception of the smartphone. They were significantly more likely than the younger group to check "don't know how to use" for all technologies except for the blog and wiki. Three significant age group differences were found for 1 motivation-based barrier. Baby boomers were more likely than the younger age group to check "Doesn't seem appropriate for health care purposes" for kiosks, smartphones, and blogs.

Comparisons between the baby boomer and older age groups indicate—although younger—baby boomers were not always more favorable toward the technologies. The trend toward favorability and younger age holds true for the first technology cluster (website, email, call center, video conferencing, or texting), but does not hold true for motivation-based barriers in the second technology cluster (podcasts, kiosks, smartphones, blogs, and wikis).

For the first cluster of technologies, baby boomers were generally more favorable to the technologies (ie, checked fewer barriers) than the older age group. Eight of the 21 statistically significant differences were among knowledge-based barriers in the first cluster. Compared to the older age group, baby boomers were significantly less likely to check knowledge-based barriers for website, call center, and texting. For these 3 technologies, baby boomers were 3 times less likely to check "don't know what it is" and 3 to 6 times less likely to check "don't know how to use." For email and video conference, baby boomers and the older age group were equally likely to check "don't know what it is" but the older ager group was significantly more likely to check "don't know how to use." Although there were no significant age group differences for motivation-based barriers in the first cluster, high percentages of all 3 age groups checked barriers related to appropriateness and enjoyment for call centers and texting. For videoconferencing, the motivation-based barriers were more related to training and enjoyment.

In the second cluster of technologies cluster (podcasts, kiosks, smartphones, blogs, and wikis), significant differences between baby boomers and the older age group were evenly split between knowledge-based (7 tests were significant) and motivation-based barriers (6 tests were significant). Of these 5 technologies, baby boomers' responses were most favorable toward smartphones and podcasts. Baby boomers were significantly less likely than the older age group to check either of the knowledge-based barriers for these 2 technologies. For smartphones, only 7.0% (18/258) of baby boomers checked "don't know what is" and 22.5% (58/258) checked "don't know how to use" compared to 19.4% (27/139) and 31.6% (44/139) of the older age group. For podcasts, 14.7% (38/258) of baby boomers checked "don't

know what is" and 17.1% (44/258) checked "don't know how to use" compared to 23.7% (33/139) and 28.8% (40/139) of the older age group. This finding is consistent with the earlier analysis on readiness measures and reinforces the argument that these 2 technologies may be promising targets for consumer health information technology applications for baby boomers.

There was less consistency for the other 3 technologies in the second cluster (kiosks, blogs, and wikis) around both knowledge-based barriers. Baby boomers were significantly less likely than the older age group to check "don't know what it is" for blogs (9.7%, 25/258 vs 19.4%, 27/139N), but "don't know how to use" for kiosks (5.8%, 15/258 vs 21.6%, 30/139) and wikis (15.9%, 41/258 vs 32.4%, 45/139). Importantly, a high percentage (over 20%) of both groups indicated that they don't know what wikis were compared to 7.2% (10/139) to 8.1% (21/258) for kiosks.

Three motivation-based barriers were significant for technologies in the second cluster: "need more training on how to use," "doesn't seem appropriate for health care purposes," and "don't enjoy using." Baby boomers were significantly less likely than the older age group to select the barriers related to training for the smartphone (11.6%, 30/258 vs 21.6%, 30/139). On the other hand, baby boomers were significantly more likely to check "not appropriate" as a barrier for kiosks (24.8%, 64/258 vs 15.8%, 22/139 and "don't enjoy using" as a barrier for all technologies in this cluster except the smartphone (podcasts: 20.2%, 52/258 vs 10.8%, 15/139; kiosks: 23.3%, 60/258 vs 14.4%, 20/139; blogs: 22.9%, 59/258 vs 14.4%, 20/139; wikis: 19.8%, 51/258 vs 9.4%, 13/139).

In examining Figure 3 for smartphones, baby boomers were significantly less likely (7.0%, 18/258) than the older group (19.4%, 27/139) and equally as likely as the younger age group (4.2%, 3/72) to check "don't know what it is." Moreover, boomers (22.5%, 58/258) were significantly less likely than the older age group (31.7%, 44/139) and more likely than the younger age group (6.9%, 5/72) to check "don't know how to use." These findings indicate that baby boomers' awareness of smartphones is on par with the younger age group, but they lag behind the younger group in knowing how to use smartphones. In fact, the 2 barriers checked by the highest percentage of baby boomers were "don't know how to use" (22.5%, 58/258) and "don't enjoy using" (22.1%, 57/258). Although none of the groups seemed to enjoy using smartphones for health purposes (younger: 19.4%, 14/72; boomers: 22.1%, 57/258; older 14.4%, 20/139), baby boomers and the younger age groups were significantly less likely than the older age group to indicate that they needed training. So although 22.5% (58/258) of baby boomers indicated they don't know how to use smartphones for health purposes, enjoyment was a barrier for a higher percentage (22.1%, 57/258) than was training (11.6%, 30/258).

It is worthwhile to connect some of the readiness and barrier findings, particularly related to podcasts and smartphones. Recall from the readiness findings, podcasts and smartphones were 2 technologies baby boomers were more ready to use than the older age group and less ready to use than the younger age group. In examining the bar chart for podcasts in Figure 3, baby boomers were significantly less likely to check both of the knowledge-based barriers and significantly more likely to check "do not enjoy" using podcasts compared to the older age group. Podcasts were typically audio-only at the time of the survey. The ability to add images and video (ie, webcasting and streaming video) might increase enjoyment of podcast-like technologies.

Figure 2. Bar charts of barriers to readiness to use consumer health technology by type of technology, type of barrier, and age group for phone, website, email, call center, and video conference.

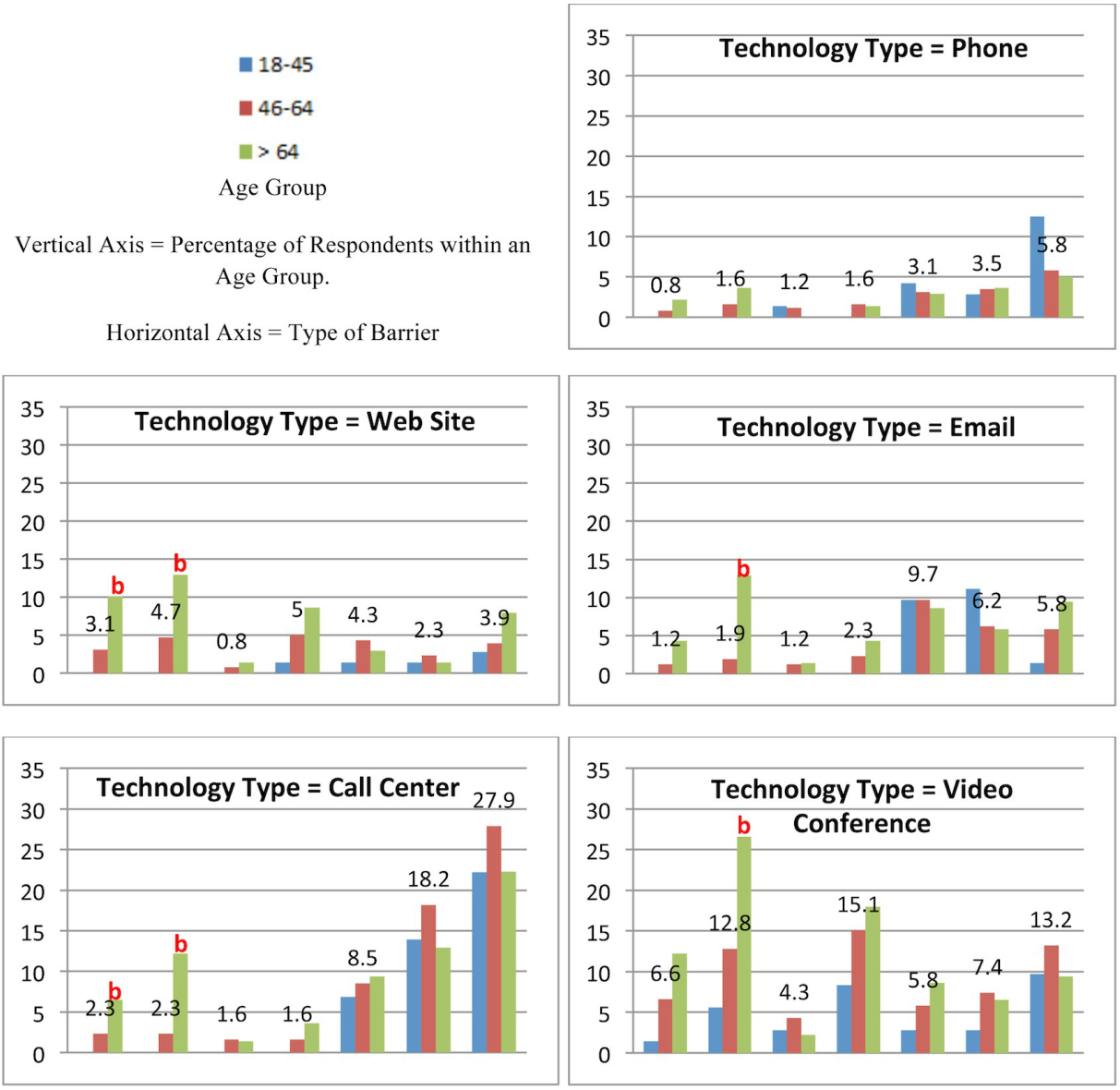
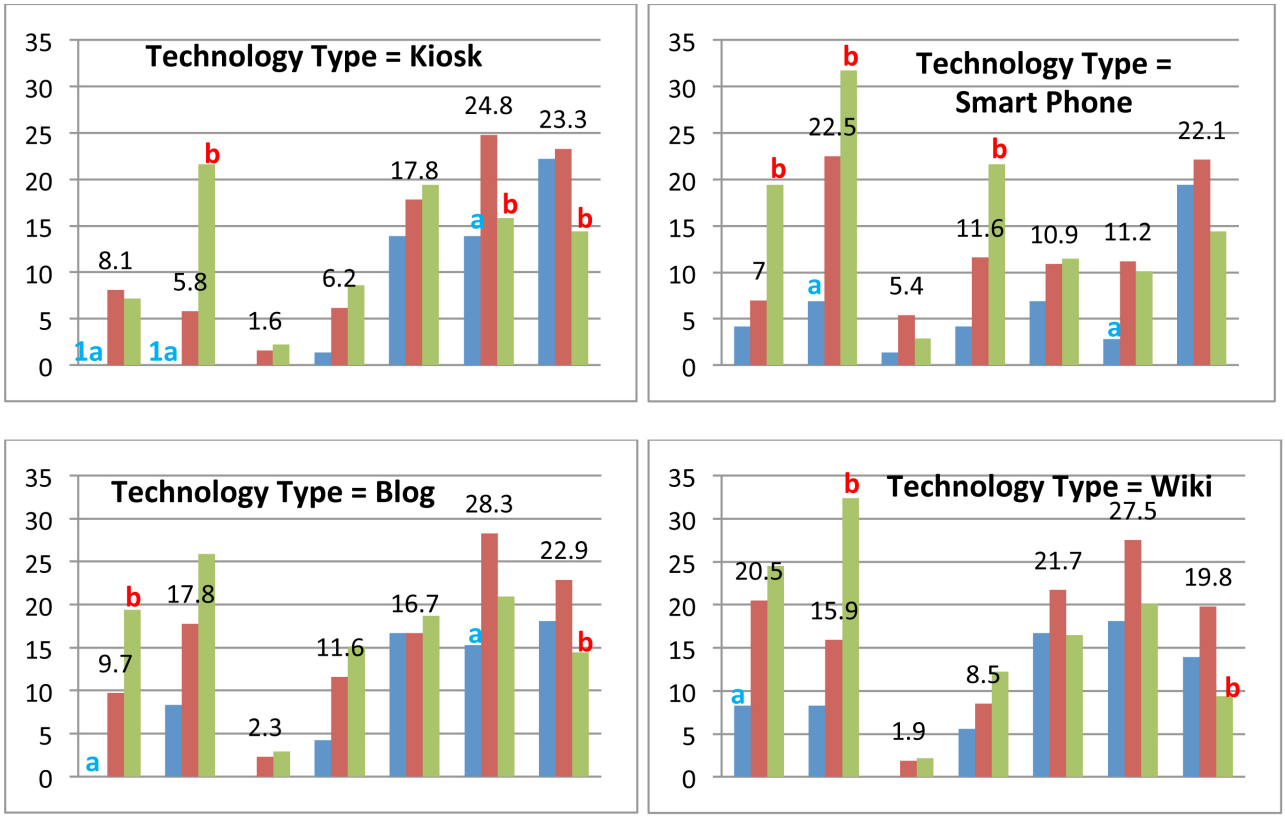


Figure 3. Bar charts of barriers to readiness to use consumer health technology by type of technology, type of barrier, and age group for texting, podcast, kiosk, and smartphone.



Discussion

Principal Findings

Our results offer a number of insights that may be useful to those interested in designing and promoting the use of consumer health technologies. First, from our data it seems that baby boomers are ready to use numerous technologies for health-related purposes. Familiarity is a reasonable explanation for this readiness. Our data indicate that most baby boomers are ready to use the phone, websites, and email for health-related purposes. Independent data indicate that most baby boomers have experience with these technologies. A Pew Research report indicates that over 75% of baby boomers interact with websites and over 90% use email [30,40]. Familiarity with a technology is an important factor in determining subsequent adoption of that and related technologies. It is useful to think of consumer health technologies as clusters of technologies in which a core technology, such as a website, is adapted for a specific purpose such as retrieving health information. Familiarity with the core technology reduces the perceived risk of adopting the specific application of the technology [68]. Adoption of a new technology is by its very nature an uncertain venture; familiarity reduces this uncertainty, which increases readiness. Approximately 50% of baby boomers in our study indicated they were ready to use call centers, video conferencing, and texting for health care purposes and they were on par with the younger age group indicating only limited barriers to use. These findings indicate that the perceived risk of adopting these technologies for health care may be low for baby boomers.

When potential adopters are not familiar with the core technology, the costs of adoption are increased. This occurs primarily through 2 mechanisms, the first of which is the previously mentioned adoption risk. The second way concerns the costs of learning to use the new technology. If one is familiar with the core technology, the learning curve is largely limited to using the core technology for the new purpose. Consider the example of someone evaluating the use of email to communicate with a health care professional. If this individual is already familiar with using email in other contexts, he or she must only learn aspects of email use related to the health care context (eg, what information is appropriate to communicate via email).

Our thinking regarding the role of familiarity is confirmed by results related to consumer health technologies with low readiness. For example, only 37% of our baby boomer respondents reported being ready to use podcasts for health purposes. Pew Research reports that only 20% of baby boomers have interacted with podcasts [40]. Similar results are evident for smartphones. Less than 35% of our baby boomers reported being ready to use smartphones for health purposes. A Pew Research report from the same period as our data collection found that less than 30% of baby boomers used smartphones [69]. Although smartphones are often used for work purposes, smartphone adoption among older adults has been shown to lag behind that of younger consumers [70]. Findings related to blogs further confirm the role of familiarity. Approximately 29% of our baby boomers were ready to use blogs for health purposes, although Pew Research reported less than 30% of baby boomers use blogs [40]. (Comparative data are not available for video conferencing, kiosk, or wiki use.)

Comparisons across age groups are also explained by differences in familiarity. For example, approximately half of our baby boomer respondents indicated a readiness to use texting for health purposes compared with approximately 20% of our older group (age >64 years). According to a Pew Research report, over 70% of baby boomers use texting generally compared with 35% for the older age group [71]. Similar results were found with websites and email; 85% and 81%, respectively, of the baby boomers in our sample were ready to use these technologies for health purposes. The proportions were significantly lower for our >64 age group (60% and 58%, respectively). A Pew report indicated over 75% of baby boomers are online, compared with 58% of those aged 65-73 years and only 30% of those older than 74 years [40].

Our readiness results offer insights for change agents interested in promoting consumer health technology use. Change agents would be well advised to focus on technologies that are already familiar to sizable portions of the target age group. For example, sending older adults medication reminders through voice calls to mobile phones may be more effective than using smartphone notification systems, at least until smartphones become more widely adopted by the target group. Tablet computers (ie, iPads) offer an interesting example of this thinking. More seniors (aged ≥65 years) own tablets or e-book readers (27%) than own smartphones (18%), which is not the case with the general population [72]. Because of this, it may be more effective for change agents promoting consumer health technology use among seniors to focus on tablets rather than smartphones.

Our results indicate numerous significant differences in readiness across age groups. These results tell us that change agents should be cautious when extrapolating success in one age group to other age groups. Smartphones offer an example. Smartphone-based health applications may well be successful with younger consumers, but our results indicate that significantly fewer older consumers are ready to use smartphones for health applications. In fact, seniors are significantly less ready to adopt any technologies studied than their younger counterparts, with the exception of telephone voice calls.

This is not to say that we should completely eliminate from consideration any technologies with low readiness. As people age, as different age groups interact, and as the general adoption of core technologies improves, it is likely that readiness will improve. We advise keeping a close eye on the diffusion of the core technologies and timing the introduction and promotion of consumer health technologies applications to lag somewhat the diffusion of the core technology.

Change agents can also take steps to overcome a lack of familiarity. The role of trial use is helpful in countering unfamiliarity. Offering consumers easy, inexpensive ways to try an unfamiliar consumer health technology lowers adoption risk. Trials also allow consumers to experience the benefits of the consumer health technology, which will further increase the probability of adoption. Vicarious trials may also be beneficial. Vicarious trials are when another's use of an innovation substitutes for one's own trial use. These trials by close others may provide information that the potential adopter can use in evaluating the innovation [73]. It is important to choose the "other" carefully. Vicarious trials are more likely to be effective when the other party is an opinion leader [36] or is similar to the potential adopter or, in the case of health care, is a supporting caregiver.

Fundamentally, the decision to adopt or reject an innovation is typically based on an explicit or implicit cost-benefit analysis [36] in which the adopting unit (a consumer in this case) weighs the benefits of adopting an innovation against the perceived costs of doing so. Because of this, those interested in promoting adoption should understand how the potential adopter will view potential benefits and what the potential adopter will view as potential costs of adopting. These perceptions vary across individuals, and often across groups of individuals. In this paper, we focus on barriers to adoption readiness, which concerns the cost side of the adoption equation. In addition to monetary costs associated with adopting a consumer health technology, consumers also face real or perceived nonmonetary costs. These nonmonetary costs are the focus of our study.

Interestingly, the dominant barrier seems idiosyncratic to a particular technology. For example, in the case of kiosks, not being appropriate and not enjoying their use were the most frequently mentioned barriers (both 21%), but few participants reported not knowing how to use the kiosk as a barrier. In contrast, for smartphones, appropriateness was not an issue for a most respondents (10% reporting as a barrier), but not knowing how to use the technology was cited by 23% of the participants. Not surprisingly, the most frequently mentioned barriers for call centers were not being appropriate and not enjoying their use; few participants cited a lack of awareness, a lack of knowledge, or difficulty as being barriers to call center use. For consumer health technology designers and change agents, the message here is that it is important to not paint consumer health technologies with a broad brush with respect to barriers. The problematic barrier depends very much on the particular technology. Knowing which barriers apply to a particular consumer health technologies enables modifications in design or promotional activities that specifically address the dominant barriers. For example, making kiosks easier to use is not likely to substantially improve adoption. It would be more effective to address appropriateness. Further, it may be useful to consider patterns of barriers. Consider kiosks, blogs, and wikis. For all these technologies, security and appropriateness are frequently cited as barriers. It is possible that addressing the security barrier will also address the appropriateness barrier. More research is needed to verify this connection.

Turning attention to intergenerational differences in barriers (Figures 2 and 3), it is apparent that there is considerable variance across age groups in knowledge-based barriers. The main message from these findings is that age segmentation is important for those promoting use of consumer health technology. Seniors (aged >64 years) cite knowledge-based barriers more frequently than the other age groups for all technologies except voice phone calls. There are fewer differences in knowledge-based barriers between baby boomer and younger consumers. This is also true for motivation-based barriers. Awareness of intergenerational barriers is important for both consumer health technologies designers and change

agents. Those who seek to promote consumer health technologies use to seniors must first address the knowledge-based barriers before turning attention to motivation-based barriers. This is especially important when the consumer health technology faces sizable awareness barriers. Awareness precedes persuasion in the innovation-decision process; consumers who are unaware of an innovation are unlikely to adopt that innovation.

Knowledge-based barriers seem to be less of an issue for baby boomers, particularly with respect to awareness. Not knowing how to use a technology is more of an issue for baby boomers than is a lack of awareness. Thus, change agents interested in promoting consumer health technologies use by baby boomers should focus more on helping consumers understand how to use the consumer health technologies. Appropriateness for health care use of certain technologies is also an issue for many of our baby boomer participants. Approximately 20% or more of baby boomers reported appropriateness as a barrier for texting, kiosks, blogs, and wikis.

The results of our research may serve as a starting point for future investigations. For example, the knowledge and motivation barriers we identified could be used in predictive models of actual adoption. Another interesting possibility is to compare groups within baby boomers, such as exploring gender differences, differences in household income, or differences across age groups within the boomer generation. Additionally, our findings could be used to inform consumer health technology design and promotional message selections, which could then be tested for their impact on adoption. Finally, future research should investigate how baby boomers use consumer health technologies, particularly in comparison to other groups, which may further facilitate the appropriate design of consumer health technologies targeted to specific groups. For example, a recent study found that baby boomers reported a significantly higher tolerance for having more Web components on a page than younger generations suggesting that younger generations would be more likely to miss key information if a Web page fails to present information using a limited number of clear focal points that are located on the first screen [51]. Such findings have not been assessed for health care websites.

Limitations

As is the case with any study, there are a number of limitations to our research. First, we presented a limited number of barriers to technology use on the survey. Although our choices were based on the consumer health technologies and innovation adoption literature, these are evolving areas, so there may be important barriers that we did not investigate.

Second, there was no attempt to balance the number of respondents in each age group. Although demographic data on the sampling frame is proprietary and unavailable to us, we can extrapolate from the US Centers for Disease Control on prescription drug use and insurance coverage [74] and US Census data [75] to assess the representativeness of our data. To compute an expected distribution, we multiplied the population in each age group by the percentage in that population with health insurance. This gave us an expected number of individuals with health insurance. We then multiplied this number by the percentage of individuals in each group that take 1 or more prescription drugs. This gives us an expected number of people in each age group that both have insurance and take 1 or more prescription drugs. We then divided this number for each group by the sum of the groups to get the total proportion of individuals who have insurance and take prescription drugs. This yields a useful approximation of what proportion we should expect in each age group. Based on these calculations, it is likely that our sample overrepresents baby boomers (sample: 55% vs expected: 37%) and underrepresents the younger (sample: 15% vs expected: 22%) and older (sample: 30% vs expected: 41%) age groups. However, our underrepresentation of those aged 65 or older was still unexpected. Although this does not invalidate our results, it is an area of concern.

Third, the gender balance for the younger age group was skewed toward females, which was not the case with the other groups. To address the gender imbalance limitation, we computed chi-square statistics comparing “no problem” responses according to gender. The only significant difference ($P<.10$) was for wikis. Fifth, 39% of the contingency tables comparing consumer health technology barriers by age groups had small expected cell counts. We addressed this issue by using the Fisher’s exact test to test for significant differences, which should be kept in mind when interpreting our results.

Finally, caution should be taken in generalizing findings to uninsured populations. This sample was drawn from subscribers to a pharmacy benefit management company. Because the sample was drawn from enrollees in a pharmacy benefit management company, it is logical to assume all the respondents had some form of health insurance coverage. It is not known whether the respondent’s insurance company would pay for consumer health technologies used by the respondent. Furthermore, there are some implications that respondents may be able to bear some burden of consumer health technologies cost. Although income was not available for 31% of the sample, of those that reported income, 68% indicated they made over US $50,000 annually. In addition, 70% of the sample lived in the continental United States in a state east of the Mississippi River. This compares to 58% of the US population.

Conclusions

Fulfilling the promise of consumer health technologies to impact health care cost and enable health care consumers requires adoption by health care consumers. Given their large numbers and growing health care needs, it is particularly important to understand what consumer health technologies baby boomers are ready to adopt. It is also important to understand what barriers block adoption for technologies with low adoption readiness. This paper addressed these issues.

Based on our analysis, most baby boomers are ready to adopt some types of consumer health technology (telephone voice calls, websites, and email). They were equally split on being ready to adopt call centers, video conferencing, and texting for health purposes. Baby boomers seem reluctant to adopt podcasting, kiosks, smartphone apps, blogs, and wikis. Specific adoption barriers vary according to the technology. For example, appropriateness and enjoyment seem to be the biggest barriers

to adoption of call centers and texting, but knowing how to use and need for training are the biggest barriers for video conferencing. Further, baby boomers seem less ready to adopt some consumer health technologies than their younger counterparts, but are more ready to adopt than their elders. Differences between baby boomers and other consumers seem related to awareness, knowledge of how to use the technology, and the appropriateness and enjoyment of using technology for consumer health-related purposes.

Those interested in promoting use of consumer health technologies among baby boomers should consider these results when developing and choosing technologies, applications, and promotional tactics. Specifically, based on the results of this study, in combination with what is already known about innovation adoption, efforts to promote baby boomers' use of consumer health technologies should focus on applications where the benefits most clearly outweigh the costs of adoption. That is, consumer health technologies that are (1) familiar to the baby boomers, (2) have clearly perceived benefits, and (3) require relatively little effort to use. Such an approach addresses both the benefit and cost sides of the adoption equation. Further, this approach is consistent with innovation adoption theory. Familiarity is associated with perceptions of compatibility. Easily communicated benefits increase perceptions of relative advantage and result in demonstrability. Low effort reduces perceptions of complexity. However, it should be noted that the relative strength of perceptions varies according to the adopter and the innovation.

Acknowledgments

This study was supported by a research grant funded by Express Scripts, Inc.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Health care barriers instrument.

