Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 13 (2011), Issue 2   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Original Paper

Effectiveness of a Web-based Intervention Aimed at Healthy Dietary and Physical Activity Behavior: A Randomized Controlled Trial About Users and Usage

Saskia M Kelders[1,2], MSc; Julia E.W.C Van Gemert-Pijnen[1], PhD; Andrea Werkman[3], PhD; Nicol Nijland[1], PhD; Erwin R Seydel[1], PhD

[1]Department of Psychology, Health and Technology, Faculty of Behavioural Sciences, University of Twente, Enschede, Netherlands

[2]National Institute for Public Health and the Environment, Bilthoven, Netherlands

[3]Netherlands Nutrition Centre, Den Haag, Netherlands

Corresponding Author:
Saskia M Kelders, MSc
Department of Psychology, Health and Technology
Faculty of Behavioural Sciences
University of Twente
CI H403
PO Box 217
Enschede, 7500 AE
Netherlands
Phone: 31 6 51070689
Fax: 31 53 489 2388
Email: s.m.kelders@utwente.nl

Abstract

Background: Recent studies have shown the potential of Web-based interventions for changing dietary and physical activity (PA) behavior. However, the pathways of these changes are not clear. In addition, nonusage poses a threat to these interventions. Little is known of characteristics of participants that predict usage.

Objective: In this study we investigated the users and effect of the Healthy Weight Assistant (HWA), a Web-based intervention aimed at healthy dietary and PA behavior. We investigated the value of a proposed framework (including social and economic factors, condition-related factors, patient-related factors, reasons for use, and satisfaction) to predict which participants are users and which participants are nonusers. Additionally, we investigated the effectiveness of the HWA on the primary outcomes, self-reported dietary and physical activity behavior.

Methods: Our design was a two-armed randomized controlled trial that compared the HWA with a waiting list control condition. A total of 150 participants were allocated to the waiting list group, and 147 participants were allocated to the intervention group. Online questionnaires were filled out before the intervention period started and after the intervention period of 12 weeks. After the intervention period, respondents in the waiting list group could use the intervention. Objective usage data was obtained from the application itself.

Results: In the intervention group, 64% (81/147) of respondents used the HWA at least once and were categorized as “users.” Of these, 49% (40/81) used the application only once. Increased age and not having a chronic condition increased the odds of having used the HWA (age: beta = 0.04, $P = .02$; chronic condition: beta = 2.24, $P = .003$). Within the intervention group, users scored better on dietary behavior and on knowledge about healthy behavior than nonusers (self-reported diet: $\chi^2_2 = 8.4$, $P = .02$; knowledge: $F_{1,125} = 4.194$, $P = .04$). Furthermore, users underestimated their behavior more often than nonusers, and nonusers overestimated their behavior more often than users (insight into dietary behavior: $\chi^2_2 = 8.2$, $P = .02$). Intention-to-treat analyses showed no meaningful significant effects of the intervention. Exploratory analyses of differences between pretest and posttest scores of users, nonusers, and the control group showed that on dietary behavior only the nonusers significantly improved (effect size $r = -.23$, $P = .03$), while on physical activity behavior only the users significantly improved (effect size $r = -.17$, $P = .03$).

Conclusions: Respondents did not use the application as intended. From the proposed framework, a social and economic factor (age) and a condition-related factor (chronic condition) predicted usage. Moreover, users were healthier and more knowledgeable about healthy behavior than nonusers. We found no apparent effects of the intervention, although exploratory analyses showed that choosing to use or not to use the intervention led to different outcomes. Combined with the differences between groups at baseline, this seems to imply that these groups are truly different and should be treated as separate entities.

Trial registration: Trial ID number: ISRCTN42687923; http://www.controlled-trials.com/ISRCTN42687923/ (Archived by WebCite at http://www.webcitation.org/5xnGmvQ9Y)

(J Med Internet Res 2011;13(2):e32) doi:10.2196/jmir.1624

KEYWORDS

Randomized controlled trial; usage; eHealth; intervention; attrition; Internet; adherence; retention

Introduction

The increasing prevalence of overweight is a problem in modern society. It is closely related to a number of chronic conditions such as type 2 diabetes mellitus and places a great burden on the health care system. Losing weight and especially preventing weight regain is challenging. It might be more cost-efficient to prevent people from becoming overweight by focusing on healthy dietary and physical activity (PA) behavior [1-3]. To achieve this goal, interventions aimed at the general public are needed that must not only inform people about the risks of unhealthy dietary and physical activity habits but must also stimulate people to adopt healthier behaviors related to diet and physical activity [2,4]. Previous research has shown that tailored and interactive interventions can achieve this goal [2,4-7]. The Internet provides an opportunity for these interventions to reach a broad population. Besides, by using a Web-based application, the content of the intervention can be tailored to the users, and the intensity can be varied according to the needs and wishes of these users [8-9]. Research has already shown the potential of these applications for the achievement of weight loss and weight management [6,10-14]. However, most studies are focused on applications aimed at treatment or secondary prevention. Many questions remain about the users and the effectiveness of Web-based applications for the prevention of health problems by stimulating healthy behaviors.

The problem of attrition [15] poses a threat to most eHealth interventions but might pose an even bigger threat to Web-based interventions for prevention, considering that people who do not experience an urgent health problem might be less internally motivated to change their behavior [16]. Until recently, the characteristics of the users and nonusers of Web-based applications have gained only very limited attention [17-19]. It is important to know who the users