[PDF File (Adobe PDF File), 93KB - jmir_v16i9e200_app1.pdf]

References

1. Keehan S, Sisko A, Truffer C, Smith S, Cowan C, Poisal J, National Health Expenditure Accounts Projections Team. Health spending projections through 2017: the baby-boom generation is coming to Medicare. Health Aff (Millwood) 2008;27(2):w145-w155 [FREE Full text] [doi: 10.1377/hlthaff.27.2.w145] [Medline: 18303038]
2. Martini EM, Garrett N, Lindquist T, Isham GJ. The boomers are coming: a total cost of care model of the impact of population aging on health care costs in the United States by Major Practice Category. Health Serv Res 2007 Feb;42(1 Pt 1):201-218 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00607.x] [Medline: 17355589]
3. Zhang P, Zhang X, Brown J, Vistisen D, Sicree R, Shaw J, et al. Global healthcare expenditure on diabetes for 2010 and 2030. Diabetes Res Clin Pract 2010 Mar;87(3):293-301. [doi: 10.1016/j.diabres.2010.01.026] [Medline: 20171754]
4. Chatterji S, Kowal P, Mathers C, Naidoo N, Verdes E, Smith JP, et al. The health of aging populations in China and India. Health Aff (Millwood) 2008;27(4):1052-1063 [FREE Full text] [doi: 10.1377/hlthaff.27.4.1052] [Medline: 18607041]
5. National Institute on Aging, National Institutes of Health, US Department of State. Why Population Aging Matters: A Global Perspective. 2007 Mar 01. URL: http://www.nia.nih.gov/sites/default/files/WPAM.pdf [accessed 2014-08-25] [WebCite Cache ID 6S5eliDcc]
6. American Hospital Association, First Consulting Group. When I'm 64: How boomers will change health care. 2007. URL: http://www.aha.org/content/00-10/070508-boomerreport.pdf [accessed 2013-01-28] [WebCite Cache ID 6E11wlz2Z]
7. US Census Bureau. Selected characteristics of baby boomers 42 to 60 years old in 2006. 2006. URL: http://www.census.gov/population/age/publications/files/2006babyboomers.pdf [accessed 2014-08-25] [WebCite Cache ID 6S5eynObP]
8. US Securities and Exchange Commission. 2006. Oldest baby boomers turn 60! URL: http://www.sec.gov/spotlight/seniors/oldestboomers2007.htm [accessed 2014-08-25] [WebCite Cache ID 6S5f54KE4]
9. Sanderson E. Medical Careers Review. The rising health care needs of aging baby boomers URL: http://medical-careers-review.toptenreviews.com/the-rising-health-care-needs-of-aging-baby-boomers.html [accessed 2013-11-05] [WebCite Cache ID 6KtTPiaQe]
10. King DE, Matheson E, Chirina S, Shankar A, Broman-Fulks J. The status of baby boomers' health in the United States: the healthiest generation? JAMA Intern Med 2013 Mar 11;173(5):385-386. [doi: 10.1001/jamainternmed.2013.2006] [Medline: 23381505]
11. Forkner-Dunn J. Internet-based patient self-care: the next generation of health care delivery. J Med Internet Res 2003 Jun;5(2):e8 [FREE Full text] [doi: 10.2196/jmir.5.2.e8] [Medline: 12857664]
12. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]

13. Coughlin JF. Old Age, New Technology, and Future Innovations in Disease Management and Home Health Care. Home Health Care Management & Practice 2006 Apr 01;18(3):196-207. [doi: 10.1177/1084822305281955]
14. Johnson MO. The shifting landscape of health care: toward a model of health care empowerment. Am J Public Health 2011 Feb;101(2):265-270. [doi: 10.2105/AJPH.2009.189829] [Medline: 21164096]
15. Sundar S, Rice R, Kim H, Sciamanna C. Online health information: Conceptual challenges theoretical opportunities. In: Thompson T, Parrott R, Nussbaum J, editors. The Routledge Handbook of Health Communication. 2nd ed. New York: Routledge; 2011:181-202.
16. Pak R, Price MM, Thatcher J. Age-sensitive design of online health information: comparative usability study. J Med Internet Res 2009;11(4):e45 [FREE Full text] [doi: 10.2196/jmir.1220] [Medline: 19917549]
17. Wen KY, Kreps G, Zhu F, Miller S. Consumers' perceptions about and use of the internet for personal health records and health information exchange: analysis of the 2007 Health Information National Trends Survey. J Med Internet Res 2010;12(4):e73 [FREE Full text] [doi: 10.2196/jmir.1668] [Medline: 21169163]
18. Choi N. Relationship between health service use and health information technology use: Analysis of the us national health interview survey. Journal of Medical Internet Research 2011;13(2):e33. [doi: 10.2196/jmir.1753PMID:]
19. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
20. Agency for Healthcare Research and Quality. 2013. Consumer health IT applications URL: http://healthit.ahrq.gov/key-topics/consumer-health-it-applications [accessed 2014-08-25] [WebCite Cache ID 6S5fLAy0B]
21. Moorman C. An Innovation Adoption Approach to the Dissemination of Health Information to Consumers. Washington, DC: Agency for Health Care Policy Research; 1994.
22. Fogg B. Persuasive Technology, Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann; 2003.
23. Fogg B. Mobile Persuasion: 20 Perspectives on the Future of Behavior Change. Standford, CA: Stanford Captology Media; 2007.
24. Fogg BJ. Creating persuasive technologies: An eight-step design process. 2009 Presented at: Paper presented at: 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, CA.
25. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2011 Apr 8:e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
26. Jones R. Development of a questionnaire and cross-section survey of patient ehealth readiness and ehealth inequalities. Journal of Medical Internet Research 2013;15(9):e9. [doi: 10.2196/med20.2559]
27. Mihailidis A, Cockburn A, Longley C, Boger J. The acceptability of home monitoring technology among community-dwelling older adults and baby boomers. Assist Technol 2008;20(1):1-12. [doi: 10.1080/10400435.2008.10131927] [Medline: 18751575]
28. Steel DM, Gray MA. Baby boomers' use and perception of recommended assistive technology: a systematic review. Disabil Rehabil Assist Technol 2009 May;4(3):129-136. [doi: 10.1080/17483100902767175] [Medline: 19378208]
29. Selwyn N, Gorard S, Furlong J, Madden L. Older adults' use of information and communications technology in everyday life. Ageing and Society 2003;23(5):561-582. [doi: 10.1017/S0144686X03001302]
30. Zickuhr K. Generations and their gadgets. Washington, DC: Pew Internet & American Life Project; 2011 Feb 03. URL: http://www.pewinternet.org/2011/02/03/generations-and-their-gadgets/ [accessed 2014-08-25] [WebCite Cache ID 6S5fjVtkw]
31. Smith A. Smartphone ownership 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jun 05. URL: http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/ [accessed 2014-08-25] [WebCite Cache ID 6S5fsrxBm]
32. Martin LG, Freedman VA, Schoeni RF, Andreski PM. Health and functioning among baby boomers approaching 60. J Gerontol B Psychol Sci Soc Sci 2009 May;64(3):369-377 [FREE Full text] [doi: 10.1093/geronb/gbn040] [Medline: 19299256]
33. MetLife Insurance Company Mature Market Institute. 2011. The MetLife study of caregiving costs to working caregivers: Double jeopardy for baby boomers caring for their parents URL: https://www.metlife.com/assets/cao/mmi/publications/studies/2011/mmi-caregiving-costs-working-caregivers.pdf [WebCite Cache ID 6E11YJXbY]
34. Freedman V, Calkins M, DeRosiers R, van Haitsma K. Barriers to implementing technology in residential long-term care settings - executive summary. Washington, DC: US Department of Health and Human Services; 2005. URL: http://aspe.hhs.gov/daltcp/reports/2005/techbarr.htm [accessed 2014-08-25] [WebCite Cache ID 6S5g6Q8P0]
35. Tak SH, Benefield LE, Mahoney DF. Technology for long-term care. Res Gerontol Nurs 2010 Jan;3(1):61-72 [FREE Full text] [doi: 10.3928/19404921-20091103-01] [Medline: 20128544]
36. Rogers EM. Diffusion of Innovations. New York: Free Press; 2003.
37. Morris MG, Venkatesh V. Age differences in technology adoption decisions: implications for a changing work force. Personnel Psychology 2000 Jun;53(2):375-403. [doi: 10.1111/j.1744-6570.2000.tb00206.x]
38. Van Dijk J, Hacker K. The digital divide as a complex and dynamic phenomenon. The Information Society 2003;19:315-326. [doi: 10.1080/01972240390227895):315-326]
39. McLeod E. The use (and disuse) of mobile phones by baby boomers. International Journal of Emerging Technologies and Society 2009;7(1):28-38.

40. Zickhur K. Generations 2010. Washington, DC: Pew Internet & American Life Project; 2010 Dec 16. URL: http://www.pewinternet.org/2010/12/16/generations-2010/ [accessed 2014-08-25] [WebCite Cache ID 6S5gUAC9G]
41. Jones R. Development of a questionnaire and cross-sectional survey of patient ehealth readiness and ehealth inequalities. In: Medicine 2.0. 2013 Presented at: Medicine 2.0; Sept 23-24, 2013; London, UK p. 20.
42. Sharit J, Czaja S, Perdomo D, Lee C. A cost-benefit analysis methodology for assessing product adoption by older user populations. Applied Ergonomics 2004;35(2):81-92. [doi: 10.1016/j.apergo.2003.12.003PMID:]
43. Cody MJ, Dunn D, Hoppin S, Wendt P. Silver surfers: Training and evaluating internet use among older adult learners. Communication Education 1999 Oct;48(4):269-286. [doi: 10.1080/03634529909379178]
44. Craik F, Salthouse T. The Handbook of Aging and Cognition. 3rd ed. New York: Psychology Press; 2008.
45. Melenhorst AS, Rogers WA, Bouwhuis DG. Older adults' motivated choice for technological innovation: evidence for benefit-driven selectivity. Psychol Aging 2006 Mar;21(1):190-195. [doi: 10.1037/0882-7974.21.1.190] [Medline: 16594804]
46. Emani S, Yamin C, Peters E. Patient perceptions of a personal health record: A test of the diffusion of innovation model. Journal of Medical Internet Research 2012;14(6):e150. [doi: 10.2196/jmir.2278PMID:]
47. Chau PYK, Hu PJ. Information Technology Acceptance by Individual Professionals: A Model Comparison Approach. Decision Sciences 2001 Dec;32(4):699-719. [doi: 10.1111/j.1540-5915.2001.tb00978.x]
48. Melenhorst A, Bouwhuis D. When do older adults consider the internet? An exploratory study of benefit perception. Gerontechnology 2004;3(2):89-101. [doi: 10.4017/gt.2004.03.02.004.00]
49. Strother JB. Call centers in health care: Effect on patient satisfaction. 2006 Presented at: IEEE International Professional Communication Conference; Oct 23-25, 2006; New York.
50. Dickinger A, Arami M, Meyer D. The role of perceived enjoyment and social norm in the adoption of technology with network externalities. Eur J Inf Syst 2008 Feb;17(1):4-11. [doi: 10.1057/palgrave.ejis.3000726]
51. Djamasbi S, Siegel M, Skorinko J, Tullis T. Online viewing and aesthetic preferences of generation Y and the baby boom generation: testing user web site experience through eye tracking. International Journal of Electronic Commerce 2011 Jun 29;15(4):121-158. [doi: 10.2753/JEC1086-4415150404]
52. Lou H, Chau PY, Li D. Understanding individual adoption of instant messaging: an empirical investigation. Journal of the Association for Information Systems 2005;6(4):5.
53. Hsieh JA, Rai A, Keil M. Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. MIS Quarterly 2008;32(1):97-126.
54. Zeman L, Johnson D, Arfken C, Smith T, Opoku P. Lessons learned: Challenges implementing a personal digital assistant (pda) to assess behavioral health in primary care. Families, Systems & Health: The Journal of Collaborative Family HealthCare 2006;24(3):286-298. [doi: 10.1037/1091-7527.24.3.286PMID:]
55. Katz SJ, Moyer CA. The emerging role of online communication between patients and their providers. J Gen Intern Med 2004 Sep;19(9):978-983 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.30432.x] [Medline: 15333064]
56. Phang C, Sutanto J, Kankanhalli A, Yan L, Tan B, Hock-Hai T. Senior citizen's acceptance of information systems: A study in the context of e-government services. IEEE Transactions on Engineering Management 2006;53(4):555-569. [doi: 10.1109/TEM.2006.883710PMID:]
57. 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]
58. Matthews AK, Sellergren SA, Manfredi C, Williams M. Factors influencing medical information seeking among African American cancer patients. J Health Commun 2002;7(3):205-219. [doi: 10.1080/10810730290088094] [Medline: 12166874]
59. Zulman D. Trust in the internet as a health resource among older adults: Analysis of data from a nationally representative study. Journal of Medical Internet Research 2011;13(1):e19. [doi: 10.2196/jmir.1552PMID:]
60. Karahanna E, Agarwal R, Angst C, Corey M. Reconceptualizing compatibility beliefs in technology acceptance research. MIS Quarterly 2006;30(4):781-804. [doi: 10.2307/25148754]
61. Van Slyke C, Belanger F, Johnson R, Hightower R. Gender-based differences in consumer e-commerce adoption. Communications of the Association for Information Systems 2010;26(2):17-34.
62. van der Heijden H. User acceptance of hedonic information systems. MIS Quarterly 2004;28(4):1.
63. Thong JY, Hong S, Tam KY. The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance. International Journal of Human-Computer Studies 2006 Sep;64(9):799-810. [doi: 10.1016/j.ijhcs.2006.05.001]
64. Brown SA, Venkatesh V. Model of adoption of technology in the household: A baseline model test and extension incorporating household life cycle. MIS Quarterly 2005;29(3):399.
65. Wright KB. Researching internet-based populations: Advantages and disadvantages of on-line survey research, on-line questionnaire authoring software packages, and web survey services. Journal of Computer-Mediated Communication 2005;10(3):11.
66. Ellis PD. The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results. Cambridge, UK: Cambridge University Press; 2010.
67. Huck SW. Reading Statistics and Research. Boston, MA: Pearson Education, Inc; 2012.

68. Wejnert B. Integrating models of diffusion of innovations: A conceptual framework. Annual Review of Sociology 2002;28:297-326.
69. Smith A. 35% of Americans Adults Own a Smartphone. Washington, DC: Pew Internet & American Life Project; 2011 Jul 11. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Smartphones.pdf [accessed 2014-08-25] [WebCite Cache ID 6S5hd9eWC]
70. Nielsen. 2012. New US Smartphone Growth by Age and Income URL: http://blog.nielsen.com/nielsenwire/online_mobile/survey-new-u-s-smartphone-growth-by-age-and-income/ [WebCite Cache ID 6E137KMMU]
71. Duggan M. Cell phone activities 2013. 2013 Oct 19. URL: http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/ [accessed 2014-08-25] [WebCite Cache ID 6S5hnPsxs]
72. Smith A. Older adults and technology use. Washington, DC: Pew Internet & American Life Project; 2014 Apr 03. URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2014-08-25] [WebCite Cache ID 6S5hssfD9]
73. Karahanna E, Straub D, Chervany N. Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Quarterly 1999;23(2):183-213.
74. Health, United States, 2013: With Special Feature on Prescription Drugs. Hyattsville, MD: National Center for Health Statistics; 2014. URL: http://www.cdc.gov/nchs/data/hus/hus13.pdf [accessed 2014-07-31] [WebCite Cache ID 6RToNSnhl]
75. Karahanna E, Straub D, Chervany N. Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Quarterly 1999;23(2):183-213.

Abbreviations

SMS: short message service

Edited by G Eysenbach; submitted 05.11.13; peer-reviewed by A Fruhling, S Ryu, R Jones; comments to author 28.11.13; revised version received 15.06.14; accepted 03.07.14; published 08.09.14.

Please cite as:
LeRouge C, Van Slyke C, Seale D, Wright K
Baby Boomers' Adoption of Consumer Health Technologies: Survey on Readiness and Barriers
J Med Internet Res 2014;16(9):e200
URL: http://www.jmir.org/2014/9/e200/
doi:10.2196/jmir.3049
PMID:25199475

Original Paper

Applying Computer Adaptive Testing to Optimize Online Assessment of Suicidal Behavior: A Simulation Study

Derek Paul De Beurs[1,2], MSc clinical psychology, MSc Biostat; Anton LM de Vries[3], PhD; Marieke H de Groot[1], PhD; Jos de Keijser[4], PhD; Ad JFM Kerkhof[1,2], PhD

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

[2]EMGO Institute for Health and Care Research, Amsterdam, Netherlands

[3]Perziq.com, Amsterdam, Netherlands

[4]GGZ Foundation for Mental Health Care Friesland and Groningen University, Groningen, Netherlands

Corresponding Author:
Derek Paul De Beurs, MSc clinical psychology, MSc Biostat
EMGO Institute for Health and Care Research
van der Boechorststraat 1
Amsterdam,
Netherlands
Phone: 31 205982589
Fax: 31 205982588
Email: dp.de.beurs@vu.nl

Abstract

Background: The Internet is used increasingly for both suicide research and prevention. To optimize online assessment of suicidal patients, there is a need for short, good-quality tools to assess elevated risk of future suicidal behavior. Computer adaptive testing (CAT) can be used to reduce response burden and improve accuracy, and make the available pencil-and-paper tools more appropriate for online administration.

Objective: The aim was to test whether an item response–based computer adaptive simulation can be used to reduce the length of the Beck Scale for Suicide Ideation (BSS).

Methods: The data used for our simulation was obtained from a large multicenter trial from The Netherlands: the Professionals in Training to STOP suicide (PITSTOP suicide) study. We applied a principal components analysis (PCA), confirmatory factor analysis (CFA), a graded response model (GRM), and simulated a CAT.

Results: The scores of 505 patients were analyzed. Psychometric analyses showed the questionnaire to be unidimensional with good internal consistency. The computer adaptive simulation showed that for the estimation of elevation of risk of future suicidal behavior 4 items (instead of the full 19) were sufficient, on average.

Conclusions: This study demonstrated that CAT can be applied successfully to reduce the length of the Dutch version of the BSS. We argue that the use of CAT can improve the accuracy and the response burden when assessing the risk of future suicidal behavior online. Because CAT can be daunting for clinicians and applied scientists, we offer a concrete example of our computer adaptive simulation of the Dutch version of the BSS at the end of the paper.

(J Med Internet Res 2014;16(9):e207) doi:10.2196/jmir.3511

KEYWORDS

suicide; psychometrics; computing methodologies; Internet; suicidal ideation; risk assessment

Introduction

Background

Suicide ideation is defined as the presence of thoughts, plans, and wishes in an individual to end his/her own life [1]. Assessment of suicide ideation is argued to be important because it may precede an attempt and it could provide information on the seriousness and lethality of the suicidal intention [2]. The Beck Scale for Suicide Ideation (BSS) is the 19-item self-report version of the Scale for Suicide ideation, a systematic interviewing tool developed for the assessment of suicide ideation and risk of future suicidal behavior [1,3]. The BSS inquires about suicidal thoughts and attitudes of subjects toward

them. Because the BSS is widely accepted and has strong psychometric properties, the BSS is frequently used in research and clinical practice to assess risk of future suicidal behavior [4].

The role of the Internet in suicide prevention is increasing [5,6]. Online self-help interventions are offered to recover from suicide ideation [7], researchers collect data on suicidal behavior in real time via mobile applications [8,9], and mental health institutions monitor suicidal behavior of patients via online questionnaires [10]. Attrition in online interventions and studies is a well-known problem [11]. To optimize online assessment of patients and thereby limit attrition, there is a need for a shorter and more accurate questionnaire to assess risk of suicidality. Traditional pencil-and-paper mental health questionnaires have a large respondent burden because they require patients to answer questions that do not provide any additional information. In our example, the BSS has 19 items and a score range from 0 to 38. However, a prospective study showed that subjects who scored >2 were 7 times more likely to show future suicidal behavior than those that scored 2 or less [2]. It seems that when assessing risk of future suicidal behavior, if a subject scores >2 there is no need to complete the other items. Computer adaptive testing (CAT) [12] allows us to reduce the number of items in a questionnaire without losing discriminatory validity. Its applicability has been demonstrated in depression [13] and anxiety [14], but not yet in the assessment of the risk of suicidal behavior. Because answering several items on suicidal behavior online can be burdensome for patients, especially at baseline or on intake [15,16], a shorter assessment of suicidal ideation is preferable.

Computer Adaptive Testing

CAT is possible because of item response theory (IRT) and the wide availability of the Internet. IRT is based on a computerized iterative process that, for each item, regresses the patient's response on a latent trait score (theta; suicide ideation in our example), the estimated value of which maximizes the likelihood of the patient's pattern of responses [17]. More concretely, a patient answers an item online and based on the response to that single item, the computer follows an IRT-based algorithm that offers the patient the next most informative item. After the patient's score has been estimated at the predefined level of precision, no more items are administered. So, only the fewest possible items are offered per patient, resulting in less respondent burden and even more accurate outcomes [17]. Due to these advantages, IRT and CAT are currently being applied in health outcomes research to develop or improve existing measures. For example, The Patient-Reported Outcomes Measurement Information System (PROMIS), a large project funded by the National Institute of Health to develop valid, reliable, and standardized questionnaires to measure patient outcomes [12] relies heavily on IRT and CAT modeling.

Current Study

The goal of the current study was to investigate whether we can use CAT to shorten the BSS without losing discriminatory validity. We followed the 5 steps of the psychometric analysis plan as used in the PROMIS project [12]. We provided descriptive statistics, evaluated the assumptions for the IRT, fitted an IRT model to our data, tested for item bias, and stimulated a CAT on our data. Because this paper is the first to apply IRT and CAT in the field of suicidology, we explain every step of the process in depth. We have ended this paper with a concrete example of a shortened version of the Dutch version of the Beck Scale for Suicide Ideation (BSS-NL). An overview of the 5 psychometric steps are:

1. Descriptive statistics
2. Testing of assumptions about the IRT model
3. Fitting of the IRT model to the data
4. Evaluating differential item functioning (DIF)
5. Computer adaptive testing

Methods

Measurement Procedure

We used the data collected at baseline in the Dutch Professionals in Training to STOP suicide (PITSTOP suicide) study [18]. In the study, mental health professionals were trained in guideline adherence via an e-learning-supported Train-the-Trainer program. Although the intervention was aimed at improving suicide prevention skills of professionals [19], the primary outcome of the study was a change in suicide ideation of patients as measured with the Dutch version of the BSS, the BSS-NL. The BSS was translated into Dutch making use of forward and back translation, and was recently used in a clinical trial study [20]. The preferred mode of data collection among patients was via the routine outcome monitoring (ROM) system, an online system by which data on the effectiveness of treatment in everyday clinical practice are systematically collected [3]. In departments not using ROM, graduate students and/or research assistants used paper-and-pencil questionnaires to collect data. The main *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition) *DSM-IV* diagnosis of each patient was assessed at intake via a structured interview by a mental health professional.

All eligible patients were informed about the study and all provided informed consent.

Software

All analyses were performed in R [21]. Descriptive statistics and principal components analysis (PCA) were obtained via the psych package [22]. The confirmatory factor analysis (CFA) models were estimated using the lavaan package [23]. Graded response models (GRM) were fitted using the latent trait modeling (LTM) package [24]. The mokken package was used to estimate monotonicity [25]. DIF was checked via the lordif package [26]. The CatIRT package was used for the CAT simulation [27].

We followed the 5 steps as used in the PROMIS study as listed in the Introduction.

Step 1: Descriptive Statistics

Descriptive statistics were described. Cronbach alpha [28] was used to test internal consistency reliability, with .8 as acceptable minimum.

Step 2: Testing Assumptions About the Item Response Theory Model

Before fitting the IRT model, the basic assumptions for IRT models were tested. The assumptions for IRT are unidimensionality, local independency, and monotonicity [17].

For unidimensionality, we performed a PCA to examine whether a 1-dimensional test explained at least 20% of the variance and whether the ratio of explained variance of the first factor to the second was 4 or higher [29]. Next, we used a CFA to test unidimensionality by using various fit indexes [12]. The residual matrix produced by this single factor CFA was used to test the second assumption, local independence. Correlations >.2 were flagged and considered as possible violations of local independence [12]. Finally, monotonicity was examined by fitting a nonparametric IRT model that resulted in IRT probability curves. Nonmonotonic items with a scalability coefficient <.3 [25] were flagged [12] and described.

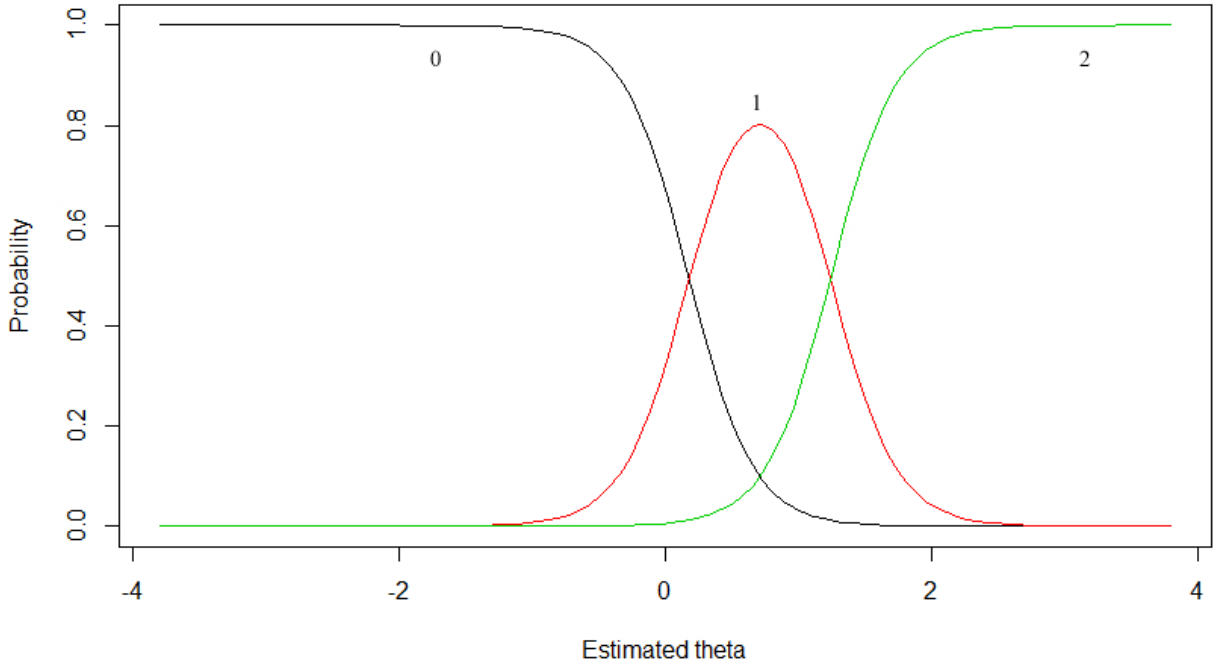
Step 3: Fit an Item Response Theory Model to the Data

There are a great number of different IRT models [30]. For questions with ordered-response categories, the GRM [31] was proposed. Because the BSS has 3 ordered-response options (0, 1, and 2; a higher score represents a higher level of suicide ideation), we fitted a GRM to our data.

As an introduction to the GRM, we provided an example of a GRM for item 7 (frequency of thinking about suicide) of the BSS. Figure 1 shows a function per response category (0, 1, and 2) that corresponds to the chance that a participant chooses that option given a certain score of theta. Because each item of the BSS has 3 response options (0, 1, and 2) per item, 3 curves are presented. The combined chance of all 3 response curves at any certain level of theta is always 1. In other words, in Figure 1, if a patient has a theta of –2, the chance that a patient will choose option 0 is approximately 1. If a patient has a theta of 1.5, he/she has an approximately zero chance of endorsing option 0, 0.65 chance of endorsing option 1, and 0.35 (1–[0+0.65]) of endorsing option 2. Patients that score theta ≥2 will most likely endorse option 2. Every single item is defined by a discrimination parameter (alpha) and 2 location parameters (β_1 and β_2). The item parameters of the current example are estimated to be α=4.117, β_1=0.171, and β_2=1.243. The discrimination parameter reflects the true difference in theta per item and is comparable to a factor loading. The betas (threshold parameters) indicate the location on the scale of the latent continuum where the item best discriminates among individuals.

To evaluate the fit of the IRT model to the data, category response curves (CRC) for each single item, such as in Figure 1, were plotted.

Figure 1. Example of category response curve for item 7 (frequency of thinking about suicide). Number and colors (0: black; 1: red; 2: green) reflect answer options.



Step 4: Evaluation of Differential Item Functioning

Differential item functioning (DIF) exists if patients from 2 groups (eg, men and women) who are equal in terms of the level of theta do not have the same probability of endorsing a test item [32]. Similarly to the PROMIS study, 4 important covariates were considered: gender, age (18-49 years, 50-69 years), education (low level of education/college or advanced degree), and method of administration (computer vs paper and pencil). The IRT item parameters (discrimination and threshold parameters) are assumed to be linear invariant with respect to group membership. Any difference found in CRC then indicates that patients with the same level of theta but from different groups have a different probability of endorsing an item. Items that show DIF at an alpha level of 0.01 were flagged. Because statistical power is dependent on sample size, trivial but nonzero differences are likely to be found to be significant in our large sample. Therefore, we also reported effect sizes to further

investigate the magnitude of the DIF. McFadden's pseudo R^2 <.13 are negligible, and effect sizes between .13 and .26 are moderate [26].

Step 5: Computer Adaptive Testing

The Package CatIRT performs a post hoc CAT. The IRT parameters obtained in step 3 were used for our CAT simulation unless the DIF analysis suggested using different parameters for subgroups of patients. As a starting point, we set an entry level, which is normally chosen to be 0 [13]. The first item to be selected was the item with the most information at this initial level of suicide ideation. The next item was selected via maximum Fisher information method, related to the theta estimated on basis of the just-selected item and the response to that item. Finally, we determined a stopping rule. As a stopping rule, we used a value of theta that reflects BSS >2 (θ >-1). Our CAT was terminated if the confidence interval surrounding an estimate of theta was fully within 1 of the categories (elevated risk/no elevated risk). We used a confidence interval of 99%. Because questionnaires in mental health care tend to peak at the relative higher levels of the clinical outcome [13,17,33], we also added a second stopping rule to prevent subjects without suicide ideation from having to complete all 19 items. The second stopping rule was use a maximum of 6 items. We compared the classification differences when using none, 1, or 2 stopping rules.

Results

Overview

We applied the CAT to the 505 patients within the PITSTOP suicide trial that completed the full 19 items. Initially, data were collected via the ROM. After the start of the study, it appeared difficult for most departments to collect our data via the ROM. In total, only 43% (217/505) of the data was collected using the ROM. As an alternative, research assistants and clinicians were instructed to complete the questionnaire via paper and pencil. Of the 505 patients, 93 (18.4%) patients had a total BSS score of 0, and 254 (50.3%) had a score <8; 128 (25.3%) had depression as their primary diagnosis and 50 (9.9%) had a personality disorder. Mean age was 42 (SD 9.2) years. At baseline, 183 of 505 (36.2%) patients stated they had attempted suicide at least once.

Step 1: Descriptives

The overall Cronbach alpha was .94. Average score on the BSS was 10.4 (SD 9.4). As Table 1 shows, removing 1 item did not lead to a substantial improvement of the internal consistency. The item-rest or remainder correlations (R_{rest}) were also satisfactory.

Table 1. Descriptive statistics of the single items of the BSS-NL.

Single item content	Category			Mean (SD)	Cronbach α	R_{rest}
	0	1	2			
1. Wish to live	242	195	68	0.66 (0.79)	0.94	.67
2. Wish to die	199	196	110	0.82 (0.70)	0.93	.74
3. Reasons living/dying	278	157	70	0.59 (0.84)	0.93	.69
4. Desire to kill oneself	246	171	88	0.69 (0.76)	0.93	.80
5. Passive suicidal desire	248	174	83	0.67 (0.77)	0.93	.63
6. Duration of suicide ideation	319	113	73	0.51 (0.90)	0.93	.76
7. Frequency of thinking about suicide	310	160	35	0.46 (0.96)	0.93	.80
8. Acceptance of idea of suicide	285	153	67	0.57 (0.85)	0.93	.73
9. Control over suicide action	344	140	21	0.36 (1.08)	0.93	.68
10. Reasons for not committing suicide	314	138	53	0.48 (0.93)	0.93	.70
11. Reasons for wanting to commit suicide	223	47	235	1.02 (0.67)	0.93	.48
12. Specific plan to commit suicide	318	118	69	0.51 (0.91)	0.93	.71
13. Access to suicide method	335	31	139	0.61 (0.82)	0.93	.60
14. Courage/ability to commit suicide	279	152	74	0.59 (0.83)	0.93	.74
15. Expectation to commit suicide	318	153	34	0.44 (0.98)	0.93	.76
16. Preparations for suicide	391	89	25	0.28 (1.19)	0.93	.63
17. Writing of suicide note	394	72	39	0.30 (1.16)	0.93	.49
18. Final acts in anticipation of death	354	106	45	0.39 (1.04)	0.93	.37
19. Conceal ideation	308	123	74	0.54 (0.88)	0.93	.47

Step 2: Testing of Assumptions of the Item Response Theory Model

When fitting a 1-factor PCA, we found that 50% of the proportional variance was explained by the first factor. The ratio between a 1- and a 2-factor model indicated that the first factor model explained 14 times more variance than the second factor. When fitting a confirmatory analysis we found a comparative fit index of 0.999, a Tucker-Lewis index of 0.989, a root-mean-square error of approximation of 0.045 (90% CI 0.038-0.053), and a standardized root-mean-square residual of 0.059.

Step 3: Fitting of a Graded Response Model

Overview

Table 2 shows that all 19 items had an alpha higher than 1. Item 7 (frequency of thinking about suicide) seems to discriminate best between patients with a higher or lower level of suicidal ideation, as indicated by the high alpha of 4.117.

Table 2. Graded response model parameters for the Dutch version of the Beck Scale for Suicide Ideation (BSS-NL).

Item and content	Parameter		
	α	β_1	β_2
1. Wish to live	2.366	0.029	0.556
2. Wish to die	3.197	–0.159	0.180
3. Reasons living/dying	3.036	0.034	0.937
4. Desire to kill oneself	4.082	–0.123	0.691
5. Passive suicidal desire	2.270	–0.188	0.898
6. Duration of suicide ideation	3.434	0.276	1.054
7. Frequency of thinking about suicide	4.117	0.171	1.243
8. Acceptance of idea of suicide	3.437	0.071	0.985
9. Control over suicide action	3.165	0.386	1.587
10. Reasons for not committing suicide	3.048	0.263	1.258
11. Reasons for wanting to commit suicide	1.557	0.001	0.100
12. Specific plan to commit suicide	3.003	0.305	1.122
13. Access to suicide method	2.479	0.550	0.616
14. Courage and ability to commit suicide	3.780	0.045	0.932
15. Expectation to commit suicide	3.825	0.232	1.369
16. Preparations for suicide	2.532	0.754	1.815
17. Writing of suicide note	1.786	1.016	1.952
18. Final acts in anticipation of death	1.098	0.887	2.414
19. Hide, conceal, or lie about suicide ideation	1.436	0.334	1.466

Category Response Curves

Of the 19 items, 17 showed CRC plots as expected. Items 11 (reasons for wanting to commit suicide) and 13 (access to suicide method) showed CRCs that warranted extra inspection. Table 3 shows the mean overall theta of participants per response option for 3 different items: item 7, which had a good CRC, and for items 11 and 13, which showed unsatisfactory CRCs. For items 11 and 13, the difference in mean theta for responses 1 and 2 was small and their confidence intervals overlapped, indicating that a higher score on 1 of these items does not necessarily reflect a higher level of suicidal ideation.

Table 3. The mean theta of patients that endorsed response option 0, 1, and 3 for items 7, 11, and 13.

Item and response	Mean θ	95% CI
7. Frequency of thinking about suicide		
0	–1.7	–2.0, –1.4
1	1.4	1.3, 1.5
2	2.6	2.4, 2.8
11. Reasons for wanting to commit suicide		
0	–2.3	–2.6, –1.9
1	0.9	0.6, 1.2
2	1.0	0.8, 1.1
13. Access to suicide method		
0	–1.5	–1.7, –1.2
1	1.3	1.1, 1.6
2	1.6	1.4, 1.7

Step 4: Differential Item Functioning

No items were flagged for DIF when analyzing differences in gender, age, or education. When analyzing the effect for the administration method, 7 items were flagged. However, R^2 were all <.13.

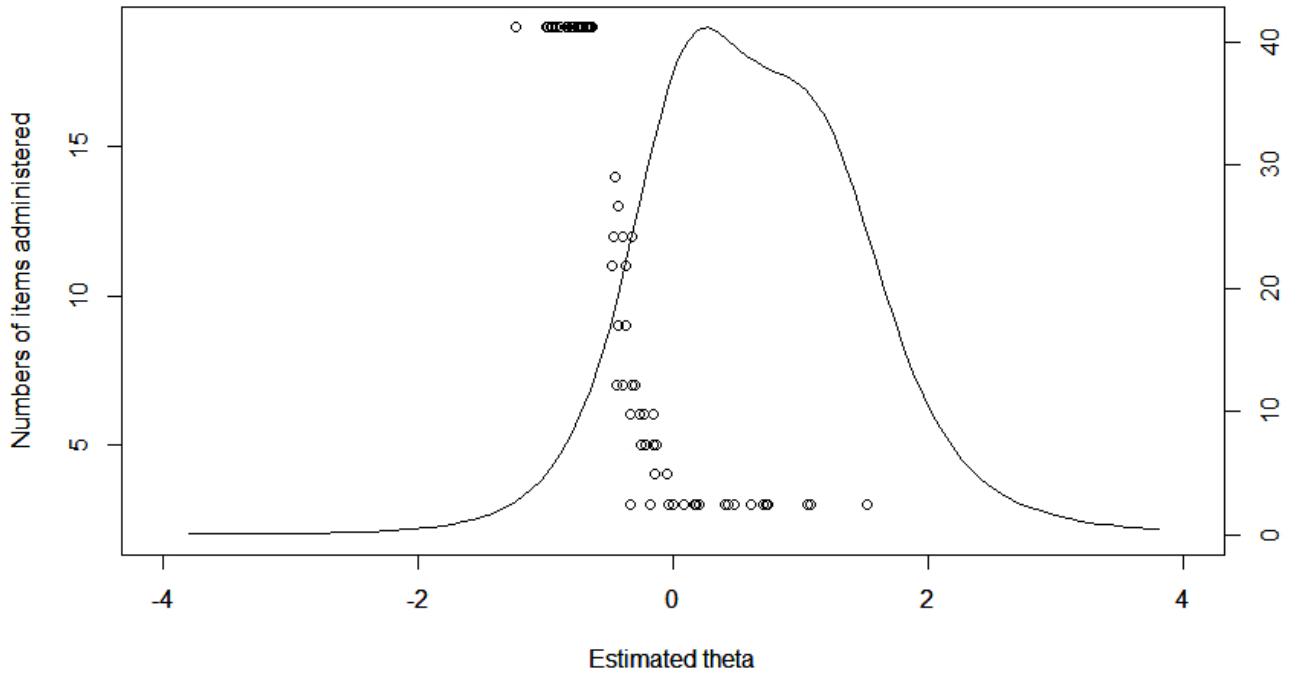
Step 5: Computer Adaptive Testing

When administrating all 19 items, 345 patients were classified as being at risk (Table 4). When allowing the number of items to vary between 3 and 19, CAT simulations showed that, on average, 10 items were sufficient to meet the same classification as the first model. For a large number of patients with a low trait of suicidal ideation, all items were exhausted before the stopping rule was met (Figure 2). When using a maximum of 6 items, 336 instead of 345 patients were classified as having an elevated risk (Table 4).

Table 4. Classification of risk for several stopping rules for Beck Scale for Suicide Ideation scores >2.

Stopping rule		Mean (SD)	Number of patients with low risk of future suicidal behavior	Number of patients with elevated risk of future suicidal behavior
Min items	Max items			
19	19	19 (0)	160	345
3	19	9.7 (7.7)	162	343
3	6	4.2 (1.4)	169	336

Figure 2. Relationship between level of theta and the number of administered items under stop rule 1. The curve represents test information as a function of theta.



Discussion

Principal Findings

Our simulation showed that an IRT model can be fitted to the BSS-NL and that CAT can successfully be applied to reduce the length of the BSS-NL when assessing risk of future suicidal behavior. PCA and confirmatory factor analysis found the scale to be highly unidimensional. No local independence or violation of monotonicity was found. Therefore, all assumptions for IRT modeling were met. For 17 of 19 items, IRT parameters were satisfactory indicating that most items are well suited to provide differential information on a patient's level of suicidal ideation. When using CAT with a maximum of 6 items, only 9 of 505 (1.7%) patients were classified in a different category when compared to the classification under all 19 items. Importantly, this simulation demonstrated that CAT makes it possible to administer only 4 items, on average, instead of the full 19 without losing discriminatory validity.

Improvement of the Items 11 and 13

We found items 11 (reasons for wanting to commit suicide) and 13 (access to suicide method) to show unsatisfactory item parameters. Further inspection revealed that, for both items, patients with comparable levels of suicide ideation were equally likely to endorse either option 1 or 2. For example, consider item 13. Our data showed that patients with low suicidal trait were more likely to endorse option 0 (have no access to means) and patients with higher levels of suicidality were equally likely to endorse option 1 (it takes time to find means) or 2 (I have access to means). Due to this overlap, patients with the same level of suicidal ideation might end up with different summed total scores. Therefore, when using the full-scale version of the BSS, we advise rephrasing the response options of both items, offering them as dichotomous items or excluding them.

Strengths and Limitations

Because this is a simulation study, real-time CAT studies are needed to determine the most accurate item parameters. Few clinical studies have implemented CAT in real time, but those studies that did showed a good comparison with simulation studies (eg, [34]). Next, it is necessary to compare the parameters of the current study with, for example, data collected with the original English-language version of the BSS. For our simulation, we used a fixed theta as cut-off score instead of the established BSS score >2. Future prospective studies must examine the most plausible theta cut-off to predict elevated risk of suicidal behavior. Also, we had no long-term follow-up data on whether patients actually engaged in any suicidal behavior after the assessment. Therefore, we were not able to compare the predictive validity of the CAT with the predictive validity of the full test. An additional limitation of CAT approaches might be that CAT data would not be comparable to normative data. By standardizing outcomes as done in meta-analysis [35], scores assessing the same outcome but measured in a variety of ways can still be compared.

With the BSS-NL, it seems to be difficult to investigate small differences in patients with a low suicidal trait. This has been found more often in mental health assessments [13,17,33]. A hybrid CAT approach, such as the 2-stage semiadaptive testing strategy recommended by Choi et al [35], might also be appropriate and result in even more accurate classification.

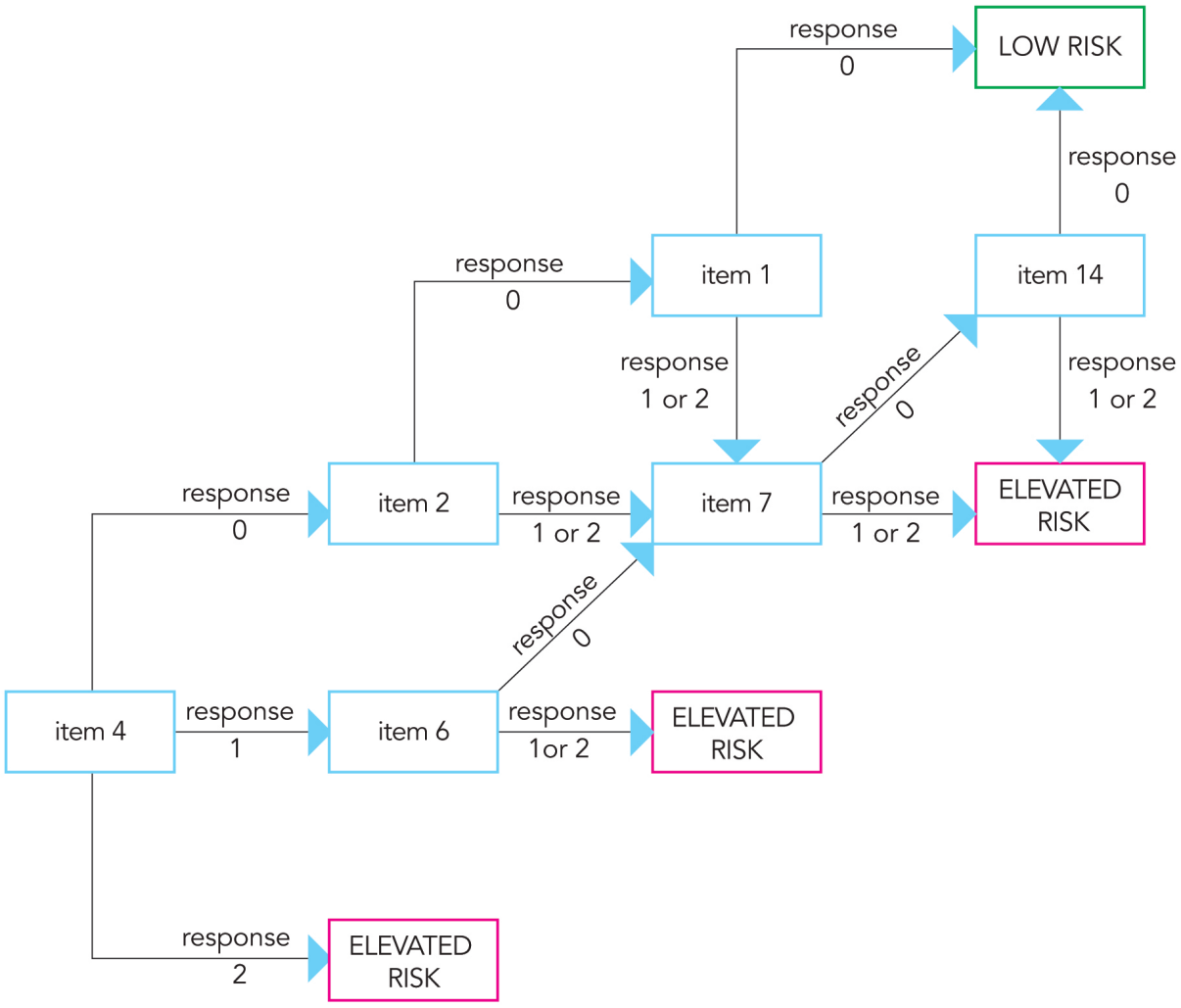
Finally, although wireless Internet and reliable hardware are widely available, the current state of ICT in (Dutch) mental health care reduces the feasibility of large-scale CAT implementation. Even for our normal (non-CAT) assessments, many of the research assistants within our study had to resort to paper-and-pencil testing because computerized testing was technically not possible. Obviously, this precludes CAT.

Strengths of this study pertain to its large sample size; 505 patients from various psychiatric departments completed the BSS. Therefore, the external validity of the findings is considerable. Another strength of the current paper is the application of modern psychometric techniques within the field of clinical psychology/psychiatry. For several reasons, such as lack of interest in new techniques and insufficient mathematical training, the integration of new techniques in psychology/psychiatry has been suboptimal at least [36]. By thoroughly explaining every step of our analysis and by focusing on the actual application of IRT and CAT in the clinical field, this paper hopes to stimulate the use of contemporary psychometric techniques.

Concrete Example of the Computer Adaptive Testing

As stated previously, due to the mathematical and computational modeling, IRT and CAT can be a bit daunting for clinicians and applied scientists. Therefore, we provide a concrete example of our last CAT simulation (theta bound=–1, max items=6) (Figure 3). In our simulation, all patients started with item 4. Based on the answer to item 4, either item 2 or item 6 is selected, or the person is at an elevated risk (if the participant answers with response 2). For example, if a patient choose response 1 for item 4 (“I have a weak desire to kill myself”), the next most informative item would be item 6 (“length of periods of thinking about killing oneself”). If a patient answers that item with either moderate or long periods (responses 1 or 2), they would be categorized to be at high risk. If a patient selects response 0 (very brief periods), the next item would be item 7 (frequency of suicidal thoughts). Following this algorithm, a high-risk patient needs only 1 or 2 items to be classified as having an elevated risk.

Figure 3. Concrete example of a result of a CAT simulation.



Conclusions

One of the main advantages of CAT is the reduction of respondent burden. Answering items on suicidal behavior online can be difficult for patients, resulting in a high dropout rate. Because attrition is a well-known problem in eHealth, reducing response burden of online assessment of suicidal behavior is important. It should be noted that our CAT simulation showed

that the number of items can be considerably reduced when using the BSS-NL to assess elevated risk of suicidal behavior. Our simulation showed that 4 items, on average, were sufficient. Obviously, for CAT to be widely accepted and implemented, many more (prospective) studies should be done and ICT within mental health or research settings should be drastically improved. However, considering the need for rapid yet accurate online assessment of suicide risk in both clinical and research practice, we argue that IRT and CAT are likely to play important roles in the development of better measurement methods for the assessment of risk of suicidal behavior.

Acknowledgments

This study is funded by the Dutch organization for health research and development (ZonMW).

Authors' Contributions

AK, MdG, and JdK obtained funding for this study. DdB carried out the study. DdB and AdV drafted the manuscript. AK, MdG, and JdK contributed to the execution of the study and to the manuscript writing.

Conflicts of Interest

None declared.

References

1. Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. J Consult Clin Psychol 1979 Apr;47(2):343-352. [Medline: 469082]
2. Brown GK, Beck AT, Steer RA, Grisham JR. Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. J Consult Clin Psychol 2000 Jun;68(3):371-377. [Medline: 10883553]
3. Beck AT, Steer RA, Ranieri WF. Scale for Suicide Ideation: psychometric properties of a self-report version. J Clin Psychol 1988 Jul;44(4):499-505. [Medline: 3170753]
4. Brown GK. A review of suicide assessment measures for intervention research with adults and older adult. 2001 Jan 01. URL: http://ruralccp.org/lyra-data/storage/asset/brown-nd-27cb.pdf [accessed 2014-09-01] [WebCite Cache ID 6SGBncew4]
5. Mishara BL. In: Mishara BL, Kerkhof AJ, editors. Suicide Prevention and New Technologies. New York: Palgrave Macmillan; Jan 01, 2013:1-185.
6. De Jaegere E. Analysis of Available Technology-Based Suicide Prevention Programmes. 2013 Jan 01. URL: http://www.zorg-en-gezondheid.be/uploadedFiles/Zorg_en_Gezondheid/Nieuws/2014/Analysis%20of%20TBSP%20programmes.pdf [accessed 2014-09-05] [WebCite Cache ID 6SMPnPKZN]
7. van Spijker B. Reducing the Burden of Suicidal Thoughts Through Online Self-Help [dissertation]. Amsterdam: VU University Amsterdam; 2012 Jan 01. URL: http://dare.ubvu.vu.nl/bitstream/handle/1871/35540/dissertation.pdf;jsessionid=0292571E72767707968B94D6CDEF289F?sequence=1 [accessed 2014-09-05] [WebCite Cache ID 6SMQ4ARzX]
8. Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delespaul P, van Os J. Experience sampling research in psychopathology: opening the black box of daily life. Psychol Med 2009 Sep;39(9):1533-1547. [doi: 10.1017/S0033291708004947] [Medline: 19215626]
9. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. J Abnorm Psychol 2009 Nov;118(4):816-827. [doi: 10.1037/a0016948] [Medline: 19899851]
10. de Beurs E, den Hollander-Gijsman ME, van Rood YR, van der Wee NJ, Giltay EJ, van Noorden MS, et al. Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother 2011;18(1):1-12. [doi: 10.1002/cpp.696] [Medline: 20238371]
11. 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]
12. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, PROMIS Cooperative Group. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care 2007 May;45(5 Suppl 1):S22-S31. [doi: 10.1097/01.mlr.0000250483.85507.04] [Medline: 17443115]
13. Smits N, Cuijpers P, van Straten A. Applying computerized adaptive testing to the CES-D scale: a simulation study. Psychiatry Res 2011 Jun 30;188(1):147-155 [FREE Full text] [doi: 10.1016/j.psychres.2010.12.001] [Medline: 21208660]
14. Walter OB, Becker J, Bjorner JB, Fliege H, Klapp BF, Rose M. Development and evaluation of a computer adaptive test for 'Anxiety' (Anxiety-CAT). Qual Life Res 2007;16 Suppl 1:143-155. [doi: 10.1007/s11136-007-9191-7] [Medline: 17342455]
15. Jorm AF, Kelly CM, Morgan AJ. Participant distress in psychiatric research: a systematic review. Psychol Med 2007 Jul;37(7):917-926. [doi: 10.1017/S0033291706009779] [Medline: 17224097]

16. Younes N, Chee CC, Turbelin C, Hanslik T, Passerieux C, Melchior M. Particular difficulties faced by GPs with young adults who will attempt suicide: a cross-sectional study. BMC Fam Pract 2013;14:68 [FREE Full text] [doi: 10.1186/1471-2296-14-68] [Medline: 23706018]
17. Embretson SE. Item Response Theory for Psychologists. New Jersey: Psychology Press; Jan 01, 2000.
18. de Beurs DP, de Groot MH, Bosmans JE, de Keijser J, Mokkenstorm J, Verwey B, et al. Reducing patients' suicide ideation through training mental health teams in the application of the Dutch multidisciplinary practice guideline on assessment and treatment of suicidal behavior: study protocol of a randomized controlled trial. Trials 2013;14:372 [FREE Full text] [doi: 10.1186/1745-6215-14-372] [Medline: 24195781]
19. de Beurs DP, de Groot MH, de Keijser J, Verwey B, Mokkenstorm J, Twisk JW, et al. Improving the application of a practice guideline for the assessment and treatment of suicidal behavior by training the full staff of psychiatric departments via an e-learning supported Train-the-Trainer program: study protocol for a randomized controlled trial. Trials 2013;14:9 [FREE Full text] [doi: 10.1186/1745-6215-14-9] [Medline: 23302322]
20. van Spijker BA, van Straten A, Kerkhof AJ. The effectiveness of a web-based self-help intervention to reduce suicidal thoughts: a randomized controlled trial. Trials 2010;11:25 [FREE Full text] [doi: 10.1186/1745-6215-11-25] [Medline: 20214777]
21. Team RDC. R project. Vienna, Austria: R Project for Statistical Computing; 2009. URL: http://cran.r-project.org/ [accessed 2014-09-01] [WebCite Cache ID 6SGCtRX0l]
22. Revelle W. An overview of the psych package. 2014 Aug 11. URL: http://personality-project.org/r/overview.pdf [accessed 2014-09-01] [WebCite Cache ID 6SGD1i29X]
23. Rosseel Y. Journal of Statistical Software. 2012 May. lavaan: an R package for structural equation modeling URL: http://www.jstatsoft.org/v48/i02/paper [accessed 2014-09-05] [WebCite Cache ID 6SMQSrJ8n]
24. Rizopoulos D. Ltm: An R package for latent variable modeling item response theory analyses. Journal of Statistical Software 2006;17(5):1-25.
25. van der Ark LA. Journal of Statistical Software. 2007 May. Mokken scale analysis in R URL: http://www.jstatsoft.org/v20/i11/paper [accessed 2014-09-05] [WebCite Cache ID 6SMQfAa1A]
26. Choi SW, Gibbons LE, Crane PK. lordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations. J Stat Softw 2011 Mar 1;39(8):1-30 [FREE Full text] [Medline: 21572908]
27. Nydick SW. Package CATIRT. 2014 Apr 02. URL: http://cran.r-project.org/web/packages/catIrt/catIrt.pdf [accessed 2014-09-01] [WebCite Cache ID 6SGDOwjx0]
28. Gadermann AM, Guhn M, Zumbo BD. Practical Assessment, Research & Evaluation. 2012 Jan. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide URL: http://pareonline.net/getvn.asp?v=17&n=3 [WebCite Cache ID 6SMQvxShi]
29. Reise SP, Waller NG. Item response theory and clinical measurement. Annu Rev Clin Psychol 2009;5:27-48. [doi: 10.1146/annurev.clinpsy.032408.153553] [Medline: 18976138]
30. van der Linden WJ. Handbook of Modern Item Response Theory. New York: Springer; 1997.
31. Samejima F. Psychometrika monograph supplement. Richmond, VA: Psychometric Society; 1969 Jan 01. Estimation of latent ability using a response pattern of graded scores URL: http://www.psychometricsociety.org/sites/default/files/pdf/MN17.pdf [accessed 2014-09-05] [WebCite Cache ID 6SMRDEbpk]
32. Holland PW, Wainer H. Differential Item Functioning. New York: Routledge; 1993.
33. Young MA, Halper IS, Clark DC, Scheftner W, Fawcett J. An item-response theory evaluation of the Beck Hopelessness Scale. Cogn Ther Res 1992 Oct;16(5):579-587. [doi: 10.1007/BF01175143]
34. Kocalevent RD, Rose M, Becker J, Walter OB, Fliege H, Bjorner JB, et al. An evaluation of patient-reported outcomes found computerized adaptive testing was efficient in assessing stress perception. J Clin Epidemiol 2009 Mar;62(3):278-87, 287.e1. [doi: 10.1016/j.jclinepi.2008.03.003] [Medline: 18639439]
35. Choi SW, Reise SP, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res 2010 Feb;19(1):125-136 [FREE Full text] [doi: 10.1007/s11136-009-9560-5] [Medline: 19941077]
36. Borsboom D. The attack of the psychometricians. Psychometrika 2006;71(3):425-440.

Abbreviations

BSS: Beck Scale for Suicide Ideation
BSS-NL: Dutch version of the BSS
CAT: computer adaptive testing
DIF: differential item functioning
IRT: item response modeling
PCA: principal component analysis
PITSTOP suicide: Professionals in Training to STOP suicide

PROMIS: The Patient-Reported Outcomes Measurement Information System
Rrest: item-rest correlation

Edited by G Eysenbach; submitted 06.05.14; peer-reviewed by P Batterham, E De Jaegere; comments to author 16.07.14; accepted 25.07.14; published 11.09.14.

Please cite as:
De Beurs DP, de Vries ALM, de Groot MH, de Keijser J, Kerkhof AJFM
Applying Computer Adaptive Testing to Optimize Online Assessment of Suicidal Behavior: A Simulation Study
J Med Internet Res 2014;16(9):e207
URL: http://www.jmir.org/2014/9/e207/
doi:10.2196/jmir.3511
PMID:25213259

Original Paper

The Management of Acute Adverse Effects of Breast Cancer Treatment in General Practice: A Video-Vignette Study

Moyez Jiwa[1], MD, FRCP, FRACGP; Anne Long[2*], BSc, BMBS, FRACP; Tim Shaw[2*], BSc, PhD; Georgina Pagey[1], MBBS, FRACGP; Georgia Halkett[3], PhD, FIR, BMedRad; Vinita Pillai[1], BSc, MSc; Xingqiong Meng[1], MBBS, MAppE, PhD

[1]Curtin University, Perth, Australia

[2]Sydney Medical School, Sydney, Australia

[3]Curtin Univeristy, Perth, Australia

[*]these authors contributed equally

Corresponding Author:
Moyez Jiwa, MD, FRCP, FRACGP
Curtin University
GPO Box U1987
Perth, 6845
Australia
Phone: 61 8 9266 1768
Fax: 61 8 9266 1608
Email: m.jiwa@curtin.edu.au

Abstract

Background: There has been a focus recently on the use of the Internet and email to deliver education interventions to general practitioners (GPs). The treatment of breast cancer may include surgery, radiotherapy, chemotherapy, and/or hormone treatment. These treatments may have acute adverse effects. GPs need more information on the diagnosis and management of specific adverse effects encountered immediately after cancer treatment.

Objective: The goal was to evaluate an Internet-based educational program developed for GPs to advise patients with acute adverse effects following breast cancer treatment.

Methods: During phase 1, participants viewed 6 video vignettes of actor-patients reporting 1 of 6 acute symptoms following surgery and chemotherapy and/or radiotherapy treatment. GPs indicated their diagnosis and proposed management through an online survey program. They received feedback about each scenario in the form of a specialist clinic letter, as if the patient had been seen at a specialist clinic after they had attended the GP. This letter incorporated extracts from local guidelines on the management of the symptoms presented. This feedback was sent to the GPs electronically on the same survey platform. In phase 2, all GPs were invited to manage similar cases as phase 1. Their proposed management was compared to the guidelines. McNemar test was used to compare data from phases 1 and 2, and logistic regression was used to explore the GP characteristics that were associated with inappropriate case management.

Results: A total of 50 GPs participated. Participants were younger and more likely to be female than other GPs in Australia. For 5 of 6 vignettes in phase 1, management was consistent with expert opinion in the minority of cases (6%-46%). Participant demographic characteristics had a variable effect on different management decisions in phase 1. The variables modeled explained 15%-28% of the differences observed. Diagnosis and management improved significantly in phase 2, especially for diarrhea, neutropenia, and seroma sample cases. The proportion of incorrect management responses was reduced to a minimum (25.3%-49.3%) in phase 2.

Conclusions: There was evidence that providing feedback by experts on specific cases had an impact on GPs' knowledge about how to appropriately manage acute treatment adverse effects. This educational intervention could be targeted to support the implementation of shared care during cancer treatment.

(J Med Internet Res 2014;16(9):e204) doi:10.2196/jmir.3585

KEYWORDS

breast cancer; treatment; general practice; adverse effects; patient care planning

Introduction

Breast cancer was the most common cancer in Australian women in 2009 (excluding nonmelanoma skin cancer) [1]. In Australia, 1 in 11 women will develop breast cancer in their lifetime [2]; 89% of women with breast cancer survive more than 5 years and die of unrelated causes [3]. The treatment of breast cancer may include surgery, radiotherapy, chemotherapy, and/or hormone treatment [4]. Treatment depends on prognosis, stage of disease, treatment options, and adverse effects, as well as the patient and her partner's preferences [5].

Following adjuvant treatment, women may experience bowel disturbance, neutropenia, radiation dermatitis, and fatigue [6-7]. Posttreatment, brief follow-up is provided in the tertiary settings in some instances. Patients are also encouraged to see their general practitioner (GP) about any new or ongoing problems [8]. Previous studies have demonstrated that women consult a GP routinely in the months and years after treatment of breast cancer [9]. Physical adverse effects are a significant feature during breast cancer treatment [10]. Breast cancer patients are more likely to contact their GP for gastrointestinal symptoms than for other symptoms [11]. However, there is no evidence that these patients are advised correctly by GPs, and patients experience substantial unmet need for reassurance and advice [12]. There is also strong interest in the specialist sector to improve the support received by patients who have been treated for cancer [13]. To address the needs of patients treated for breast cancer, the GP needs to know how to effectively treat the adverse effects of therapy and understand the indications for urgent referral for specialist care. There is some evidence that GPs need more information on the diagnosis and management of specific adverse effects encountered immediately after cancer treatment [8].

There has been a recent focus on the use of the Internet and email to deliver education interventions to GPs. The use of video vignettes to explore medical decision making and to test other innovations is promising [14]. Preliminary evidence using video vignettes suggests that some innovations are not effective [15]. The value of letters from specialists about patients currently under the care of a GP is known to have an educational value; therefore, the study reported here incorporates this style of feedback education as an intervention delivered via the Internet in conjunction with video vignettes [16].

Methods

Overview

Following approval from the Curtin Human Research Ethics Committee, Perth, Western Australia (RD-68-12), participants were recruited from a previously established network of 150 GPs across Australia [15]. GPs were emailed invitations and nonresponders were followed up with personal invitations. Participants were remunerated with AUD $50 for their contribution.

Materials

A total of 12 video vignettes were developed, 1 pair for each potential adverse effect related to treatment of breast cancer (see Textbox 1 for exemplars). Each vignette depicted a patient with clear indications for specific management, including referral, prescription, reassurance, and/or investigation [15]. The vignettes were developed by 3 GPs, a radiation oncologist, a medical oncologist, and a surgeon with reference to what they considered the most common complication immediately after treatment. The expert panel referenced best practice guidelines in the development of management options for each case with suggestions for prescription, referral for specialist treatment, and laboratory investigation. Each case had more than 1 correct management option (see Table 1).

Textbox 1. Details of patients presented in the video vignettes (a=phase 1, b=phase 2).

Consultation 1a: “Marion Jones” 65-year-old female patient who had a lumpectomy 1 month ago. She is generally well. She suffers from hypertension and is treated with ramipril 5 mg. No allergies. Nonsmoker. She is now reporting a swelling at or near the surgical site with leakage of clear fluid through the lump. The area is not painful. Diagnosis: seroma

Consultation 1b: “Anne O’Brien” 60-year-old female patient had a left mastectomy and lymph node dissection 10 weeks ago. Now presents with a persistent painless swelling in her left axilla. She has been treated for depression with fluoxetine 20 mg/day and takes tamoxifen. Her husband is very worried about this lump and feels the cancer may be back. On examination, this is a soft fluctuant mass with mild erythema above the lesion. Diagnosis: seroma

Consultation 2a: “Christine Wilkins”, 50-year-old female patient had chemotherapy for breast cancer 4 days ago. Now presents with lethargy, moderate sore throat, and fever >38 °C. On examination, her throat is inflamed. She has no past medical history of note. She has no rigors or chills. She is allergic to penicillin and cephalosporins. Diagnosis: postchemotherapy infection (possible neutropenia)

Consultation 2b: “Marilyn Michaels” 56-year-old female patient had chemotherapy for breast cancer 9 days ago. Now presents with lethargy, slight dysuria, and fever >38 °C, BP <100 systolic, pulse >100. On examination, her urine is clear—no hematuria or proteinuria. She has no past medical history of note. She has no rigors or chills. Diagnosis: postchemotherapy infection (possible neutropenia)

Consultation 3a: “Margaret Enright” 59-year-old female patient had chemotherapy for breast cancer 9 days ago. Now presents with diarrhea for past 3 days. She has no abdominal pain and there is no blood in the motions. However, she has an episode of diarrhea every 2 hours and she feels weak and tired. She is not taking any medication and is allergic to penicillin. She lives with her sister who is working in the city and not able to be with her all day. On examination, she looks pale and tired. Her blood pressure is normal. Her abdomen is soft, nontender. Diagnosis: postchemotherapy diarrhea

Consultation 3b: “Francis Burn” 50-year-old female patient had chemotherapy for breast cancer 6 days ago. She has had 3 episodes of loose, watery bowel movements every day for the past 3 days. Her partner is concerned because her appetite is reduced and she has been feeling weak and tired. On examination, she is apyrexial but looks pale and tired. Her abdomen is soft, nontender. She is not clinically dehydrated, her blood pressure is normal. Diagnosis: postchemotherapy diarrhea

Consultation 4a: “Michelle Sands” 70-year-old female patient had chemotherapy for breast cancer 2 days ago now presents with vomiting 10 times a day. She has mild abdominal pain and feels thirsty but has lost her appetite. She denies any chest or abdominal pain. She is tired because she says she has not slept very well. Her son, who lives a few minutes away, is worried about her. She takes metformin 500 mg 3 times daily and amlodipine 5 mg/day. Diagnosis: postchemotherapy vomiting

Consultation 4b: “Sandra Speers” 55-year-old female patient had chemotherapy for breast cancer 4 days ago now presents with vomiting 4 times a day. She has mild lower abdominal pain and feels thirsty but has lost her appetite. She is tired because she has not slept very well. Her daughter is worried about her. She has had a myocardial infarction 6 months ago and is maintained on atenolol 50 mg/day and amlodipine 10 mg/day. She denies chest pain or breathlessness. On examination, there are no clinical signs of abdominal pathology. Diagnosis: postchemotherapy vomiting

Consultation 5a: “Susan Smith” 60-year-old female patient has been receiving full left breast radiation therapy for 5 weeks. She presents with a painful, itchy erythema and patchy, moist desquamation. She is aware that this is probably an adverse effect of the radiotherapy and wonders if anything can be done to make her more comfortable. She is also concerned that the therapy will scar her skin. She wonders if she is still fit to receive more radiotherapy. She is also worried if the therapy is having a bad effect on her ribs, heart, and lungs. Diagnosis: radiation dermatitis

Consultation 5b: “Doris Daniels” 50-year-old female is receiving radiotherapy to the right breast and axilla. She now presents with painful, itchy erythema and patchy, moist desquamation, especially in the skin folds. Paracetamol does not help the pain. She can’t wear a brassiere. She has been using topical propylene glycol but doesn’t find it particularly helpful. She is going to her sister’s wedding in a couple of days and would like to know if anything can help her be more comfortable. Diagnosis: radiation mastitis

Consultation 6a: “Alex Horner” 40-year-old female is receiving adjuvant chemotherapy for breast cancer. She now presents with constipation of 3 days duration. She has had very minimal bowel movements, mostly mucous and flatulence. She feels bloated and uncomfortable. On examination, she is bloated and there are minimal bowel sounds. On rectal examination, she is impacted with hard feces. Diagnosis: postchemotherapy constipation

Consultation 6b: “Michelle Marshall” 43-year-old female receiving chemotherapy for adjuvant breast cancer. She now complains of generalized colicky abdominal pain and constipation of 6 days duration. She has had no bowel movements in that time. She is very uncomfortable and occasionally troubled by the pain. She has had minimal oral intake and appears dehydrated. On examination, she is bloated and slightly tender throughout her abdomen. On rectal examination, she is impacted with feces and bleeding slightly from a hemorrhoid. Diagnosis: postchemotherapy constipation

Table 1. Specific recommendations for management of symptoms or problems after treatment for breast cancer.

Symptom or problem	Adverse effect		
	Mild	Moderate	Severe
Seroma			
Characterized by	Slight fluctuant swelling in operative site	Obvious fluctuant swelling in operative site	Tense (nonfluctuant) and uncomfortable swelling in operative site
Action to be taken by GP	Arrange for patient to see surgeon within 7 days	Arrange for patient to see surgeon within 4 days	Arrange for patient to see surgeon within 2 to 3 d
Infection	All the following:	Any of the following:	Any of the following:
Characterized by	Fever <38 °C Does not feel unwell BP normal Pulse <100	Fever >38 °C Mild symptoms plus BP <100 systolic Pulse >100	Fever >38 °C Symptoms including: Rigors/chills Dizziness BP <100 systolic Pulse >100
Action if <7 d or >14 d since chemotherapy	See patient <48 h Paracetamol regularly Review 1-2 d if worse	See patient <24 h Paracetamol regularly Oral antibiotics	Call medical oncologist. Send immediately to closest emergency department (ED). If patient from rural area and closest hospital >2 h away, consider giving IV/IM broad spectrum antibiotic prior to sending
Action if 7-14 d since chemotherapy	See patient <24 h Paracetamol regularly Oral antibiotics Review 1 d, if worse call medical oncologist	Call medical oncologist	Call medical oncologist Send immediately to closest ED. If patient from rural area and closest hospital >2 h away consider giving IV/IM broad spectrum antibiotic prior to sending
Diarrhea	≤4/d	5-8/d	>8/d±abdominal pain or dehydrated and cannot cope at home
Action to be taken	See patient <48 h Imodium after every bowel movement (BM) up to 8/d. If not controlled on this, then to return for review sooner and to trial codeine	See patient <24 h Trial of loperamide (2 after first BM and 1 after each subsequent), if insufficient then codeine phosphate 30 mg after each BM max 4/d	Call medical oncologist
Vomiting			
Characterized by	≤ 6/d still managing oral intake	7-10/d with mild oral intake	>10/d no oral intake
Action to be taken	See patient <48 h prochlorperazine maleate suppositories ± Ondansetron wafer 8 mg bid	See patient < 24 h prochlorperazine maleate suppositories ± Ondansetron 8 mg bid Call Medical Oncology Consider admission	Call medical oncologist and consider admission
Radiation dermatitis	All of the following	Any of the following	Any of the following
Characterized by	Faint skin erythema ± dry desquamation ± moderate pruritus	Brisk erythema ± patchy moist desquamation limited to skin folds/creases	Confluent moist desquamation ± erythema extending beyond treatment area ± systemically unwell

Symptom or problem	Adverse effect		
	Mild	Moderate	Severe
Action to be taken	Avoid additional skin irritants (eg, sun exposure, perfumed soaps/perfumes, adhesive tapes) Wash area daily in warm/tepid water Moisturize daily-tid with water-soluble moisturizer Simple analgesics Consider sparing use of topical steroids for patchy pruritus as long as there is no evidence of desquamation or skin breakdown See patient <48 h re: compliance	As above re: avoidance of skin irritants, simple hygiene, measures and analgesics Daily or twice daily nonadherent dressings for comfort See patient daily until resolution Contact Radiotherapy Department <48 h for advice/dressings	Analgesia as appropriate Daily or twice daily nonadherent dressings Commence empirical systemic antibiotics—only if indicated by evidence of systemic infection, recommend radiation oncologist review Contact radiation oncologist/Radiotherapy Department <24 h
Constipation			
Characterized by	≤2 d of nil/minimal BM	3-5 d of nil/minimal BM	>5 d of nil/minimal BM ± colic
Action to be taken	Recommend docusate sodium & senna II bid and or macrogol 3350 sachet I bid and glycerine suppository See patient <48 h	Increase docusate sodium and senna and macrogol 3350 consider addition of sorbitol etc or Macrogol AND sodium citrate suppositories See patient <24 h	Call medical oncologist Consider saline enema

The vignettes were then prepared as a short video monolog by an actor-patient (see Multimedia Appendix 1). The video included an off-camera commentary by an actor-doctor describing relevant signs to be found on clinical examination. Each of the 6 pairs was then randomly assigned to phase 1 or phase 2 of the study. Participation in the study was via the Internet. Participants were asked four questions after watching each video vignette:

1. What is your diagnosis?
2. Would you prescribe something? If so, what?
3. Would you refer the patient? If so, to whom?
4. Would you order tests? If so, which tests?

The responses were recorded via the Internet platform used to administer the survey. After responding to the first 6 of 12 videos, participants were provided written feedback in the style of a letter from a specialist clinic, highlighting the recommended guidelines for managing the adverse effect in each case.

The project was completed in two phases. In phase 1, participants were invited to view the first set of 6 videos and describe their management of the standardized patient depicted. All participants received expert feedback on the management of the cases viewed within 2 weeks of the project coordinator receiving their proposed management plan. The feedback was in the form of a letter written as if the patient had attended a specialist clinic immediately after consulting the GP. The letters did not refer to the GPs proposed plan for the patient, but stated "For Marion Jones [Consultation 1a], I would recommend the following...". The letter also outlined the protocol for the general management of her symptoms if they were mild, moderate, or severe. In most cases, there was more than 1 action that the GP could have taken to manage the case as per specialist guidelines.

The letters were sent via the Internet using the same Qualtrics survey platform. Once the study coordinator was alerted by the system that the participant had opened the letter for each case, they were sent a link to phase 2. In phase 2, all participating GPs were invited to view the second set of 6 videos and to describe their management of the standardized patient depicted. The phase 2 vignettes matched those in phase 1 by diagnosis (see Textbox 1).

Statistical Analysis

We hypothesized that the proportion of those who managed cases as per the expert recommendations would be greater after feedback (60% vs 30%). Therefore, a sample of 42 participants per group was deemed sufficient in this exploratory study [17].

The McNemar test was used to determine phase differences in the proportion of cases diagnosed and managed correctly. The phase 1 data offered the opportunity to investigate the GP characteristics that were associated with an incorrect response. GPs' characteristics associated with inappropriate case management were explored by using logistic regression models using phase 1 data. This helped to identify which groups of practitioners might best be targeted for the intervention and may be different for each of the case types. A full regression model included the following variables: age, sex, country of graduation, years after graduation, years of GP experience, status as established GP or GP registrar (trainee general practitioner), recognized speciality qualification with the Royal Australian College of General Practitioners (Fellow of the Royal Australian College of General Practitioners, FRACGP), the remoteness of their primary practice, the number of GPs at their primary practice, status as a principal (practice owner) within their primary practice, number of patients seen per week, total patient

care hours per week, and whether they conducted consultations in languages other than English. Regression models were constructed using both backwards elimination and forward selection. Univariate modeling was performed before the stepwise regressions, and the results were used to guide the reduction of the full models. Variables with a *P* value less than .05 were retained in the final model and reported. Stata version 12.1 (StataCorp LP, College Station, TX, USA) was used to perform the analysis. Logistic regression models were adjusted for the lack of independence between individual participants by estimating the clustered standard errors to account for intragroup correlation (vce option in Stata). No special technique was used to handle the missing values because there are very few missing values in this study. For participants' demographics, only the variable "sessions per week" had 2 missing values (4% of total), and there was no missing value for the outcome variables.

Results

A total of 50 GPs consented to participate and completed the study. GPs self-reported their demographic characteristics (see Table 2). Those who participated in the study were younger than Australian GPs generally (mean age 43.4 years vs 50.5 years) and a greater proportion were female (52% vs 39.1%), registrars (18% vs 3.8%), and Australian graduates (76% vs 65.9%) [18-20]. Most participants (62%-100%) correctly diagnosed cases in this study, especially in phase 2 (see Table 3). However, there were significant differences in the management of cases between the two phases and between cases (see Table 4).

Table 2. Participant demographic information (N=50).

GP characteristics	Study sample	National population[a]	
		Mean	%
Demographics			
Age (years), mean (SD)	43.4 (11.0)	50.5	
Years after graduation, mean (SD)	19.5 (11.2)		
Years of GP experience, mean (SD)	14.8 (11.3)		
GPs at primary clinic, mean (SD)	7.7 (4.0)		
GP sessions worked/week (n=48), mean (SD)	6.5 (3.1)		
Sex, n (%)			
Male	24 (48)		60.9
Female	26 (52)		
Graduated in Australia, n (%)	38 (76)		65.9
Registrars (GPs in training), n (%)	9 (18)		3.8
FRACGP (Fellows of the Royal Australian College of GPs), n (%)	30 (60)		56.8
Practice demographics			
Accredited, n (%)	49 (98)		88.6
Location (Australian state/territory), n (%)			
Australian Capital Territory	1 (2)		1.5
New South Wales	7 (14)		33.1
Queensland	3 (6)		19.5
South Australia	6 (12)		8.4
Victoria	13 (26)		25.1
Western Australia	20 (40)		9.1
Clinic remoteness, n (%)			
Major city	38 (76)		71.1
Nonmajor city	12 (24)		28.9
GP position, n (%)			
Principal	9 (18)		
Nonprincipal	32 (64)		
Others	9 (18)		
Patient consultations			
Patient consultations per week, n (%)			
<100	27 (54)		
100-149	14 (28)		
≥150	9 (18)		
Patient consultations hours per week, n (%)			
<11	7 (14)		1.2
11-20	8 (16)		12.2
21-40	27 (54)		53
≥41	8 (16)		33.5
Non-English consultations, n (%)			
No	43 (86)		
Yes, <25%	7 (14)		24.5

[a]Sourced from national data when available [18-20].

Table 3. Correct diagnosis of cases per phase of study (N=50).

Diagnosis	Phase 1, n (%) (n=50)	Phase 2, n (%) (n=50)	*P* value[a]
Constipation	40 (80)	39 (78)	.80
Diarrhea	46 (92)	50 (100)	.046
Radiation dermatitis	47 (94)	47 (94)	>0.99
Postchemotherapy infection	44 (88)	49 (98)	.03
Postoperative seroma	31 (62)	49 (98)	<.001
Vomiting	47 (94)	50 (100)	.08
Total (n=300)	255 (85.0)	284 (94.7)	<.001

[a]*P* values derived from McNemar test.

Table 4. Correct management of cases by phase of study (N=50).[a]

Management	Phase 1, n (%) (n=50)	Phase 2, n (%) (n=50)	*P* value[b]
Constipation			
Refer to oncologist	3 (6)	8 (16)	.14
Prescribe medication	38 (76)	44 (88)	.11
Diarrhea			
Prescribe Medication	32 (64)	45 (90)	.003
Radiation dermatitis			
Refer to breast care nurse / radiation oncologist / specialist	33 (66)	25 (50)	.09
Prescribe specific creams	3 (6)	19 (38)	<.001
Advise patient	13 (26)	38 (76)	<.001
Postchemotherapy infection			
Refer to oncologist	18 (36)	27 (54)	.07
Collect throat swab	18 (36)	33 (66)	.003
Postoperation seroma			
Refer to surgeon	23 (46)	40 (80)	.001
Organize ultrasound	35 (70)	39 (78)	.32
Aspiration	16 (32)	13 (26)	.49
Vomiting			
Refer to oncologist	20 (40)	5 (10)	<.001
Prescribe medication	22 (44)	46 (92)	<.001
Order tests	20 (40)	4 (8)	<.001

[a]In each case there is more than 1 correct answer.

[b]*P* values derived from McNemar test.

Regression analysis was carried out to determine the variables associated with management of adverse effects that were not consistent with expert opinion in phase 1. GPs failed to manage the cases as recommended by experts with reference to 2 explanatory variables:

1. GP demographics
2. Individual vignettes

In phase 1, as shown in Table 5, Australian graduates were statistically less likely to offer an inappropriate referral. Male GPs and participants who identified themselves as neither practice owner nor employee (most likely locum practitioners) were less likely to provide an inappropriate prescription. GPs who worked more sessions per week were more likely to offer an inappropriate prescription. Older GPs were less likely to

order a unnecessary test. In contrast, those who consulted for more than 20 hours per week were more likely to order an unnecessary test. Compared to managing constipation or diarrhea, participants were less likely to make an inappropriate referral. Patients with radiation dermatitis, postchemotherapy infection, or vomiting were more likely to be given an inappropriate prescription and test.

These variables explained 15%-28% of the differences observed (Pseudo R^2=15%, 28%, and 15% for inappropriate referral, prescription, and test order, respectively) (see Table 5). Results in Table 5 are odds ratios and 95% confidence intervals derived from 3 logistic regression models adjusted for clustering of GPs. Only variables with *P* values <.05 were included in the final model and reported in Table 5.

Overall, all three aspects of management had improved significantly in phase 2 as shown in Table 6.

Table 5. Factors associated with incorrect management (inconsistent with expert opinion) in phase 1 (N=50).

Explanatory variable	Inappropriate referral		Inappropriate prescription		Unnecessary tests	
	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Age	—		—		0.98 (0.96, 1.00)	.04
GP sessions worked/week	—		1.12 (1.03, 1.22)	.008	—	
Sex	—				—	
Female			1.00			
Male			0.44 (0.26, 0.74)	.002		
Graduated in Australia			—		—	
No	1.00					
Yes	0.47 (0.23, 0.97)	.04				
Patient consultations hours/week	—		—			
<21					1.00	
21-40					1.87 (1.08, 3.22)	.03
>41					2.01 (1.14, 3.56)	.02
Position	—				—	
Principal			1.00			
Nonprincipal			0.49 (0.21, 1.11)	.09		
Others			0.25 (0.09, 0.71)	.009		
Cases						
Constipation	1.00		1.00		1.00	
Diarrhea	0.29 (0.07, 1.25)	.10	1.90 (0.78, 4.62)	.16	15.92 (6.11, 41.47)	<.001
Radiation dermatitis	0.03 (0.01, 0.11)	<.001	59.41 (14.05, 251.26)	<.001	2.19 (0.88, 5.46)	.09
Postchemotherapy infection	0.11 (0.03, 0.38)	<.001	4.66 (1.82, 11.91)	.001	3.15 (1.23, 8.07)	.02
Postoperative seroma	0.07 (0.02, 0.30)	<.001	0, 30 (0.08, 1.04)	.06	1.62 (0.53, 5.00)	.40
Vomiting	0.09 (0.03, 0.32)	<.001	4.66 (1.90, 11.43)	.001	7.32 (2.49, 21.59)	<.001

Table 6. Incorrect management by phase.

Management	Phase 1, n (%) (n=300)	Phase 2, n (%) (n=300)	*P* value[a]
Refer to specialist	194 (64.7)	148 (49.3)	<.001
Prescription	138 (46.0)	76 (25.3)	<.001
Order test	126 (42.0)	103 (34.3)	.03

[a]*P* values derived from McNemar test.

Discussion

These data indicate that although most participants correctly diagnosed the conditions presented throughout the study, limited numbers knew how to manage the acute adverse effects of breast cancer treatment. Australian graduates performed better, but those who worked longer hours were more likely to make questionable decisions in this study. The latter may reflect research that longer hours have a negative impact on job performance [21]. This study did not test performance with real patients or in conditions of varying levels of fatigue; therefore, the comments remain speculative. We also note that practitioners who worked longer hours were more likely to order unnecessary tests. It is possible that this group is more comfortable with trying to manage cases on their own rather than refer back to an oncologist. However, we were unable to explore this hypothesis with the data collected.

The management of radiation dermatitis, postchemotherapy infection, and vomiting proved the most challenging. For almost every case, the management improved following feedback. These differences were marked for seroma, postchemotherapy infection, and diarrhea. This is an important observation which suggests that, if this study had been conducted with real patients, there was scope for significant harm because of diagnostic or management failures. Participants were more likely to diagnose and refer a seroma after feedback. Such differences in the management of acute adverse effects by GPs have not been reported previously because most patients are likely to consult their specialist within days or weeks of treatment rather than a GP [22].

Some adverse effects, such as persistent vomiting after chemotherapy, are likely to be emergencies; others, such as seromas, are distressing to patients, but unlikely to be life threatening. Some adverse effects, such as a postchemotherapy infection, can cause significant harm if they go unrecognized [23]. It has been suggested that GPs should play a much more active part during the treatment phase of the patient's cancer journey [24]. If this is to be the case, then GPs need to be trained to manage the common acute effects of cancer treatment and at the very least these conditions need to feature in the differential diagnosis of patients presenting with symptoms soon after treatment of breast cancer [6,7].

Differences in the proposed management between the participants and the expert panel were less marked in phase 2 (after feedback). Such improvements, if they were noted in actual clinical practice, would lead to a reduction in adverse incidents, and better outcomes and satisfaction for patients. For example, as shown in Table 4, in phase 1 only 6% of participants prescribed the appropriate treatment of radiation dermatitis, whereas in phase 2 this proportion increased to 38%. In the case of the possible neutropenia, a significant proportion would arrange appropriate investigations in phase 2. This increases the potential for shared care between health sectors and makes it more likely treatment would be offered sooner rather than later. This is especially the case where patients may suffer avoidable harm if the practitioner in the community is able to recognize the need for urgent specialist advice for someone receiving lifesaving treatment. There were still significant numbers of participants whose proposed management of the vignettes was not consistent with expert opinion. In this study, it was not clear whether this was because participants disagreed with the suggested treatment plans or failed to assimilate the feedback into the phase 2 responses. Although there was a marked improvement in the management of cases, it would be unsafe to assume this was entirely related to the feedback received after phase 1. In the case of chemotherapy-induced vomiting in phase 2, although the participants were more likely to prescribe an antiemetic, they were less likely to refer back to the oncologist or to order the relevant tests after feedback. This was unexpected. It is possible that the scenario presented was considered "mild" because the patient was reported to have vomited only 4 times a day, in which case it may have been deemed unnecessary to refer to the oncologist or arrange laboratory tests to check the renal function. A future study involving this scenario would need to make it clearer that the patient was in need of specialist advice. Therefore, more severe symptoms would need to be presented in the vignette.

A recent literature review reported that two other factors are also likely to be important in the context of a cancer diagnosis: attitudes and beliefs [24]. These issues were not evaluated in this study. For example, we were unable to report the participants' attitudes toward the management of patients with acute adverse effects and whether they felt this role extended to investigating and treating acute conditions that may have resulted from specialist treatment [25]. A diversity of opinions in regards to this issue have been described among Australian GPs in previous reviews [26]. Nor could we confirm that all participants had easy access to the relevant specialists and/or would have had the option to refer a patient urgently with a condition they had not previously encountered to such an expert. The available evidence suggests that this is not a safe assumption and that management plans would be impacted by the clinicians' experience in their local context [13].

This pilot study had a modest sample size, which was estimated based on the hypothesis that the participants would be twice as likely to manage patients as per expert opinion following feedback on a similar previous case. This was not true of all management modalities. In some cases, any significant improvement in phase 2, as shown in Table 4, was much more modest. Therefore, a much larger study would be required to robustly demonstrate that this mode of education is likely to increase GPs' knowledge. Because no other educational intervention was offered in a randomized experimental design, the conclusions that can be drawn from these data are also limited.

This study with vignette-based feedback showed promising results that managing the common adverse effects of cancer treatment could be delegated to general practice. Such an intervention could support the application of shared care models of care. A larger study, including management of adverse effects in real patients, needs to be conducted before it can be safely recommended. However, noting that some patients with potentially life-threatening adverse effects may not be managed appropriately suggests a need for safeguards to protect patients in a study with bona fide patients.

Acknowledgments

This study was funded by a grant from the Royal Australian College of General Practitioners. We would like to thank the experts who assisted in developing the cases and advising on appropriate management.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example of video vignette.

[MOV File, 78MB - jmir_v16i9e204_app1.mov]

References

1. Australian Institute of Health &Welfare Australasian Association of Cancer Registries. Cancer in Australia: an overview, 2012. Canberra: AIHW & AACR; 2012 Jan 01. URL: http://www.aihw.gov.au/cancer/cancer-in-australia-overview-2012/ [accessed 2014-08-28] [WebCite Cache ID 6SACpLLIa]
2. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol 2001 Sep;2(9):533-543. [doi: 10.1016/S1470-2045(01)00486-7] [Medline: 11905707]
3. Australian Institute of Health and Welfare & Cancer Australia. Breast cancer in Australia: an overview. Canberra: AIHW; 2012 Jan 01. URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=10737423006 [accessed 2014-08-28] [WebCite Cache ID 6SAD3dkyl]
4. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Thürlimann B, Senn HJ, Panel members. Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. Ann Oncol 2005 Oct;16(10):1569-1583 [FREE Full text] [doi: 10.1093/annonc/mdi326] [Medline: 16148022]
5. Duric V, Stockler M. Patients' preferences for adjuvant chemotherapy in early breast cancer: a review of what makes it worthwhile. Lancet Oncol 2001 Nov;2(11):691-697. [doi: 10.1016/S1470-2045(01)00559-9] [Medline: 11902540]
6. Wood AJ, Shapiro CL, Recht A. Side effects of adjuvant treatment of breast cancer. Side effects of adjuvant treatment of breast cancer. N Engl J Med 2008:1997-2008. [doi: 10.1056/NEJM200106283442607]
7. Srivastava V, Basu S, Shukla VK. Seroma formation after breast cancer surgery: what we have learned in the last two decades. J Breast Cancer 2012 Dec;15(4):373-380 [FREE Full text] [doi: 10.4048/jbc.2012.15.4.373] [Medline: 23346164]
8. Raupach JC, Hiller JE. Information and support for women following the primary treatment of breast cancer. Health Expect 2002 Dec;5(4):289-301. [Medline: 12460218]
9. Collins RF, Bekker HL, Dodwell DJ. Follow-up care of patients treated for breast cancer: a structured review. Cancer Treat Rev 2004 Feb;30(1):19-35. [doi: 10.1016/S0305-7372(03)00141-5] [Medline: 14766124]
10. Carelle N, Piotto E, Bellanger A, Germanaud J, Thuillier A, Khayat D. Changing patient perceptions of the side effects of cancer chemotherapy. Cancer 2002 Jul 1;95(1):155-163 [FREE Full text] [doi: 10.1002/cncr.10630] [Medline: 12115329]
11. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG, Schouten HC, van Kleef M, Patijn J. Quality of life and non-pain symptoms in patients with cancer. J Pain Symptom Manage 2009 Aug;38(2):216-233. [doi: 10.1016/j.jpainsymman.2008.08.014] [Medline: 19564094]
12. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]
13. Jefford M, Rowland J, Grunfeld E, Richards M, Maher J, Glaser A. Implementing improved post-treatment care for cancer survivors in England, with reflections from Australia, Canada and the USA. Br J Cancer 2013 Jan 15;108(1):14-20 [FREE Full text] [doi: 10.1038/bjc.2012.554] [Medline: 23257892]
14. Jiwa M, Meng X. Video consultation use by Australian general practitioners: video vignette study. J Med Internet Res 2013;15(6):e117 [FREE Full text] [doi: 10.2196/jmir.2638] [Medline: 23782753]
15. Jiwa M, Halkett G, Meng X, Pillai V, Berg M, Shaw T. Supporting patients treated for prostate cancer: a video vignette study with an email-based educational program in general practice. J Med Internet Res 2014;16(2):e63 [FREE Full text] [doi: 10.2196/jmir.3003] [Medline: 24571952]
16. Westerman RF, Hull FM, Bezemer PD, Gort G. A study of communication between general practitioners and specialists. Br J Gen Pract 1990 Nov;40(340):445-449 [FREE Full text] [Medline: 2271276]
17. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]

18. Australian Institute of Health and Welfare. www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129546076. Canberra: AIHW; 2014. Medical Workforce 2010 URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129546076 [accessed 2014-08-28] [WebCite Cache ID 6SADByQi0]
19. General Practice Registrars Australia Ltd. Annual report, 2011-12. URL: http://issuu.com/gprapublications/docs/annual_report_2012_web_resolution_2/1?e=2280775/1564037 [accessed 2014-08-28] [WebCite Cache ID 6SA2jt2h3]
20. Britt H, Miller GC, Charles J, Henderson J, Bayram C, Pan Y. General practice activity in Australia 2008-09. Canberra: AIHW; 2009. URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442456272 [accessed 2014-08-28] [WebCite Cache ID 6SADSOPyT]
21. Lockley SW, Barger LK, Ayas NT, Rothschild JM, Czeisler CA, Landrigan CP, Harvard Work Hours, HealthSafety Group. Effects of health care provider work hours and sleep deprivation on safety and performance. Jt Comm J Qual Patient Saf 2007 Nov;33(11 Suppl):7-18. [Medline: 18173162]
22. Anvik T, Holtedahl KA, Mikalsen H. "When patients have cancer, they stop seeing me"--the role of the general practitioner in early follow-up of patients with cancer--a qualitative study. BMC Fam Pract 2006;7:19 [FREE Full text] [doi: 10.1186/1471-2296-7-19] [Medline: 16549036]
23. Crawford J, Dale DC, Lyman GH. Chemotherapy-induced neutropenia: risks, consequences, and new directions for its management. Cancer 2004 Jan 15;100(2):228-237 [FREE Full text] [doi: 10.1002/cncr.11882] [Medline: 14716755]
24. Jiwa M, McManus A, Dadich A. The impact of knowledge, attitudes and beliefs on the engagement of primary and community-based healthcare professionals in cancer care: a literature review. Curr Med Res Opin 2013 Nov;29(11):1475-1482. [doi: 10.1185/03007995.2013.838154] [Medline: 23998506]
25. Jiwa M, McManus A, Dadich A. Craft based: Continuity of cancer care: Where do primary care practitioners fit in? Cancer Forum 2013 Mar;37(1):31-34.
26. Johnson CE, Lizama N, Garg N, Ghosh M, Emery J, Saunders C. Australian general practitioners' preferences for managing the care of people diagnosed with cancer. Asia Pac J Clin Oncol 2014 Jun;10(2):e90-e98. [doi: 10.1111/ajco.12047] [Medline: 23279791]

Abbreviations

ED: emergency department
GP: general practitioner
BM: bowel movement

Edited by G Eysenbach; submitted 05.06.14; peer-reviewed by S Cutrona; comments to author 30.06.14; revised version received 10.07.14; accepted 13.08.14; published 03.09.14.

Please cite as:
Jiwa M, Long A, Shaw T, Pagey G, Halkett G, Pillai V, Meng X
The Management of Acute Adverse Effects of Breast Cancer Treatment in General Practice: A Video-Vignette Study
J Med Internet Res 2014;16(9):e204
URL: http://www.jmir.org/2014/9/e204/
doi:10.2196/jmir.3585
PMID:25274131

Original Paper

Comparison of Virtual Patient Simulation With Mannequin-Based Simulation for Improving Clinical Performances in Assessing and Managing Clinical Deterioration: Randomized Controlled Trial

Sok Ying Liaw[1], PhD; Sally Wai-Chi Chan[2], PhD; Fun-Gee Chen[3], MBBS, MMed (Anaesth); Shing Chuan Hooi[4], MBBS, PhD; Chiang Siau[3], MBBS, MMed(Anaesth)

[1]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore
[2]School of Nursing and Midwifery, Faculty of Health and Medicine, The University of Newcastle, Newcastle, Australia
[3]Department of Anaesthesia, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[4]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Corresponding Author:
Sok Ying Liaw, PhD
Alice Lee Centre for Nursing Studies
Yong Loo Lin School of Medicine
National University of Singapore
Alice Lee Centre for Nursing Studies. Level 2, Clinical Research Centre, Block MD 11, 10 Medical Drive
Singapore 117597,
Singapore
Phone: 65 65167451
Fax: 65 67767135
Email: nurliaw@nus.edu.sg

Abstract

Background: Virtual patient simulation has grown substantially in health care education. A virtual patient simulation was developed as a refresher training course to reinforce nursing clinical performance in assessing and managing deteriorating patients.

Objective: The objective of this study was to describe the development of the virtual patient simulation and evaluate its efficacy, by comparing with a conventional mannequin-based simulation, for improving the nursing students' performances in assessing and managing patients with clinical deterioration.

Methods: A randomized controlled study was conducted with 57 third-year nursing students who were recruited through email. After a baseline evaluation of all participants' clinical performance in a simulated environment, the experimental group received a 2-hour fully automated virtual patient simulation while the control group received 2-hour facilitator-led mannequin-based simulation training. All participants were then re-tested one day (first posttest) and 2.5 months (second posttest) after the intervention. The participants from the experimental group completed a survey to evaluate their learning experiences with the newly developed virtual patient simulation.

Results: Compared to their baseline scores, both experimental and control groups demonstrated significant improvements (P<.001) in first and second post-test scores. While the experimental group had significantly lower (P<.05) second post-test scores compared with the first post-test scores, no significant difference (P=.94) was found between these two scores for the control group. The scores between groups did not differ significantly over time (P=.17). The virtual patient simulation was rated positively.

Conclusions: A virtual patient simulation for a refreshing training course on assessing and managing clinical deterioration was developed. Although the randomized controlled study did not show that the virtual patient simulation was superior to mannequin-based simulation, both simulations have demonstrated to be effective refresher learning strategies for improving nursing students' clinical performance. Given the greater resource requirements of mannequin-based simulation, the virtual patient simulation provides a more promising alternative learning strategy to mitigate the decay of clinical performance over time.

(J Med Internet Res 2014;16(9):e214) doi:10.2196/jmir.3322

KEYWORDS

simulation; education; virtual patient; deterioration; clinical performance; patient safety

Introduction

The nurses' role in recognizing, responding to, and reporting a patient's clinical deterioration is critical in the prevention of the patient's progression to cardiopulmonary arrest. A literature review identified an educational need to prepare nursing students to manage physiological deterioration of patients before they start their ward nursing practice [1]. A 6-hour mannequin-based simulation program, known as RAPIDS (Rescuing a Patient in Deteriorating Situations) was implemented as part of a core learning activity in an undergraduate nursing program. A randomized controlled study has shown that the RAPIDS program using ABDCDE (Airway, Breathing, Circulation, Disability, and Expose/Examine) and SBAR (Situation, Background, Assessment, and Recommendation) mnemonics effectively developed nursing students' clinical competencies in assessing, managing deterioration, and communicating patient's deterioration to the team doctor [2]. However, it is unclear how best to maintain these competencies post-RAPIDS training.

A qualitative study was conducted to examine the effect of the RAPIDS program on the nursing students' performances in the clinical practice. The study showed the benefits of mannequin-based simulation in preparing the nursing students for their encounters with deteriorating ward patients. To optimize their retention and transfer of learning in RAPIDS, the study recommended regular reinforcement with follow-up simulation training using the ABCDE and SBAR mnemonics [3]. Previous studies have supported the use of mannequin-based simulation in the acquisition of clinical skills but have also demonstrated their limitations [1,4]. Because mannequin-based simulation involves a small number of students at one time, a considerable amount of faculty time is required to conduct repeated sessions. Besides faculty time, the availability of simulation facilities and scheduling issues are major challenges faced by educators when implementing mannequin-based simulation. These challenges made it difficult to be certain whether it is the best follow-up learning method to maintain or enhance previously acquired skills.

Virtual patient simulation has fewer of these resource constraints compared to mannequin-based simulation. It is capable of creating high-fidelity simulation by applying the features identified in a systematic review. With the capacity for exhibiting a high level of interactivity and realism, a wide range of clinical scenarios with guided reflection can be designed into the virtual patient simulation [5]. In addition, it can cater to a large number of learners simultaneously and be used by learners repeatedly when needed. Being accessible anytime and anywhere, it can also be integrated into curricula in a more flexible manner [6]. Although the use of virtual patient simulations have been widely adopted for training health professionals [7-9], more research is required to inform how to effectively design and integrate them into curricula [10,11].

A virtual patient simulation was designed and developed for use as an instructional learning strategy to revise RAPIDS training. We conducted a randomized controlled study to determine the efficacy of virtual patient simulation, by comparing it with mannequin-based simulation, in improving the nursing students' clinical performances in assessing and managing deterioration. A survey was also conducted to evaluate learners' perception towards the newly developed instructional strategy.

Methods

Design and Development of Virtual Patient Simulation

The virtual patient simulation, known as e-RAPIDS, was developed at National University of Singapore (NUS) by a group of academic staff, clinicians, and educational technologists. This single user interactive multimedia simulation was created using Flash software and run on a secure server hosted by NUS. The contents were developed based on the following learning objectives: (1) Demonstrate a systematic approach using the ABCDE mnemonic to assess and manage clinically deteriorated patient, and (2) Demonstrate effective communication skills using the SBAR tool to report patient deterioration to doctor. Five simulation scenarios associated with acute medical conditions (acute coronary syndrome, hypoglycemia, hypovolemic shock, sepsis, septic shock) were used. Common deteriorating conditions such as airway obstruction, breathlessness, hypotension, tachycardia, oliguria, altered consciousness, and abnormal temperature were embedded in these scenarios. All the scenarios applied the same clinical case history of a virtual patient who was admitted to a hospital with multiple medical conditions and comorbidities (Figure 1). The complexity of the case history allowed sequential simulation of deteriorating events at different phases of the virtual patient's hospitalization. Appropriate clinical findings and responses of the virtual patient were developed for each scenario.

Figure 2 presents the path of virtual patient simulation scenarios. The learner can choose to participate in any scenarios by clicking on the patient's day of admission (part A in Figure 2). The deteriorating events occur on first, third, sixth, eighth, and tenth day of admission. Once inside the virtual ward, the learner receives a handover report on the case history and the latest update of the virtual patient's condition (part B in Figure 2). After the handover, the learner, who plays the role of the nurse, is presented with a virtual patient with deteriorating conditions. The learner is required to assess and manage the virtual patient's deteriorating condition by clicking on the actions from the ABCDE control menus (part C in Figure 2). There are over 30 actions programmed into the simulation. The clicking of a specific action may lead to an arrow sign that directs the learner to click on specific equipment or an item in the virtual ward. Immediate feedback, including information and physiological changes, was programmed into the system to respond to the learner's actions. The information generated from an action is

delivered through the virtual nurse's verbalization with texts displayed in the form of speech bubbles. The physiological parameters including heart rate, blood pressure, and oxygenation are displayed on an electronic monitor in the virtual ward. SBAR control menus are used in the program to aid the learner in reporting about the patient's deterioration.

At the end of each scenario, the learner enters a "debriefing" screen (part D in Figure 2), which consists of (1) five debriefing questions, (2) an evaluation tool adapted from a validated and reliable tool known as RAPIDS tool, and (3) a performance score. The debriefing questions lead the learner to reflect on the deteriorating situation and actions they have taken. Using a checklist format and brief explanation, the evaluation tool provides feedback to the learners on the appropriate and inappropriate actions taken in the simulation scenario. A score is automatically calculated from the evaluation tool based on the learner's actions in the scenario (part D in Figure 2).

Figure 1. Clinical history of the virtual patient.

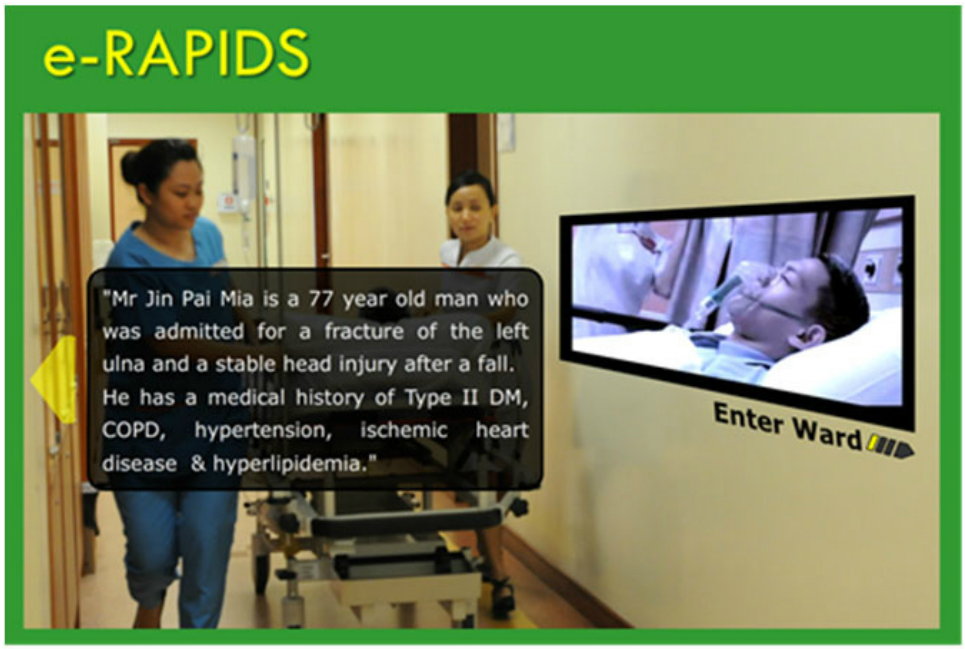
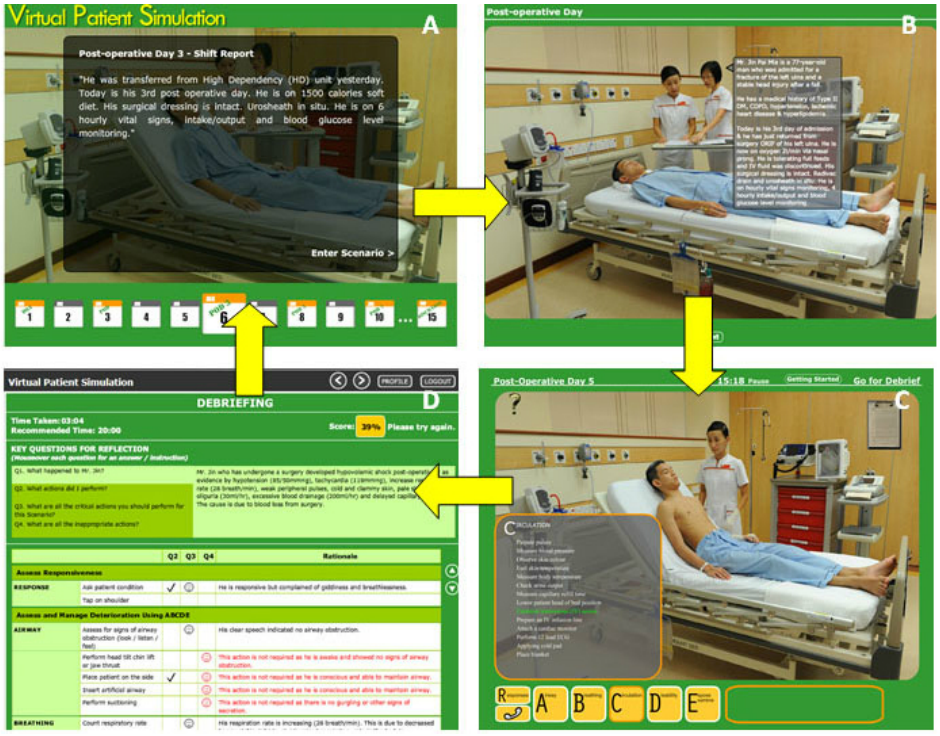


Figure 2. User follows A to D to navigate each scenario: A, Click on the patient's day of admission to enter a scenario; B, Receive patient information during hand-off report; C, Emulate the role of nurse to assess and manage deteriorating patient by clicking on the ABCDE options menus; D, Self-reflection through a list of questions.

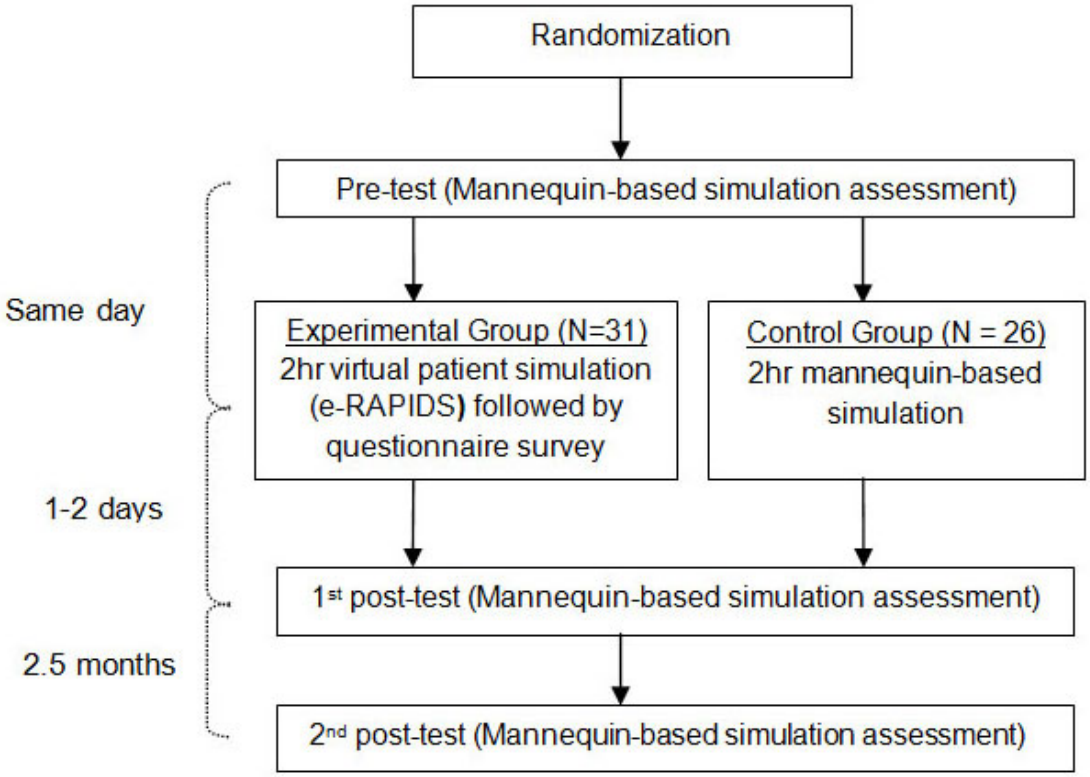


Evaluation of Virtual Patient Simulation

Design and Sample

A prospective, randomized controlled trial with a pretest-posttest design was conducted (see Figure 3). The study was approved by the National University of Singapore (NUS) Institutional Human Research Ethics Board. A total of 97 senior nursing students in their third year of nursing studies, who had undertaken a 6-hour mannequin-based RAPIDS simulation program about 8 months prior, from August to December 2011 at NUS, were invited to participate in the study through email. After being given a participant information sheet that explained the purpose of the study, 61 students consented to participate. The students were assured that their decision to participant or not to participate would not affect their training. They were randomly assigned to experiment (n=31) and control group (n=30) using a random number table. However, 4 participants from the control group withdrew from the study as they were unable to join the scheduled simulation, leaving that group with 26 students.

Figure 3. Flow of data collection.



Baseline Testing

After randomization, both groups underwent baseline testing using a mannequin-based simulation assessment at the university simulation center. The participants from both groups undertook the tests individually. To reduce bias, the participants were required to wear caps, gowns, and masks to blind their identities from the raters who may have known the participants' training background. Following an orientation to the simulated ward and brief introduction of the case history, a mannequin-based test scenario with signs and symptoms of clinical deterioration was presented to the participants. The participants were each given 15 minutes to assess and manage the deteriorating patient simulator with signs of shock. As all the participants undertook the same test scenario, they were reminded about the confidentiality of the scenario before they left the laboratory. The entire process of the testing scenario was videotaped.

Intervention

The interventions for both groups were conducted immediately after the baseline testing. Both interventions were conducted concurrently on the same day in the university's simulation center. The duration of intervention (2 hours) was equal in the experimental and control groups. The participants in the experimental group were brought individually into a room with a computer set up to use the virtual patient simulation as described above. They were instructed to undertake all the simulation scenarios in the virtual patient simulation. The participants from the control group were placed into groups of 6 to participate in the mannequin-based simulation, led by a trained simulation facilitator. The participants worked through two simulation scenario (sepsis and septic shock) of a patient with deteriorating conditions. These scenarios were developed to model situations that led the learners to use ABCDE and SBAR mnemonics to perform thorough nursing assessment and management, and to call for help. Each simulation scenario began with a role play followed by a debriefing session. While

3 participants undertook the role play, the rest of the learners observed the scene. To allow all participants to have hands-on experience, the participants reversed their roles as observers and role players in the two scenarios. During debriefing, the facilitator used questioning techniques to lead the learners to reflect their simulation performances. In addition, using the ABCDE and SBAR mnemonic checklists, the facilitator provided specific feedback to the learners.

Survey and Posttests

On completion of the virtual patient simulation, the participants in the experimental group were asked to complete a questionnaire to evaluate their learning experiences. All participants from both groups undertook two posttests using mannequin-based simulation assessment. The first posttest was conducted one or two days after the intervention. This was followed by the second posttest about 2.5 months later. The scenario and instrument used for the posttest simulation assessment were similar to the baseline testing. The entire process of the simulation assessment was videorecorded.

Instruments

The recorded simulation performances were observed and rated by an academic staff using the RAPIDS tool, which consists of two domains: (1) ABCDE domain consists of binary checklist items and a global rating scale item to evaluate the performance in assessing and managing patient deterioration, and (2) SBAR domain consists of binary checklist items and a global rating scale to measure the communications skills in reporting patient deterioration. The psychometric properties of the RAPIDS tool, including content and construct validity, and interrater reliability were tested and supported in a previous study [12]. An excellent interrater reliability between 2 raters, with high intraclass correlation coefficient of .92 (95% CI 0.82-0.96), was obtained in this study.

A 19-item questionnaire with four subscales (system quality, information quality, user satisfaction, and net benefit) was adapted and modified from the e-learning systems success (ELSS) scale to evaluate the participants' perception of the virtual patient simulation. Each item is rated on a 7-point Likert-type scale ranging from "strongly disagree" to "strongly agree". The scale was developed and tested by Wang et al [13] in a previous study to measure the success of an e-learning system in an organizational context. A high internal consistency of this scale was obtained in this study (Cronbach alpha=.904).

Data Analysis

The sample size calculation was based on a previous study measured using the RAPIDS tool. A sample of 15 participants per arm gave an effect size of 3.16 that could achieve more than 80% power at 5% alpha level [14]. Allowing for an attrition rate of 50%, particularly from the second posttest, the aim was to recruit 30 participants to each arm. Chi-square tests and *t* tests were used to examine any differences in demographic characteristics between the two groups. Interrater reliability was assessed by intraclass correlation coefficient. Repeated measures analysis of variance (ANOVA) was used to analyze time and group differences in the performance scores. Descriptive statistics using means and standard deviation were performed to examine the participants' perception of the virtual simulation.

Results

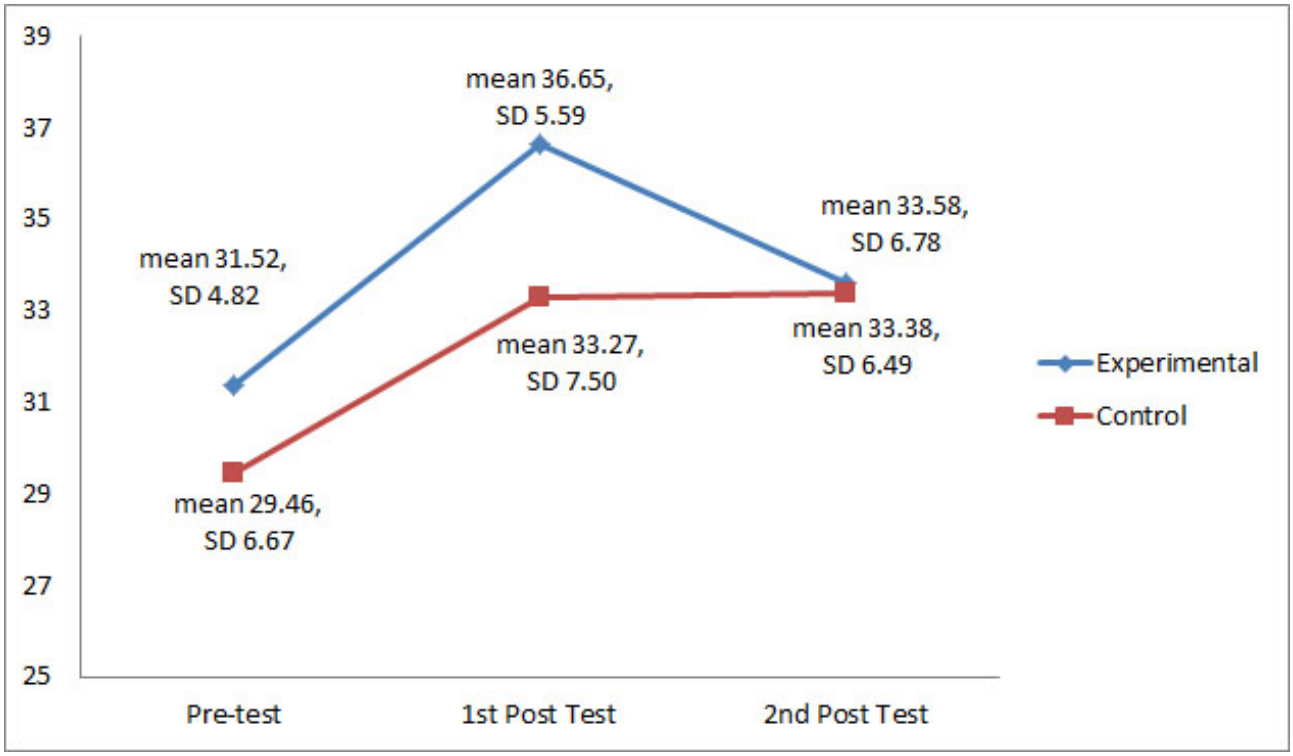
Most of the participants were female (89.5%, 51/57), Chinese (78.9%, 45/57), and an average of 21.86 years old (SD 1.13). The intervention and control group did not differ significantly in demographic variables of gender (P=.29), ethnic (P=.07), and age (P=.14). These results supported the randomization and homogeneity of the participants between the groups.

With the possible maximum performance mean scores of 54, the score of 30.58 (SD 5.78) indicated that most participants had an average performance score. Repeated-measures ANOVA (within-subject analysis) indicated a significant increase in the first posttest scores from the pretest scores for the experimental (P<.001) and control groups (P<.05). As shown in Figure 4, the second posttest scores for the experimental group decreased significantly (P<.05) from the first posttest. However, no significant difference (P=.94) between the first and second posttest scores was found for the control group. The second posttest scores for the experimental (P<.05) and control groups (P<.01) were significantly higher from the pretest scores.

Pairwise comparison between the two groups at the three time-points indicated that both groups increased significantly from pretest to first posttest (P<.001) and from the pretest to second posttest (P<.001). There was no significant difference between the first and second posttests for both groups (P=.12). Repeated-measures ANOVA (between-subject difference) showed that there was no significant difference in performance mean scores over time between the experimental and control group (P=.17).

The mean scores from the participants' rating (experimental group) on a 7-point scale indicated that that they were satisfied with the virtual patient simulation (mean 6.06, SD 0.71), highly positive about the quality of the system (mean 6.01, SD 0.56) and information (mean 6.06, SD 0.50), and perceived highly the net benefits (mean 6.28, SD 0.59) of the program.

Figure 4. Performance mean scores and standard deviation at pretest, first posttest, and second posttest.



Discussion

Principal Findings

The virtual patient simulation for this study was developed through a systematic and comprehensive evaluation of a mannequin-based simulation program known as RAPIDS program. Programmatic research was conducted previously from the beginning, from performing needs analysis to evaluating the long-term outcomes of the RAPIDS program in clinical practice [1-3]. The scenarios were developed based on common acute care events and opinions of clinical experts. A variety of conceptual frameworks were applied in developing and implementing the virtual patient simulation. The ABCDE and SBAR mnemonics, accepted guidelines for the care of critically ill patients, form the foundation for the training contents. Kolb's experiential learning guided the design and process of the virtual patient simulation in which the learners began the learning process in each scenario with a simulation experience followed by self-reflection [2].

The effectiveness of the virtual patient simulation was evaluated by comparing it with mannequin-based simulation. Comparing between these two simulations modalities, we acknowledged that the variations in structures and designs were recognized as confounding variables. It is important to address the confounding variables for the observed results [15]. Using a self-directed learning approach, the virtual patient simulation provided learners control of their training agenda, allowing repeated "deliberate practice" and receiving standardized feedback. The mannequin-based simulation used a collaborative learning approach that required the learners to perform hands-on simulation in small groups and engaged in a group debrief led by a trained facilitator. With this confounding variable, it is crucial to provide evidence to support the validity of the outcome measure [15]. The outcome measure selected for this study was closely aligned with the learning objectives. In both the virtual patient simulation and mannequin-based simulation, the ABCDE and SBAR mnemonics provided the frameworks to guide learning of the training contents. They were also constructs for the RAPIDS evaluation tool, which was previously validated and used to measure the performance outcomes during the simulation-based assessment [1]. The validity of the RAPIDS tool has been well established in a previous study, based on consensus among a panel of clinical experts [13].

The performance scores (baseline) from the simulation-based assessment, conducted 8 months following the full-scale RAPIDS simulation course, indicated that there is much room for the participants to improve their performance in assessing and managing patient deterioration. We found that both virtual patient simulation and mannequin-based simulation are effective refresher learning strategies to improve the nursing students' clinical performance. Earlier studies did not support the use of virtual patient as part of a blended approach to advanced cardiac life support training (ACLS) [16,17]. It was suggested that virtual patient simulation may not be the best modality for learning ACLS, which is based on algorithmic approaches and psychomotor skill development, but better used to develop decision making skills [18]. In this study, we designed virtual patient simulation to promote clinical reasoning skills through deliberate practice with multiple and varied simulation scenarios. Despite the significant improvement of the posttest scores in both the experimental and controls groups, the brief refresher training had generally not achieved an optimal level of skill improvement. The learning from the virtual patient simulation could be maximized by allowing the learners to undertake the learning on a regular basis. The positive evaluation of the virtual

patient simulation by the nursing students supported the use of this learning strategy in the nursing undergraduate curriculum in our institution.

We found that both virtual patient and mannequin-based simulation were effective in improving nursing students' ability to assess and manage a deteriorating patient. This finding is consistent with a previous study that demonstrated the effectiveness of both virtual simulation and mannequin-based simulation for improving teamwork skills [4]. We incorporated the features for effective learning identified by a systematic review into our simulation design [19]. These principles included curriculum integration, range of training levels, clinical variation, multiple learning strategies, deliberate practice, and feedback. The virtual patient simulation was integrated into a final year nursing module within the existing undergraduate curricula. It was used as a self-directed learning strategy to maintain competence following a RAPIDS course. A variety of simulation scenarios was developed for virtual patient simulation to provide clinical variation. These scenarios were sequenced across a range of training levels, progressively from simple to complex problem solving, to facilitate deliberate practice. In each scenario, the learners engaged in active learning by playing the role of a nurse in assessing and managing the deteriorating virtual patient. The learners received direct real-time feedback from the computer software with pre-programmed patient responses based on their nursing actions. At the end of each scenario, the learner also received feedback via the scoring system and checklist that appears in the debriefing screen. As our study provided positive learning outcomes of the virtual patient simulation, we call upon educators to apply the principles of simulation in the design and implementation of virtual patient learning experiences.

Although the use of virtual patient simulation yielded immediate improvement in the participants' clinical performance, this improvement was not sustained at 2.5 months. The participants in the mannequin-based simulation, however, demonstrated a more consistent and sustained improvement at 2.5 months, with little decay over time. This retention of learning suggests that hands-on simulation provided deeper learning compared to multimedia teaching modalities, which is consistent with the result of a previous study [20]. In our study, the virtual patient simulation exposed the participants to a larger number of cases with repetitive practice and provided a significant amount of didactic information in the feedback section. These aspects were much less integrated into the mannequin-based simulation, which also relied on collaborative learning and facilitator-led debriefing. The social interaction underlying collaborative learning has been found to promote long-term retention of the learned material [21]. In addition, learning gets deepened with guided and thoughtful reflection rather than mere simulation experiences and customized feedback [22].

As both learning tools are associated with improved outcome, comparing their resource intensity may lead to better-informed curricular decisions [23]. The virtual patient simulation was a more resource-efficient simulation modality than the mannequin-based simulator for a refresher training course. Although there were initial start-up costs for developing the virtual patient simulation, its implementation was less resource intensive than the mannequin-based simulation. The use of virtual patient simulation does not require simulation facilitators, expensive equipment, or facilities, which are associated with high costs. In addition, it can accommodate multiple users at one time and provide learning content for a large group of learners, all of which can facilitate efficiencies in curriculum planning and use of instructor time [24]. As a result, the use of virtual patient simulation would be a viable option in institutions with large numbers of learners [25]. While it is also feasible to repeat the mannequin-based simulation to maintain competence, it has a higher resource demand. The opportunities to engage in repetitive training in the virtual patient simulation are unlimited. This repetitive learning is essential if the learner is to achieve long-term retention of learning [26].

Limitations

There are limitations that warrant attention. First, when comparing the virtual patient simulation with the mannequin-based simulation, it was challenging to account for all the differences in the simulation designs and structure. As such, the comparison was confounded [27], and the findings may not be reliably generalized. Second, we did not measure the level of performance immediately after the 6-hour RAPIDS simulation course to determine the level of deterioration prior to the refresher training. Third, in the present study, both groups were given the same duration of training. We did not optimize the use of the virtual patient simulation by allowing the experimental group to undertake the learning on a regular basis. Future study is needed to find out whether frequent exposure of virtual patient simulation could prevent the decay of learning. Fourth, due to faculty and students' time constraints, the outcome measure was limited to immediate and 2.5 months following the intervention. Future study could evaluate the competence over a longer period of time. Fifth, we did not evaluate the learning experience from the control group. A comparative study on learners' experiences with the two types of simulation modalities could be considered for future study. Last, the outcome measure was limited to individual simulation-based assessment, which may bias towards the experimental group. Future studies could determine the outcomes of the virtual patient simulation on actual clinical practice.

Conclusions

A virtual patient simulation, known as e-RAPIDS, was developed as a refresher training course to enhance nursing students' clinical performance in rescuing a patient in a deteriorating situation. Although the study did not show the superiority of virtual simulation over mannequin-based simulation, both simulations have shown to be effective learning strategies for improving clinical performance in assessing and managing clinical deterioration. Overall, the learners evaluated the virtual patient simulation very positively. Given the flexibility, practicality, and scalability of the virtual patient simulation, it appears to provide a more promising learning strategy over time than the mannequin-based simulation for refreshing clinical performance. Further studies can build on this to provide more evidence on blended learning, where the

virtual simulation is integrated with the mannequin-based simulation.

Acknowledgments

The authors would like to thank Fui Ping Lim and Lai Fun Wong and nursing students of the National University of Singapore (NUS) for their assistance with this study. We also thank staff of the Centre for Healthcare Simulation for their support with this study and Hung Chew Wong for her statistical advice. This study was supported by NUS Centre for Instructional Technology and funded by a teaching enhancement grant from NUS Center for Development of Teaching and Learning to Alice Lee Centre for Nursing Studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [28].

[PDF File (Adobe PDF File), 990KB - jmir_v16i9e214_app1.pdf]

References

1. Liaw SY, Scherpbier A, Klainin-Yobas P, Rethans JJ. A review of educational strategies to improve nurses' roles in recognizing and responding to deteriorating patients. Int Nurs Rev 2011 Sep;58(3):296-303. [doi: 10.1111/j.1466-7657.2011.00915.x] [Medline: 21848774]
2. Liaw SY, Rethans JJ, Scherpbier A, Piyanee KY. Rescuing A Patient In Deteriorating Situations (RAPIDS): A simulation-based educational program on recognizing, responding and reporting of physiological signs of deterioration. Resuscitation 2011 Sep;82(9):1224-1230. [doi: 10.1016/j.resuscitation.2011.04.014] [Medline: 21664026]
3. Liaw SY, Chan SW, Scherpbier A, Rethans JJ, Pua GG. Recognizing, responding to and reporting patient deterioration: transferring simulation learning to patient care settings. Resuscitation 2012 Mar;83(3):395-398. [doi: 10.1016/j.resuscitation.2011.08.021] [Medline: 21946262]
4. Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P. Design, development, and evaluation of an online virtual emergency department for training trauma teams. Simul Healthc 2008;3(3):146-153. [doi: 10.1097/SIH.0b013e31817bedf7] [Medline: 19088658]
5. Wiecha J, Heyden R, Sternthal E, Merialdi M. Learning in a virtual world: experience with using second life for medical education. J Med Internet Res 2010;12(1):e1 [FREE Full text] [doi: 10.2196/jmir.1337] [Medline: 20097652]
6. Triola M, Feldman H, Kalet AL, Zabar S, Kachur EK, Gillespie C, et al. A randomized trial of teaching clinical skills using virtual and live standardized patients. J Gen Intern Med 2006 May;21(5):424-429 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00421.x] [Medline: 16704382]
7. Ellaway R, Poulton T, Fors U, McGee JB, Albright S. Building a virtual patient commons. Med Teach 2008;30(2):170-174. [doi: 10.1080/01421590701874074] [Medline: 18464142]
8. Huang G, Reynolds R, Candler C. Virtual patient simulation at US and Canadian medical schools. Acad Med 2007 May;82(5):446-451. [doi: 10.1097/ACM.0b013e31803e8a0a] [Medline: 17457063]
9. Papadopoulos L, Pentzou AE, Louloudiadis K, Tsiatsos TK. Design and evaluation of a simulation for pediatric dentistry in virtual worlds. J Med Internet Res 2013;15(11):e240 [FREE Full text] [doi: 10.2196/jmir.2651] [Medline: 24168820]
10. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in health professions education: a systematic review and meta-analysis. Acad Med 2010 Oct;85(10):1589-1602. [doi: 10.1097/ACM.0b013e3181edfe13] [Medline: 20703150]
11. LeFlore JL, Anderson M, Zielke MA, Nelson KA, Thomas PE, Hardee G, et al. Can a virtual patient trainer teach student nurses how to save lives--teaching nursing students about pediatric respiratory diseases. Simul Healthc 2012 Feb;7(1):10-17. [doi: 10.1097/SIH.0b013e31823652de] [Medline: 22228285]
12. Liaw SY, Scherpbier A, Klainin-Yobas P, Rethans JJ. Rescuing A Patient In Deteriorating Situations (RAPIDS): an evaluation tool for assessing simulation performance on clinical deterioration. Resuscitation 2011 Nov;82(11):1434-1439. [doi: 10.1016/j.resuscitation.2011.06.008] [Medline: 21794972]
13. Wang YS, Wang HY, Shee DY. Measuring e-learning systems success in an organizational context: Scale development and validation. Computers in Human Behavior 2007 Jul;23(4):1792-1808. [doi: 10.1016/j.chb.2005.10.006]
14. Liaw SY, Scherpbier A, Rethans JJ, Klainin-Yobas P. Assessment for simulation learning outcomes: a comparison of knowledge and self-reported confidence with observed clinical performance. Nurse Educ Today 2012 Aug;32(6):e35-e39. [doi: 10.1016/j.nedt.2011.10.006] [Medline: 22064013]
15. Cook DA. If you teach them, they will learn: why medical education needs comparative effectiveness research. Adv Health Sci Educ Theory Pract 2012 Aug;17(3):305-310. [doi: 10.1007/s10459-012-9381-0] [Medline: 22696095]

16. Perkins GD, Fullerton JN, Davis-Gomez N, Davies RP, Baldock C, Stevens H, et al. The effect of pre-course e-learning prior to advanced life support training: a randomised controlled trial. Resuscitation 2010 Jul;81(7):877-881. [doi: 10.1016/j.resuscitation.2010.03.019] [Medline: 20409627]
17. Jensen ML, Mondrup F, Lippert F, Ringsted C. Using e-learning for maintenance of ALS competence. Resuscitation 2009 Aug;80(8):903-908. [doi: 10.1016/j.resuscitation.2009.06.005] [Medline: 19570601]
18. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Med Educ 2009 Apr;43(4):303-311. [doi: 10.1111/j.1365-2923.2008.03286.x] [Medline: 19335571]
19. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
20. Delasobera BE, Goodwin TL, Strehlow M, Gilbert G, D'Souza P, Alok A, et al. Evaluating the efficacy of simulators and multimedia for refreshing ACLS skills in India. Resuscitation 2010 Feb;81(2):217-223. [doi: 10.1016/j.resuscitation.2009.10.013] [Medline: 19926385]
21. Kreijns K, Kirschner PA, Jochems W. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Computers in Human Behavior 2003 May;19(3):335-353. [doi: 10.1016/S0747-5632(02)00057-2]
22. Edelbring S. Research into the use of virtual patients is moving forward by zooming out. Med Educ 2013 Jun;47(6):544-546. [doi: 10.1111/medu.12206] [Medline: 23662870]
23. Isaranuwatchai W, Brydges R, Carnahan H, Backstein D, Dubrowski A. Comparing the cost-effectiveness of simulation modalities: a case study of peripheral intravenous catheterization training. Adv Health Sci Educ Theory Pract 2014 May;19(2):219-232. [doi: 10.1007/s10459-013-9464-6] [Medline: 23728476]
24. Cook DA, Levinson AJ, Garside S. Time and learning efficiency in Internet-based learning: a systematic review and meta-analysis. Adv Health Sci Educ Theory Pract 2010 Dec;15(5):755-770. [doi: 10.1007/s10459-010-9231-x] [Medline: 20467807]
25. Bonnetain E, Boucheix JM, Hamet M, Freysz M. Benefits of computer screen-based simulation in learning cardiac arrest procedures. Med Educ 2010 Jul;44(7):716-722. [doi: 10.1111/j.1365-2923.2010.03708.x] [Medline: 20636591]
26. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004 Oct;79(10 Suppl):S70-S81. [Medline: 15383395]
27. Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. Acad Med 2005 Jun;80(6):541-548. [Medline: 15917356]
28. 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

ABDCDE: airway, breathing, circulation, disability, and expose/examine
ACLS: advanced cardiac life support
ANOVA: analysis of variance
NUS: National University of Singapore
RAPIDS: rescuing a patient in deteriorating situations
SBAR: situation, background, assessment, and recommendation

Edited by G Eysenbach; submitted 15.02.14; peer-reviewed by S Feaster, C Georg, C Bond; comments to author 16.07.14; revised version received 04.08.14; accepted 21.08.14; published 17.09.14.

Please cite as:
Liaw SY, Chan SWC, Chen FG, Hooi SC, Siau C
Comparison of Virtual Patient Simulation With Mannequin-Based Simulation for Improving Clinical Performances in Assessing and Managing Clinical Deterioration: Randomized Controlled Trial
J Med Internet Res 2014;16(9):e214
URL: http://www.jmir.org/2014/9/e214/
doi:10.2196/jmir.3322
PMID:25230684

Original Paper

Preparing Facilitators From Community-Based Organizations for Evidence-Based Intervention Training in Second Life

Angel Felix Valladares[1], MPH; Michelle Aebersold[1], RN, PhD; Dana Tschannen[1], RN, PhD; Antonia Maria Villarruel[1], RN, FAAN, PhD

University of Michigan, School of Nursing, Ann Arbor, MI, United States

Corresponding Author:
Michelle Aebersold, RN, PhD
University of Michigan
School of Nursing
400 N Ingalls Street
4156
Ann Arbor, MI, 48109
United States
Phone: 1 (734) 615 8498
Fax: 1 (734) 647 2416
Email: mabersol@umich.edu

Abstract

Background: A major barrier to the use and scale-up of evidence-based interventions are challenges related to training and capacity building. A cost-effective and highly interactive multi-user virtual environment, Second Life (SL) is a promising alternative for comprehensive face-to-face facilitator training.

Objective: The purpose of this study was to examine the feasibility of using SL to train facilitators from community-based organizations to use ¡Cuídate! (Take Care of Yourself), one of the few evidence-based interventions developed and tested with Latino youth to reduce sexual risk behaviors.

Methods: We recruited 35 participants from community-based organizations throughout the United States to participate in the SL ¡Cuídate! Training of Facilitators. Preparation to use SL consisted of four phases: (1) recruitment and computer capacity screening, (2) enrollment, (3) orientation to the SL program, and (4) technical support throughout the synchronous training sessions. Technical difficulties, the associated cause, and the mitigation strategy implemented were recorded during each session. Participants completed evaluations including perceptions of self-efficacy and confidence to complete the necessary skills to participate in SL training.

Results: Overall, participants reported high levels of self-efficacy for all skills necessary to participate in SL training. Based on an 11-point scale (0-10), self-efficacy to download and access the software was rated the highest: mean 8.29 (SD 2.19). Interacting with items in SL had the lowest mean score: mean 7.49 (SD 2.89). The majority of technical difficulties experienced by participants were related to inadequate Internet connections or computer malfunctions.

Conclusions: Our findings support the feasibility of using SL for the ¡Cuídate! Training of Facilitators. The process used in this study to prepare participants to use SL can be used as a basis for other evidence-based intervention training in SL. This study is an important contribution to developing cost-effective and accessible training options for evidence-based interventions.

(J Med Internet Res 2014;16(9):e220) doi:10.2196/jmir.3606

KEYWORDS

evidence-based intervention; virtual environment; Second Life; health education; community-based organization; facilitator training; computer simulation; distance education; serious gaming

Introduction

Background

Multi-user virtual environments (MUVEs), such as Second Life (SL), are a promising alternative to face-to-face training for health care program delivery [1]. While rigorously tested evidence-based interventions (EBIs) have the potential to influence health outcomes, challenges related to training and capacity building access (eg, funding, high travel costs, limited to no technical assistance offered, and limited guidance on appropriate adaptations) [2,3] can impede implementation fidelity and the effectiveness of the EBI. Due to these challenges, more accessible training modalities on how to deliver the intervention accurately are needed to better disseminate EBIs and to ensure their effectiveness. Virtual environments such as SL have not been extensively used for EBI training of individuals who work in community-based organizations and deliver training to adolescents (ie, facilitators). This is important as facilitators deliver the curriculum to the adolescents and facilitate group discussions and other activities that are part of the curriculum. As a cost-effective and highly interactive MUVE, SL has potential to serve as a delivery platform for comprehensive EBI facilitator training.

A MUVE is a computer-based, simulated virtual environment allowing users to inhabit an online virtual world and interact with others via self-representations known as avatars in synchronous sessions [1]. The advantage of MUVEs like SL is that they increase access to capacity building opportunities [2] and are a promising alternative to face-to-face training due to substantial features mirroring human interactions. These features include advanced and realistic voice chat, speech gestures, and the ability to manipulate voices in a multidimensional space just as one would experience in real life [4]. This study describes the process of preparing individuals from community-based organizations to participate in a SL ¡Cuídate! (Take care of yourself) Training of Facilitators [5].

Second Life Training

SL has been previously used for small-scale health-related programs focusing on education and awareness, support, training, marketing, and promotion of health services [1,6-11]. For example, SL was used to implement a 1-hour seminar to enhance clinician knowledge of insulin therapy for type 2 diabetics for primary care physicians (n=14) [9]. Results of the training indicated a positive impact on self-efficacy to perform insulin therapy, gains in clinical competence around the use of insulin therapy, and positive feedback on the sense of presence attributed to the use of avatars. There was high variability in the SL user learning curve depending on past experience with video games, but an average of 78 minutes was spent with each participant to gain proficiency in SL. Similarly, Mitchell et al pilot-tested motivational interviewing training in SL with primary care physicians (n=13) focusing on colorectal cancer screening [10]. A self-directed approach to learning SL skills was used to reduce the amount of time necessary to prepare participants in SL. The actual amount of time participants spent on preparation was not explicitly reported. Finally, Tschannen et al used SL for diabetes self-management training for nurses. Participants reported that they were highly satisfied with the virtual simulation experience. Further, participants' ability to apply knowledge gained was comparable to those who completed a face-to-face training alternative [11]. While these studies demonstrate the utility of SL for training, none of these studies reported quantitative data on SL proficiency outcomes, or self-efficacy of SL skills. Further, training participants were limited to highly trained clinicians.

Technical Challenges

Virtual training presents a number of technical challenges, including computer system capacity and security issues [10]. For example, training participation requires minimum system specifications, installation of a separate computer program, and the development of specific computer skills to facilitate the interaction required in SL to engage fully in the training. Specifically, SL requires a high level of processing power, graphical memory, and a need for a large bandwidth with high-speed Internet access [12]. Individuals may encounter network policies that do not allow access to public virtual worlds like SL or allow downloads of software or a continuous Internet connection [9]. Despite the technological complexities, SL has proven to be an accessible medium for effective training and education.

Purpose

The purpose of this study was to examine the feasibility of using SL to train facilitators to use an EBI, ¡Cuídate! (Take Care of Yourself), one of the few sexual risk reduction interventions developed and tested with Latino youth [13]. The participants received training in SL to prepare them to deliver the ¡Cuídate! curriculum to Latino youths in their community. This paper examines the management of technical issues and participant self-evaluation of SL skills.

Methods

Overview

This is a descriptive study. The study protocol was reviewed by the institutional review board at the University of Michigan and was deemed exempt and not regulated.

A virtual training center (Figure 1) was developed in SL to deliver the ¡Cuídate! training. This training center was built on university-owned space in SL (Wolverine Region). The training center consisted of a conference room with a table and chairs and a training room with movable chairs, which were scripted to provide avatars with gestures such as hand-raising. The training was set up with Web prims to display interactive posters and Google Docs.

Preparation to use SL consisted of four phases: (1) recruitment and computer capacity screening, (2) enrollment, (3) orientation to the SL program, and (4) technical support throughout the synchronous training sessions.

Figure 1. Screenshot of the ¡Cuídate! training center in Second Life.

Phase One: Recruitment and Computer Capacity Screening

Community-based organizations throughout the United States were sent information about the SL Cuídate Training of Facilitators opportunity via social media (eg, Facebook, Twitter), distribution lists, and newsletters. To ensure agency staff could feasibly participate in the SL training, an online portal and a screening process were established. Community-based organizations submitted contact information and answered screening questions to determine if their computer system capacity met SL's system requirements. The minimum system requirements for SL include at least a Windows XP or Mac OS X 10.6 operating system, a graphics card model NVIDIA GeForce 6600 or better, an ATI Radeon 9500 or better, or an Intel 945 chipset for three-dimensional functionality, and a high-speed Internet connection [12]. To identify possible SL installation barriers, access to the Internet (ie, wired or wireless connection) and availability of information technology (IT) personnel at their facility to assist in mitigating technical difficulties was assessed.

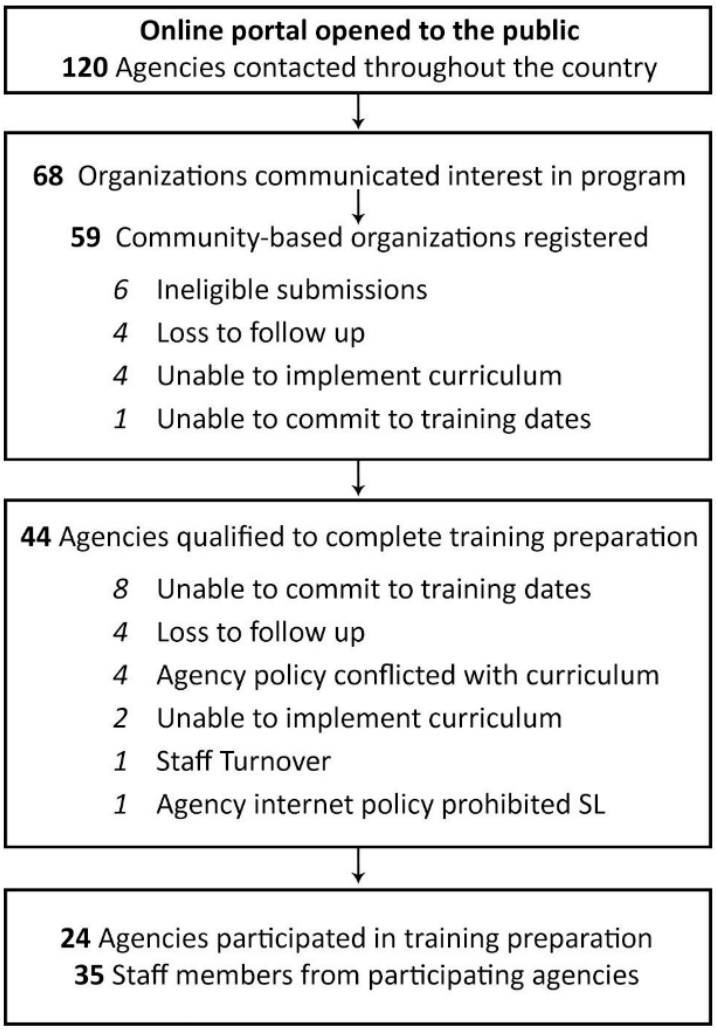
If registrants had questionable capacity (eg, older operating systems or a low-power graphics card), research staff guided each registrant through an online scanner available from System Requirements Lab, Inc. [14]. If minimum requirements were met, instructions for downloading and installing the program were provided. In rare cases, a member of the study team walked through the installation steps over the phone with them or their IT personnel.

Phase Two: Enrollment

Once an agency met the minimum computer system capacity, they were presented with a Memorandum of Understanding (MOU) that included information on the purpose of the study, responsibilities of participants, and expectations and accountability of the research team. Participants received an enrollment phone call to review the MOU and were provided with an outline of the project activities. Once the MOU was signed, a ¡Cuídate! Implementation Kit [15] and a universal serial bus (USB) microphone headset were mailed to participants. The USB microphone headsets were sent to standardize equipment for easier troubleshooting of SL voice chat difficulties during training. For their participation, community-based organizations were provided a financial incentive of US $500.00 to deliver ¡Cuídate! in their community post-participation in the Training of Facilitators.

Communication occurred with 120 different community agencies and entities throughout the United States via recruitment phone calls. Of those, personnel from 59 agencies underwent the screening process previously described. Of the agencies that met computer system capacity, 55% (24/44) committed to the training and completed the tasks required to participate. Details of the recruitment and screening process are reported in Figure 2.

Figure 2. Training participant recruitment and screening process.



Phase Three: Orientation to Second Life

Participants were oriented to SL in three steps: (1) self-directed online orientation manual, (2) a one-on-one SL review session with research staff, and (3) a group skill review on SL training day one. In the first step, we developed an online orientation manual that provided instructions on how to install SL, create an avatar, perform basic navigation functions, teleport to different parts of the environment, manipulate camera angles, and use text and voice communication features. The online manual consisted of text and screenshots of steps and guided video demonstrations embedded within the manual.

In the next step, one-on-one meetings in SL were scheduled to review and practice basic skills and to introduce participants to more complex skills. These additional skills focused on interactions with three-dimensional objects in the program including seats, working on collaborative Google documents [16], changing their display, and participating in small group voice chats. The SL ¡Cuídate! training was developed to limit participant burden with learning SL to ensure the focus of the training was on learning the ¡Cuídate! Facilitator's Curriculum.

The third step was incorporated into the first day of the SL facilitator training. This orientation session served as an introduction to the ¡Cuídate! training and provided an opportunity to practice features of the SL environment relevant to training. Best practices in SL were shared to ensure a smooth training experience. The practices included silencing microphones when they were not speaking to minimize distracting noises, ensuring their avatars remained active by participating in discussions, raising their avatar's hand when they had a question or comment, and how best to view the interactive documents and posters during the training.

Phase Four: Technical Support During Training

In the training center, a designated staff member with high SL proficiency was available to provide technical support. Participants were asked to sign in 15 minutes prior to scheduled SL ¡Cuídate! training and to send a message using text chat to the technical staff if they were having trouble during the session. The staff member would either work in the background with the participant or have them teleport (a quick navigation method in SL) to another location to troubleshoot issues, thus minimizing distractions for other participants.

Measures

Overview

To evaluate the process of preparing individuals to participate in SL training, several different descriptive methods were used. Data were gathered from the participant screening survey, technical support sessions, and a post-orientation survey. The surveys were administered using Qualtrics. Data were then categorized in three different areas for analysis: computer system capacity and experience, SL self-efficacy, and participant preparation time.

Computer System Capacity and Experience

A measure of overall computer system capacity consisted of two components: technical specifications and participant experience with online environments. During enrollment, participants were asked about technical specifications related to their operating system and Internet connectivity. Participants were also asked about previous online learning experiences (Yes/No) and if they had used SL prior to this training (Yes/No). They were also asked to report the frequency of engagement in online and gaming software on a 3-point scale (1=rarely, 2=often, 3=very often).

Second Life Self-Efficacy

Self-efficacy to use SL was assessed with nine questions adapted from a scale developed by Gist, Schwoerer, and Rosen, which was developed to measure computer and software self-efficacy [17]. The first three questions (eg, User Capability to Navigate SL) measured the participant's ability to navigate SL after using either the written instructions or the video demonstrations (both part of the orientation manual). The next six questions (eg, Self-Efficacy of SL Skills) in the scale measured the self-efficacy of participants to (1) download SL software, (2) create an avatar, (3) walk with their avatar, (4) teleport to different locations, (5) change the view in SL, and (6) interact with items in SL. The questions were on an 11-point scale, from 0 (completely unable to do the task) to 10 (totally capable of doing the task). Reliability coefficient alpha=.83 for the User Capability to Navigate SL (3 questions) and .87 for the Self-Efficacy of SL Skills (6 questions).

Participant Preparation Time

Two questions assessed participants' estimates of the time spent reviewing the SL orientation manual. Additionally, the research team recorded actual time spent supporting participants in the use of SL or addressing computer issues. Technical difficulties experienced by participants, the associated cause, and the mitigation strategy implemented were also recorded throughout the training and preparation.

Results

Overview

A total of 35 participants from 24 agencies participated in 1 of 5 cohorts of SL ¡Cuídate! training. The majority of participants were female (69%, 24/35) and ranged from 20-59 years of age. Most participants had received at least a Bachelor's degree (80%, 28/35). Slightly more than half of the participating community-based organizations had fewer than 50 employees (58.3%, 14/24), while 10 community-based organizations reported having more than 50 employees (41.7%, 10/24).

Computer System Capacity and Experience

Most participants had Windows Vista or better (74%, 26/35), while only a few reported using Windows XP (14%, 5/35), or Mac OSX 10+ (9%, 3/35). Regarding Internet connections, nearly equal numbers of individuals (40%, 14/35) reported having a wired connection or only a wireless connection (37%, 13/35), while a small percentage of participants (23%, 8/35) had access to both wired and wireless connections.

A majority of participants (77%, 27/35) reported having online learning experience as well as experience with computer and video games, but very few participants (9%, 3/35) had ever used SL. Of those reporting having played computer or video games in the past 12 months (51%, 18/35), 2 participants (11.1%) played very often, 3 (16.7%) often, and 13 (72%) had rarely played.

Second Life Self-Efficacy

Most individuals rated their confidence levels high (no mean scores were below 7) on the SL self-efficacy scale. Participants reported high capability to navigate SL after reviewing written instructions and additional video clips (mean 7.97, SD 2.19). Participants had high capability scores for skills such as downloading the SL program and creating an avatar (mean 8.29, SD 2.19, and mean 8.03, SD 2.19 respectively), but slightly lower scores for changing the view on their computer screen and interacting with items in SL (mean 7.49, SD 2.89, and mean 7.49, SD 2.92 respectively). See Table 1.

Table 1. Evaluation of SL orientation manual and self-efficacy of SL skills (N=35).

SL skills	Mean (SD)[a]
User capability to navigate SL	
Without reviewing SL orientation materials	6.80 (2.69)
Reviewing only written instructions	7.80 (1.84)
Reviewing additional video clips	7.97 (2.19)
Self-efficacy of SL skills	
Downloading the SL software	8.29 (2.19)
Creating my avatar	8.03 (2.08)
Making my avatar walk	7.91 (2.79)
Changing the view on my computer screen	7.49 (2.92)
Teleporting to different locations	8.00 (2.57)
Interacting with items in SL	7.49 (2.89)

[a]Mean and standard deviation were derived from an 11-point scale (0=unable, 10=totally capable).

Participant Preparation Time

The majority of participants (91%, 32/35) reviewed both the written instructions as well as the video demonstrations in the SL orientation manual. The amount of time participants spent using the orientation manual ranged from less than 1 hour (43%, 15/35), between 1 and 2 hours (17%, 6/35) or more than 2 hours (40%, 14/35). It is unknown how many more hours were spent using the orientation manual as the scale ended at 2 plus hours.

The one-on-one session in SL required an average of 28.8 (SD 11.8) minutes per participant with a range of 15-60 minutes. Troubleshooting technical issues required an average of 28.4 (SD 16.4) minutes per participant for 74% (26/35) of the participants. The shortest troubleshooting time (2 minutes) was used to correct simple settings issues, and the longest (75 minutes) was used to assist participants with installation difficulties.

Some participants had issues with SL installation due to either agency policies (n=3) or inadequate skill (n=3). Some agencies had policies prohibiting access to public virtual worlds like SL. The most prominent technical difficulty after installation was headset malfunction and incompatibility with users' computers. The majority of these issues were corrected by adjusting settings on the operating system or on the SL program itself. For others, there were SL malfunctions such as lags in performance, periodic interruptions to audio or an ejection from the program. Internet connectivity was thought to be the root cause of these performance issues, particularly those with wireless Internet connections. The overwhelming majority of participants who continued to receive troubleshooting during the training reported wireless Internet connections. Technical difficulties and support are highlighted in Table 2.

Table 2. Participant technical support and troubleshooting (N=26).

Technical support issue	Instances[a], N	Possible causes and mitigation strategy
Phase 2: Enrollment		
Installation barriers	6	3 participants had insufficient knowledge to install SL on their own and 3 due to their agency's network policies; all required assistance installing the program.
Second Life installation error	1	Laptop PC could not complete installation; participant used another computer for the training.
Phase 3: Orientation		
Headset functioning/ connectivity	15	Linked to Internet connection quality issues as well as settings compatibility with SL and the PC. Audio settings in the operating system and SL were reset, SL was also restarted.
Could not connect to SL Server	4	Loss of Internet connectivity or use of incompatible Internet devices (eg, cellular Internet hotspots). Participants were required to access SL at an alternate location.
Phase 4: Technical Troubleshooting in Training		
Headset functioning/connectivity	12	Continued connection issues contributed to additional headset issues in training. Audio settings were adjusted, and some restarted SL to regain audio connection.
SL user navigation issues	5	Users accidentally interact with different features of the program interrupting their view or changing their location. These issues required technical staff to help them navigate back to the correct view or location.

[a]Instances are not equal to unique individuals and include multiple troubleshooting sessions with participants.

Discussion

Principal Findings

The purpose of this study was to describe the process and related challenges of preparing individuals to participate in an EBI facilitator training program adapted for SL. The response to initial recruitment for participants was positive despite training being offered in a very non-traditional format (68/120 agencies indicated interest). Throughout the screening process, a variety of agencies were able to meet technical requirements and complete the training, half of which were small community-based agencies (<50 employees). Reasons that agencies did not participate, despite initial interest, were primarily related to inadequate computer system capacity and/or the inability to implement the training within 6 months post training. The variation in computer system capacity and Internet connections of the agencies that participated allowed for the opportunity to test the feasibility of using SL in areas that do not have available technical support staff or high-speed Internet capabilities.

In terms of SL skills orientation, the majority of participants reported high confidence levels for basic skills in SL. This supports the feasibility of using a self-directed orientation manual since participants were able to reasonably orient themselves to the SL environment without additional support. Further, the use of a self-directed orientation manual resulted in limited staff time (less than one hour) for SL orientation and limited time necessary for additional troubleshooting. The time spent on orientation is less than previous reported results of Wiecha et al who spent 78 minutes coaching each participant [9]. Findings from this study provide the overall process that can be implemented to facilitate other similar trainings in SL.

Lessons Learned

In this study, participants did identify technical challenges that impacted facilitator training. Major difficulties encountered included hardware and software compatibility issues and Internet issues (bandwidth) and not with individual ability to use the SL software (ratings for self-efficacy were high and very few troubleshooting encounters were reported requiring assistance with SL program skills). Although system requirements were identified pre-training, further testing is needed to determine specifically what system requirements should be mandated in order to prevent technical hurdles. For some, the use of a laptop interfered with a seamless experience because some laptops are not designed with the capacity (ie, adequate graphic cards, reliance on wireless) to run even less burdensome three-dimensional environments such as SL. Of these connection-related difficulties, audio input and output malfunctions were most prevalent even when using a standard audio headset and after pre-training audio testing.

One unanticipated cause of these issues was the load on an individual's network and computer system when multiple users were simultaneously logged on to the virtual environment. Smaller training cohorts (fewer than 7 participants) had fewer connection-related issues in the live training. Therefore, providing installation instructions upon initial registration may allow registrants to test their bandwidth by having them enter more populated areas of SL to see if their computers can handle a large number of users. This early bandwidth testing should reduce the number of technical difficulties that may arise in the training program from a lack of bandwidth capacity. One additional approach that may be warranted is to limit the amount of participants per training session to minimize the technical challenges that may occur.

Conclusions

Technical support was a vital component of this study. Delivering training in SL requires technical expertise to remotely troubleshoot with participants. Advanced knowledge of both computer hardware and software components is essential for the technical staff who will provide the support before and during training. By having support staff available throughout the training, major interruptions were avoided and participant burden was minimized allowing for the successful retention of all 35 participants across 3 days of training.

In summary, preparing participants to use SL for the ¡Cuídate! Training of Facilitators was feasible, and participants were confident in their ability to use the program for training after completing orientation. In addition, we developed processes to decrease technical difficulties prior to and during training. The potential advantages of conducting SL trainers include the ability to access EBI training without the high cost of travel. This is especially relevant for agencies and communities with limited resources. Decreasing access barriers through the use of technology can facilitate the dissemination and use of evidence-based interventions, resulting in improved health outcomes.

Acknowledgments

The support for this research comes from the National Institute of Health Grant Number: 1R21MH09572301A1. The authors would like to extend a special thank you to the Cuídate training participants for their willingness to venture into the virtual world and for their insight, which helped shape this training for future iterations.

Conflicts of Interest

Dr Antonia M Villarruel is the developer of the ¡Cuídate! curriculum and normally receives royalties from the publisher related to distribution. However, Dr Villarruel waived all royalties typically provided her with the distribution of the curriculum for this study.

References

1. Boulos MN, Hetherington L, Wheeler S. Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. Health Info Libr J 2007 Dec;24(4):233-245. [doi: 10.1111/j.1471-1842.2007.00733.x] [Medline: 18005298]
2. Veniegas RC, Kao UH, Rosales R, Arellanes M. HIV prevention technology transfer: challenges and strategies in the real world. Am J Public Health 2009 Apr;99 Suppl 1:S124-S130 [FREE Full text] [doi: 10.2105/AJPH.2007.124263] [Medline: 19218184]
3. Cunningham SD, Card JJ. Realities of replication: implementation of evidence-based interventions for HIV prevention in real-world settings. Implement Sci 2014;9:5 [FREE Full text] [doi: 10.1186/1748-5908-9-5] [Medline: 24387142]
4. Linden Lab. Using voice chat. 2011 Feb 24. URL: http://community.secondlife.com/t5/English-Knowledge-Base/Using-voice-chat/ta-p/700169 [accessed 2014-06-11] [WebCite Cache ID 6QG68o36V]
5. Linden Lab. Second Life. 2014. URL: http://secondlife.com/ [accessed 2014-06-11] [WebCite Cache ID 6QG5tOKiF]
6. Ghanbarzadeh R, Ghapanchi AH, Blumenstein M, Talaei-Khoei A. A decade of research on the use of three-dimensional virtual worlds in health care: a systematic literature review. J Med Internet Res 2014;16(2):e47 [FREE Full text] [doi: 10.2196/jmir.3097] [Medline: 24550130]
7. Beard L, Wilson K, Morra D, Keelan J. A survey of health-related activities on second life. J Med Internet Res 2009;11(2):e17 [FREE Full text] [doi: 10.2196/jmir.1192] [Medline: 19632971]
8. Johnston JD, Massey AP, Devaneaux CA. Innovation in weight loss programs: a 3-dimensional virtual-world approach. J Med Internet Res 2012;14(5):e120 [FREE Full text] [doi: 10.2196/jmir.2254] [Medline: 22995535]
9. Wiecha J, Heyden R, Sternthal E, Merialdi M. Learning in a virtual world: experience with using second life for medical education. J Med Internet Res 2010;12(1):e1 [FREE Full text] [doi: 10.2196/jmir.1337] [Medline: 20097652]
10. Mitchell S, Heyden R, Heyden N, Schroy P, Andrew S, Sadikova E, et al. A pilot study of motivational interviewing training in a virtual world. J Med Internet Res 2011;13(3):e77 [FREE Full text] [doi: 10.2196/jmir.1825] [Medline: 21946183]
11. Tschannen D, Aebersold M, Sauter C, Funnell MM. Improving nurses' perceptions of competency in diabetes self-management education through the use of simulation and problem-based learning. J Contin Educ Nurs 2013 Jun;44(6):257-263. [doi: 10.3928/00220124-20130402-16] [Medline: 23565600]
12. Linden Lab. System Requirements. 2014. URL: https://secondlife.com/support/system-requirements/ [accessed 2014-06-11] [WebCite Cache ID 6QG8ZZSqg]
13. Villarruel AM, Beakin B. ¡Cuídate! facilitator's curriculum. 2nd ed. New York: Select Media; 2012.
14. Husdawg. Can You Run It. 2014. URL: http://systemrequirementslab.com/cyri [accessed 2014-06-11] [WebCite Cache ID 6QG92vMMS]
15. Select Media. Evidence-Based Programs - ¡Cuídate!. 2014. URL: http://www.selectmedia.org/programs/cuidate.html [accessed 2014-06-11] [WebCite Cache ID 6QG9HPZZ4]
16. Google. Google Drive. 2014. URL: http://www.google.com/drive/index.html [accessed 2014-06-11] [WebCite Cache ID 6QG9PTgGi]

17. Gist M, Schwoerer C, Rosen B. Effects of alternative training methods on self-efficacy and performance in computer software training. Journal of Applied Psychology 1989;74(6):884-891. [doi: 10.1037/0021-9010.74.6.884]

Abbreviations

EBI: evidence-based intervention
IT: information technology
MOU: memorandum of understanding
MUVE: multi-user virtual environment
SL: Second Life
USB: universal serial bus

Edited by G Eysenbach; submitted 13.06.14; peer-reviewed by P Yellowlees, J Brixey; comments to author 12.08.14; revised version received 31.08.14; accepted 31.08.14; published 30.09.14.

Please cite as:
Valladares AF, Aebersold M, Tschannen D, Villarruel AM
Preparing Facilitators From Community-Based Organizations for Evidence-Based Intervention Training in Second Life
J Med Internet Res 2014;16(9):e220
URL: http://www.jmir.org/2014/9/e220/
doi:10.2196/jmir.3606
PMID:25270991

Original Paper

Replacing Ambulatory Surgical Follow-Up Visits With Mobile App Home Monitoring: Modeling Cost-Effective Scenarios

Kathleen A Armstrong[1,2], MD; John L Semple[1,3], MSc, MD; Peter C Coyte[2], MA, PhD

[1]Division of Plastic and Reconstructive Surgery, Department of Surgery, University of Toronto, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Department of Ambulatory Surgery, Women's College Hospital, Toronto, ON, Canada

Corresponding Author:
Kathleen A Armstrong, MD
Division of Plastic and Reconstructive Surgery
Department of Surgery
University of Toronto
149 College Street, 5th Floor, Room 508
Toronto, ON, M5T 1P5
Canada
Phone: 1 416 978 8534
Fax: 1 416 978 7316
Email: k.armstrong@mail.utoronto.ca

Abstract

Background: Women's College Hospital (WCH) offers specialized surgical procedures, including ambulatory breast reconstruction in post-mastectomy breast cancer patients. Most patients receiving ambulatory surgery have low rates of postoperative events necessitating clinic visits. Increasingly, mobile monitoring and follow-up care is used to overcome the distance patients must travel to receive specialized care at a reduced cost to society. WCH has completed a feasibility study using a mobile app (QoC Health Inc, Toronto) that suggests high patient satisfaction and adequate detection of postoperative complications.

Objective: The proposed cost-effectiveness study models the replacement of conventional, in-person postoperative follow-up care with mobile app follow-up care following ambulatory breast reconstruction in post-mastectomy breast cancer patients.

Methods: This is a societal perspective cost-effectiveness analysis, wherein all costs are assessed irrespective of the payer. The patient/caregiver, health care system, and externally borne costs are calculated within the first postoperative month based on cost information provided by WCH and QoC Health Inc. The effectiveness of telemedicine and conventional follow-up care is measured as successful surgical outcomes at 30-days postoperative, and is modeled based on previous clinical trials containing similar patient populations and surgical risks.

Results: This costing assumes that 1000 patients are enrolled in bring-your-own-device (BYOD) mobile app follow-up per year and that 1.64 in-person follow-ups are attended in the conventional arm within the first month postoperatively. The total cost difference between mobile app and in-person follow-up care is $245 CAD ($223 USD based on the current exchange rate), with in-person follow-up being more expensive ($381 CAD) than mobile app follow-up care ($136 CAD). This takes into account the total of health care system, patient, and external borne costs. If we examine health care system costs alone, in-person follow-up is $38 CAD ($35 USD) more expensive than mobile app follow-up care over the first postoperative month. The baseline difference in effect is modeled to be zero based on clinical trials examining the effectiveness of telephone follow-up care in similar patient populations. An incremental cost-effectiveness ratio (ICER) is not reportable in this scenario. An incremental net benefit (INB) is reportable, and reflects merely the cost difference between the two interventions for any willingness-to-pay value (INB=$245 CAD). The cost-effectiveness of mobile app follow-up even holds in scenarios where all mobile patients attend one in-person follow-up.

Conclusions: Mobile app follow-up care is suitably targeted to low-risk postoperative ambulatory patients. It can be cost-effective from a societal and health care system perspective.

(J Med Internet Res 2014;16(9):e213) doi:10.2196/jmir.3528

KEYWORDS

cost-effectiveness; ambulatory surgical procedures; mobile apps; ambulatory monitoring

Introduction

Technology is identified as an opportunity to constrain the growth in health care costs and eliminate barriers due to distance [1]. In Ontario (Canada), specialized surgical services tend to be concentrated within metropolitan areas. This results in many patients having to travel great distances to receive care. Women's College Hospital (WCH) in Toronto offers specialized ambulatory surgical procedures, including breast reconstruction following mastectomy for breast cancer. Ambulatory surgery means that the patient goes home within 24 hours of surgery and comes back at a later date for follow-up care. The average ambulatory breast reconstruction patient travels 76 km from home to hospital, with the furthest patient coming from 540 km away. Similarly in Ontario, 23% of all orthopedic surgery patients leave their local health care catchment to receive care [2].

Patients not only travel to receive care, they also travel to receive follow-up care. In an ambulatory (or outpatient) surgery patient population, travel for postoperative follow-up seems superfluous as the chance of postoperative complication is exceedingly low. This is because of advancements in surgery and rigorous patient selection. In general, ambulatory surgery is largely reserved from the treatment of American Society for Anesthesia (ASA) class I and II patients [3]. These patients are considered healthy or with mild systemic disease, respectively. Complication rates in this subset of breast reconstruction patients are approximately 5% [4]. If a complication occurs, it is typically a minor skin infection or wound dehiscence. Rarely (<1%), a hematoma requiring surgical evacuation may occur. These types of complications occur suddenly and present to the emergency department.

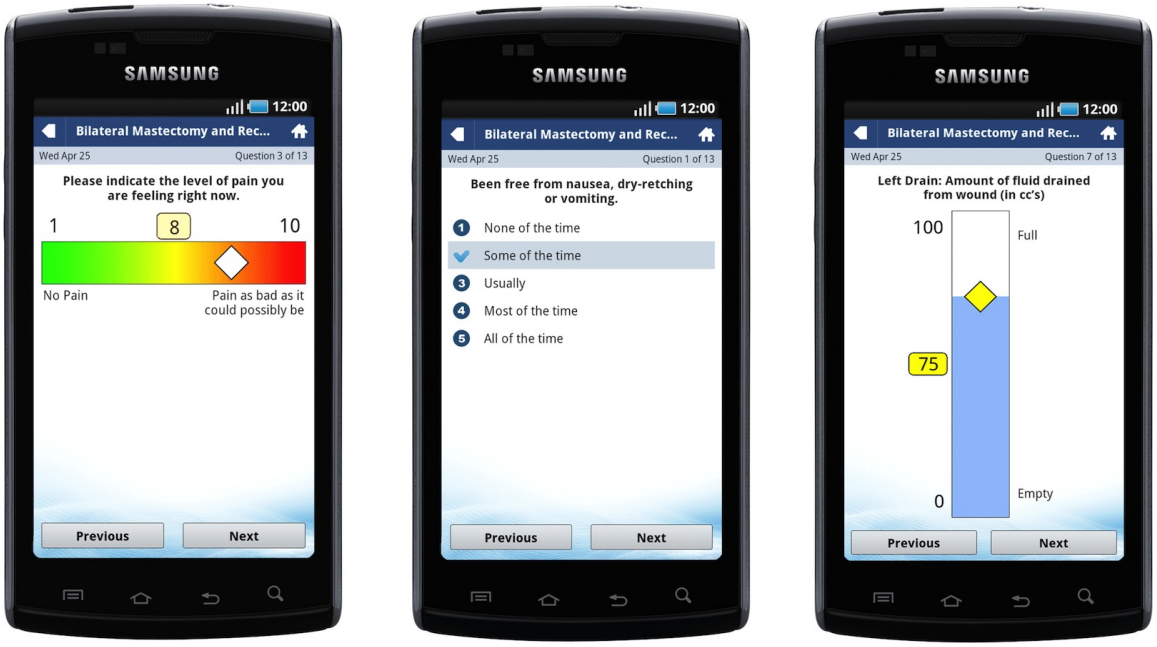
Finding solutions to limit unnecessary burden of care associated with travel is a worthwhile goal in any patient population. However, it is particularly important in patient populations where rurality and lower socioeconomic status are known barriers to breast reconstruction [5]. For these reasons, WCH has completed a feasibility study using a mobile app (QoC Health Inc, Toronto) to support postoperative care in breast reconstruction patients. This mobile app raises the bar by combining validated quality of recovery questionnaires and surgical site photos submitted at the patients' convenience in an asynchronous manner (see Figure 1). The study suggests that mobile app follow-up care adequately detects postoperative complications and eliminates the need for in-person follow-up care. This is concordant with other postoperative telemedicine studies [3,6,7].

Previous studies have found that after a tonsillectomy or adenoidectomy, telephone follow-up care with standardized questionnaires is as safe as standard follow-up care and offers considerable cost reduction and patient convenience [3]. Similar telephone follow-up has also been used successfully in elective open hernia repairs, laparoscopic cholecystectomy, and curative breast cancer surgery [6,7]. Others have shown that planned outpatient appointments after uncomplicated surgery are neither necessary nor cost-effective [8]. A "no planned follow-up" saves money for hospitals and patients [8].

Patients are highly satisfied with telephone follow-up. Still, there are some glaring disadvantages when telephone follow-up care is compared to mobile app follow-up care. Telephone follow-up relies on synchronous communications between patients and health care providers. Studies report between 15 and 27% of patients were unreachable by phone after multiple attempts to reach postoperative patients [6,9]. It is also heavily dependent on labor costs as a nurse or other health care worker is designated to call, collect, and relay the questionnaire information to the primary surgeon. This makes the questionnaire data more expensive to relay when compared to a mobile app, which transmits directly from patient to surgeon. Similarly, "no planned follow-up" is poorly received by patients and providers who value continuity of care [8]. In this way, mobile app follow-up care offers an optimal middle ground between conventional in-person follow-up care, telephone follow-up care, and no planned follow-up care.

The proposed project provides breast reconstruction patients with timely contact with their surgeon from the comfort of their home. This technology has the potential to address wait times by freeing up specialty surgeon clinic time so that they may engage in new consultations and attend to patients in the emergency department. The first step in more widespread implementation involves demonstrating cost-effectiveness.

Figure 1. The mobile app user interface.



Methods

This study used method recommendations from international health technology assessment (HTA) agencies for economic evaluations to develop a model for comparing mobile app follow-up care with in-person follow-up visits [10]. Inputs and outputs were chosen based on relevance to the decision-making perspective of the economic evaluation [10]. Cost data were derived from WCH breast reconstruction patient administrative data and QoC Health Inc mobile app billed costs. This is in keeping with HTA agency recommendations [10]. A societal perspective was adopted wherein all costs were assessed irrespective of the payer [11]. Again, this perspective was chosen based on HTA agency recommendations. This recommendation is meant to improve comparability and consistency across studies [12]. The patient/caregiver, health care system, and externally borne costs are calculated within the first postoperative month. The results are also presented using a narrower health care system perspective that may be of key interest to health administrators and policy decision makers. The effectiveness of mobile app and conventional follow-up care was measured as successful surgical outcomes at 30-days postoperative. Successful surgical outcome was clearly defined as a "surgical patient not requiring medical or surgical intervention related to the original surgery within the first 30-days postoperative". This was deemed an important outcome where no meaningful difference in health-related quality of life (HRQL) between mobile app and in-person follow-up care has been demonstrated [13]. The 30-day time horizon was chosen based on literature surrounding postoperative complications in the first 30-days [14]. It was felt to be long enough to capture all relevant costs and benefits of mobile app and in-person follow-up care. Effectiveness data were derived from clinical studies from similar ambulatory patient populations. Model parameters were input into TreeAge software.

Cost data were collected from a societal perspective using a micro-costing approach advocated for by HTA agencies [10]. All cost data were based on 2013/2014 estimates.

In-person follow-up costs incurred by the health care system include employees, compensation, drugs, surgical instruments and supplies, equipment, and other (eg, linens, telephone charges, general supplies), specialized breast center clinical assistant compensation, resident compensation, and physician fee (see Table 1). WCH provided per patient clinic costs and Ontario Health Insurance Plan (OHIP) billing codes were used to determine physician fees. Physician payment methods were verified through the hospital to ensure double counting did not occur. In keeping with cost-effectiveness analysis, non-recoupable or sunk costs were not included in the in-person follow-up arm [15].

Mobile app follow-up costs incurred by the health care system include the start-up fixed costs such as: health center setup, design/setup of procedure protocols, and training of hospital staff. The start-up costs were divided over the number of patients served over the useful lifespan of smartphone technology, which was conservatively estimated at 5 years. Current e-assessment OHIP physician billing codes are limited. There is currently no OHIP billing code for surgical e-assessments. In the future, we assume that billing codes will exist and so we applied a fee based on actual OHIP telemedicine follow-up fees. The variable costs for mobile app follow-up care included software, licensing, and technical support. The bring-your-own-device (BYOD) model variable cost was $3.50 CAD per patient per day. This costing assumes that 1000 patients are enrolled in BYOD mobile

app follow-up per year based on a QoC Health Inc business model that enrolls hospitals.

In-person follow-up costs incurred by the patient included foregone patient leisure time, the wage of a caregiver, and travel and parking costs associated with follow-up visits. We determined foregone leisure costs based on labor force participation rates and age-sex adjusted average Ontario wages. Labor force non-participants were assigned an Ontario homemaker's wage [16]. We presumed that a caregiver equivalent would be present at the first follow-up visit, and assigned a homemaker wage ($11.28 per hour) to that person [16]. This is considered a conservative estimate of true caregiver costs because most patients bring their partner with them to clinic, and those individuals would earn higher average age/sex adjusted wages. The hourly rates were multiplied by the travel time and length of the clinic visit. The clinic time was assumed to be 1 hour to include time to park, register, and meet with the health care team. Travel time and costs estimates were based on actual breast reconstruction patient distance data from home postal code to WCH. Canadian Automobile Association (CAA) Ontario-based average costs per km driven were used to calculate transportation costs. The number of clinic visits was averaged at 1.64 visits over the first postoperative month based on actual attendance by breast reconstruction patients at WCH.

Mobile app follow-up costs incurred by the patient were modeled based on a BYOD format, in which the patient loads the app on to their own mobile phone. Costs included the foregone leisure time to submit follow-up data and the cost of data submission. Each submission takes approximately 3 minutes to enter and submit. In the feasibility study, patients were asked to submit monitoring information once daily for the first 2 weeks and then once weekly for the next 2 weeks. Leisure time was not interrupted by the submission of a mobile app follow-up; therefore, there was minimal sacrifice. Each submission (including survey information and photo) used approximately 0.35 MB of data. In Ontario, 2 GB of data can be purchased for $45 CAD [17]; therefore, data costs were negligible. Patient training sessions were held while patients waited for their preoperative appointment. There were no additional patient costs associated with this time.

This modeling study used telephone follow-up studies to determine the effectiveness of mobile app follow-up when compared to in-person follow-up care. Telephone and mobile app follow-up are considered to transmit the same questionnaire-based data from patient to provider. HTA agencies recommend conducting a systematic review of the literature on key model inputs including effect data; however, clinical trials and observational studies can be used to obtain effect data if they more appropriately represent the model of interest [10]. A recent article in BMC Health Services Research systematically reviewed telephone consultations in place of face-to-face outpatient consultation for patients discharged from hospital following surgery. It reported low methodological quality and dissimilar outcomes [18]. None of the articles included in the review captured patient populations or outcomes that were comparable to the patient populations and outcomes modeled in this study. The lack of comparative data reflects the fact that type of follow-up care (mobile app, telephone, or in-person) does not impact the chance complication. For this reason, baseline equivalence in effect was modeled between the two groups. This assumption is supported by large observational studies following laparoscopic cholecystectomy, inguinal and paraumbilical hernia repair, other hernia repair, varicose vein surgery, circumcision, excision of subcutaneous lesions, carpal tunnel release, and appendectomies [6,9]. These studies found that structured postoperative telephone questionnaires conducted between 2 and 6 weeks were a safe alternative to in-person follow-up care [6,9]. Telephone questionnaire-based follow-up adequately detected patients that required further in-person assessment (5-11% of all patients) [6,9]. These studies contain similar patient populations, procedural variation, and surgical risks when compared to ambulatory breast reconstruction patients.

Three types of sensitivity analyses were performed to determine their effects on costs and outcomes. A scenario analysis was conducted for variations in the number of in-person clinic visits and crossover from mobile app follow-up to in-person follow-up. A two-way sensitivity analysis varied patient wage and mobile app follow-up effect. Additionally, a probabilistic sensitivity analysis was performed to account for uncertainty in the distribution of patient, caregiver, and clinic costs as well as uncertainty in effects (ie, complication rates).

Results

Overview

The results of this analysis are summarized in Table 1. The total cost difference between mobile app and in-person follow-up care was $245 CAD ($223 USD based on the current exchange rate), with in-person follow-up being more expensive ($381 CAD) than mobile app follow-up care ($136 CAD). This takes into account the total of health care system, patient, and external borne costs. If we examine health care system costs alone, in-person follow-up was $38 more expensive than mobile app follow-up care (please see Table 1). The baseline difference in effect was modeled to be zero based on the WCH feasibility study, as well as other ambulatory telephone follow-up studies. An incremental cost-effectiveness ratio (ICER) is not reportable in this scenario. An incremental net benefit (INB) is reportable, and reflects merely the cost difference between the two interventions for any willingness-to-pay value (INB=$245 CAD).

Table 1. Cost breakdown.

In-person follow-up	Cost (CAD $)	Mobile app follow-up	Cost (CAD $)
Health care system costs			
Fixed costs			
Compensation	103.74	Health center setup	1.39
Equipment	2.16	Design/setup procedure protocol	6.94
		Training	0.44
Variable costs			
Drugs	0.21	Platform licensing, accounts	42.00
Other (Linens)	3.83	Standard support	43.05
Clinical assistant (10 min)	10.25	Infrastructure hosting	19.95
Surgeon fee	43.46	Surgeon fee	22.00
Resident	10.56		
Health care system costs subtotal (per patient per 1.64 visits over 30 days)	$174	Health care system costs subtotal (per patient per 30 days monitoring)	$136
Patient costs			
Variable costs			
Patient leisure time	102.24	Patient leisure time	negligible
Caregiver wage	33.84	Data (approx. 350 kB per transmission with photo)	negligible
Travel (to and from clinic)	38.11		
Parking	32.80		
Patient costs subtotal (per patient per 1.64 visits over 30 days)	$207	Patient costs subtotal (per patient per 30-day monitoring)	negligible
Total societal costs[a]			
Per patient per 1.64 visits over 30 days	$381	Per patient per 30-day monitoring period	$136

[a]Total societal costs = health care system costs subtotal + patient costs subtotal

Scenario Analysis: Societal Perspective Costs With Varying Number of In-Person Visits in the First Month Postoperative

The number of in-person follow-up visits was set to a minimum value and compared to the costs of mobile app follow-up. This sensitivity analysis demonstrates that even at only 1 in-person visit per patient over the first month postoperative, mobile app follow-up care is less costly from a societal perspective. From a societal perspective, mobile app follow-up care remains cost equivalent to in-person follow-up even when 100 percent of the mobile app follow-up care patients attend 1 in-person visit during the first month.

Two-Way Sensitivity Analysis: Societal Perspective Costs With Varying Foregone Patient Leisure Time and Mobile Effectiveness

The patient's wage was set between $11.28 (homemaker) and $26.71 (age/sex adjusted) per hour wage. The mobile effect was varied between a 90-96% success rate. Table 2 demonstrates how an incremental net benefit only favors (ie, produces a negative value) in-person follow-up if a 6 percentage point difference in effect exists between the two follow-up groups and the patient makes <$19 CAD per hour. This calculation uses a willingness-to-pay (WTP) of $100,000 USD ($109,970 CAD based on the current exchange rate) per quality adjusted life year (QALY), and a 0.04 QALY difference between no complication and minor skin infection. This is a high estimate previously reported in the literature [19].

Table 2. Two-way sensitivity analysis with varying patient lost leisure time and effectiveness of mobile app follow-up care.[a]

		Patient lost leisure time (CAD $)		
Mobile Effect		$33.84	$56.98	$80.12
INB[b] @	Effect 0.96	198.75	236.70	274.65
INB @	Effect 0.94	110.77	148.72	186.67
INB @	Effect 0.92	22.80	60.75	98.70
INB @	Effect 0.90	−65.18	−27.23	10.72

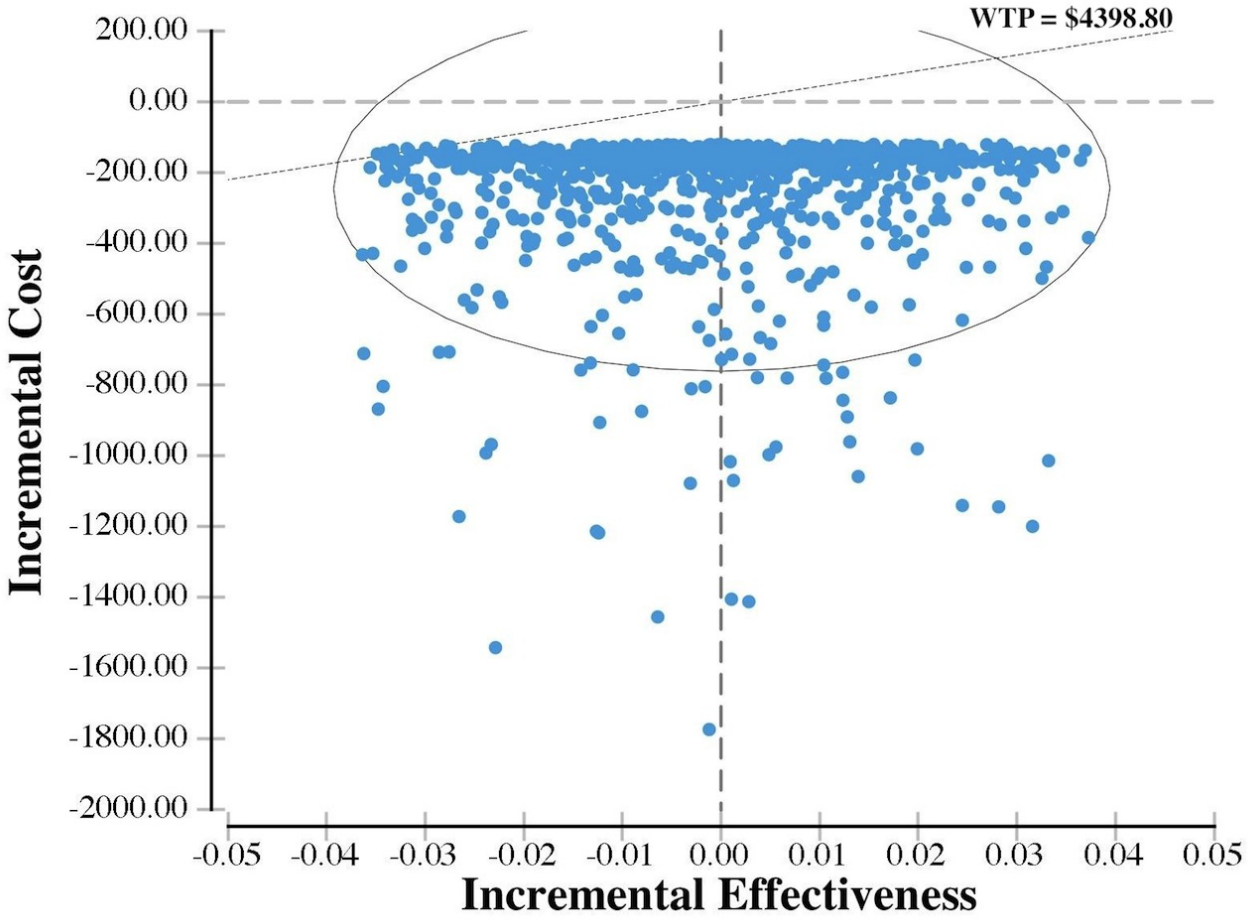
[a]Willingness-to-pay (WTP)=$4398.80 CAD per effect based on $109,970 CAD per quality adjusted life year (QALY) and 0.04 QALY assigned to one superficial skin infection [19].

[b]INB: incremental net benefit

Probabilistic Sensitivity Analysis

To further explore the robustness of the base case results, a probabilistic sensitivity analysis was performed, based on random re-sampling (Monte Carlo simulation). In the analysis, a uniform distribution was allocated to the clinic cost (by +/− 20%), mobile and in-person follow-up effect (+/− 2 percentage points). A gamma distribution was applied to the patient wage. In 10,000 simulations, the mean societal cost of mobile app follow-up care was $135.78 CAD and in-person follow-up care was $383.55 CAD per patient. The large in-person follow-up care standard deviation of $211.80 CAD accurately reflects the variation in wage and travel time among patients. In all scenarios, mobile app follow-up care was cheaper than in-person follow-up care from a societal perspective. In approximately 50% of scenarios, the effectiveness of mobile app follow-up was less than in-person follow-up (see Figure 2). This was imposed by the distributions assigned. It is important to note, scenarios that are less effective and less costly can still be considered cost-effective. Again, using a WTP of $4398.80 CAD/superficial skin infection, mobile app follow-up care is the preferred strategy in 99.1% of scenarios.

Figure 2. Graphical representation of probabilistic sensitivity analysis demonstrating the Incremental Cost-Effectiveness Ratio (ICER) values for mobile app versus in-person follow-up care.



Discussion

Principal Findings

Results from modeling cost-effectiveness show that mobile app follow-up care is cost-effective when compared to in-person follow-up care from a societal and health care system perspective. A detailed examination demonstrated that in the first month of follow-up care, an average of $136 CAD was spent in the mobile app follow-up care stream whereas $381 CAD was spent in the in-person follow-up care stream. The higher cost of in-person follow-up care is spread between the health care system and patient; however, the patient reaps the majority of the cost-savings from participating in mobile app follow-up care. This is demonstrated by comparing the societal and health care system perspective savings ($245 vs $38, respectively), as the patient savings are only captured in the societal perspective. This is an important finding as lower socioeconomic status is a known barrier to breast reconstruction [5]. The two-way sensitivity analysis demonstrates how these savings are maintained even when the patient makes a homemaker wage. Decreasing costs incurred to the patient, at least in the postoperative period, may improve access.

There is a deficit in Canadian policy promoting mobile phone communication between patients and providers. Ontarians cannot even renew a prescription over the phone, unless they choose to pay out-of-pocket for a normally insured service, because there is no telemedicine prescription renewal code [20]. This cost-effectiveness study is an important first step in demonstrating to health care administrators and policy decision-makers the benefits of investing in mobile app follow-up care. Mobile app follow-up care generates an incremental net benefit of $38 per patient from the perspective of the health care system. Decreasing the total number of in-person follow-up visits required has the potential to generate efficiency in one of two ways. Hospitals could choose to investment in smaller clinic spaces, decreasing the fixed and variable costs that accompany these spaces. Alternatively, hospitals could serve more patients in a given clinic space, including more new consultations. This is an important finding given the concern with long specialty wait times across Canada [21]. Orthopedic surgery and plastic surgery have the longest wait times. These two specialties perform a significant number of ambulatory surgeries and their patients in particular could benefit from mobile app follow-up care. Moreover, the number of patients that would benefit from mobile app follow-up care is growing yearly. At Women's College Hospital, over 5000 elective ambulatory surgeries are performed each year. These numbers are small when you look at other neighboring hospitals, where over 20,000 ambulatory surgeries are performed annually. These numbers will continue to grow as we follow trends in the United States where currently 60 to 70% of the surgical procedures are performed in the ambulatory setting [22].

Mobile app follow-up transmits the same information as telephone follow-up care, but its obvious advantages include its asynchronous nature and autonomy from health care labor force to call, collect, and relay the patient data. The ease of use allows data to be collected multiple times during the 30-day follow-up period. In our pilot study, patients submitted questionnaire and surgical site photos every day for the first 2 weeks and once a week for the following 2 weeks. This provides richer data than could ever be achieved by telephone or in-person follow-up care. At this point in time, mobile app follow-up makes sense. Usage is ubiquitous throughout North America. Mobile phone penetration is approaching 90% in the United States, and smartphones are now considered the dominant mobile device [23]. As technology is an economy of scale, the potential for cost-savings increases with user uptake.

Limitations

There are a few limitations to this study. Equivalency in the effectiveness of mobile app and in-person follow-up care is assumed based on observational studies of telephone questionnaire-based follow-up care from similar ambulatory surgery patient populations. There are no randomized control trials demonstrating equal effectiveness between mobile app and in-person follow-up care. From a clinical perspective, effect equivalence is intuitive because outcomes are dependent on patient and surgical factors (face validity).

This study did not compare the cost-effectiveness of telephone follow-up care to mobile app and in-person follow-up care. This is because most HTA agencies recommend comparing technology to usual care [10]. Mobile app and telephone follow-up care utilize the same standardized questionnaire tool; however, telephone follow-up care has obvious disadvantages including (1) the reliance on synchronous communication between the patient and health care professional, (2) no capacity to submit surgical site photography, and (3) a heavy dependence on human resources leading to higher costs.

Conclusions

Mobile app follow-up care is suitably targeted to low-risk postoperative ambulatory patients. It can be cost-effective from a societal and health care system perspective. Mobile phone penetration is approaching 90% in the United States, and smartphones are now considered the dominant mobile device. Using a ubiquitous technological platform to reduce health care costs for patients and providers in an already large and growing patient population makes sense.

Acknowledgments

This project was funded by the Canadian Institutes of Health Research e-Health Catalyst Grant.

Conflicts of Interest

John Semple is a shareholder in the company that produces the mobile app, QoC Health Inc.

References

1. Ministry of Health and Long Term Care.: Government of Ontario; 2012. Healthy Change: Ontario's Action Plan for Health Care URL: http://www.health.gov.on.ca/en/ms/ecfa/healthy_change/ [accessed 2014-03-18] [WebCite Cache ID 6OAShWp9R]
2. Canizares M, Badley E, Davis A, Mackay C, Mahomed N. Orthopaedic surgery in Ontario in the era of the wait time strategy. Toronto: Arthritis Community Research & Evaluation Unit, University Health Network; 2007. URL: http://www.acreu.ca/pdf/pub5/07-02.pdf [accessed 2014-08-14] [WebCite Cache ID 6RpPCZN5l]
3. Jones DT, Yoon MJ, Licameli G. Effectiveness of postoperative follow-up telephone interviews for patients who underwent adenotonsillectomy: a retrospective study. Arch Otolaryngol Head Neck Surg 2007 Nov;133(11):1091-1095. [doi: 10.1001/archotol.133.11.1091] [Medline: 18025311]
4. Davidge KM, Brown M, Morgan P, Semple JL. Processes of care in autogenous breast reconstruction with pedicled TRAM flaps: expediting postoperative discharge in an ambulatory setting. Plast Reconstr Surg 2013 Sep;132(3):339e-344e. [doi: 10.1097/PRS.0b013e31829ace62] [Medline: 23985645]
5. Platt J, Baxter N, Zhong T. Breast reconstruction after mastectomy for breast cancer. CMAJ 2011 Dec 13;183(18):2109-2116. [doi: 10.1503/cmaj.110513]
6. Hwa K, Wren SM. Telehealth follow-up in lieu of postoperative clinic visit for ambulatory surgery: results of a pilot program. JAMA Surg 2013 Sep;148(9):823-827. [doi: 10.1001/jamasurg.2013.2672] [Medline: 23842982]
7. Kimman ML, Dirksen CD, Voogd AC, Falger P, Gijsen BC, Thuring M, et al. Economic evaluation of four follow-up strategies after curative treatment for breast cancer: results of an RCT. Eur J Cancer 2011 May;47(8):1175-1185. [doi: 10.1016/j.ejca.2010.12.017] [Medline: 21257305]
8. Bailey J, Roland M, Roberts C. Is follow up by specialists routinely needed after elective surgery? A controlled trial. J Epidemiol Community Health 1999 Feb;53(2):118-124 [FREE Full text] [Medline: 10396473]
9. Gray RT, Sut MK, Badger SA, Harvey CF. Post-operative telephone review is cost-effective and acceptable to patients. Ulster Med J 2010 May;79(2):76-79 [FREE Full text] [Medline: 21116423]
10. Mathes T, Jacobs E, Morfeld JC, Pieper D. Methods of international health technology assessment agencies for economic evaluations--a comparative analysis. BMC Health Serv Res 2013;13:371 [FREE Full text] [doi: 10.1186/1472-6963-13-371] [Medline: 24079858]
11. Neumann PJ. Costing and perspective in published cost-effectiveness analysis. Med Care 2009 Jul;47(7 Suppl 1):S28-S32. [doi: 10.1097/MLR.0b013e31819bc09d] [Medline: 19536023]
12. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. JAMA 1996;276(16):1339-1341. [Medline: 8861994]
13. Guidelines for the economic evaluation of health technologies. Canada: Canadian Agency for Drugs and Technologies in Health; 2006. URL: http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf [accessed 2014-08-05] [WebCite Cache ID 6RaZykjgS]
14. Kazaure HS, Roman SA, Sosa JA. Association of postdischarge complications with reoperation and mortality in general surgery. Arch Surg 2012 Nov;147(11):1000-1007. [doi: 10.1001/2013.jamasurg.114] [Medline: 23165614]
15. Robinson R. Costs and cost-minimisation analysis. BMJ 1993 Sep 18;307(6906):726-728 [FREE Full text] [Medline: 8401098]
16. Explore Careers by Occupation - Job Bank.: Government of Canada; 2013. URL: http://www.jobbank.gc.ca/occupation_search-eng.do [accessed 2014-05-05] [WebCite Cache ID 6OATFQvyD]
17. Price comparison of Wireline, Wireless, and Internet Services in Canada and with Foreign Jurisdictions.: Wall Communications Inc; 2013. URL: http://www.wallcom.ca/pdfs/price-comp-report_2013update.pdf [accessed 2014-09-10] [WebCite Cache ID 6SUuNqPiu]
18. Thompson-Coon J, Abdul-Rahman A, Whear R, Bethel A, Vaidya B, Gericke C, et al. Telephone consultations in place of face to face out-patient consultations for patients discharged from hospital following surgery: a systematic review. BMC Health Serv Res 2013;13:128 [FREE Full text] [doi: 10.1186/1472-6963-13-128] [Medline: 23561005]
19. Grover R, Padula WV, Van Vliet M, Ridgway EB. Comparing five alternative methods of breast reconstruction surgery: a cost-effectiveness analysis. Plast Reconstr Surg 2013 Nov;132(5):709e-723e. [doi: 10.1097/PRS.0b013e3182a48b10] [Medline: 24165623]
20. Can my doctor charge me for refilling a prescription?. Toronto: Globe and Mail; 2011. URL: http://www.theglobeandmail.com/life/health-and-fitness/ask-a-health-expert/can-my-doctor-charge-me-for-refilling-a-prescription/article579941/ [accessed 2014-08-05] [WebCite Cache ID 6RaaqDpYo]
21. Barua B, Esmail N. Waiting your turn: Wait times for health care in Canada.: Fraser Institute; 2013. URL: http://www.fraserinstitute.org/uploadedFiles/fraser-ca/Content/research-news/research/publications/waiting-your-turn-2013.pdfs/waiting-your-turn-2013.pdf [accessed 2014-08-05] [WebCite Cache ID 6RabKXw8e]
22. Cullen K, Hall M, Golosinskiy A. Ambulatory surgery in the United States, 2006. Natl Health Stat Report 2009 Jan 28(11):1-25 [FREE Full text] [Medline: 19294964]
23. Page M, Molina M, Jones G. The Mobile Economy. London: A.T. Kearney; 2013. URL: http://www.atkearney.com/documents/10192/760890/The_Mobile_Economy_2013.pdf [accessed 2014-03-18] [WebCite Cache ID 6OATgN6cl]

Abbreviations

BYOD: bring-your-own-device
HTA: health technology assessment
ICER: incremental cost-effectiveness ratio
INB: incremental net benefit
QALY: quality adjusted life year
WCH: Women's College Hospital
WTP: willingness-to-pay

Edited by G Eysenbach; submitted 11.05.14; peer-reviewed by K Ho, M Karunanithi; comments to author 23.07.14; revised version received 05.08.14; accepted 23.08.14; published 22.09.14.

Please cite as:
Armstrong KA, Semple JL, Coyte PC
Replacing Ambulatory Surgical Follow-Up Visits With Mobile App Home Monitoring: Modeling Cost-Effective Scenarios
J Med Internet Res 2014;16(9):e213
URL: http://www.jmir.org/2014/9/e213/
doi:10.2196/jmir.3528
PMID:25245774

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org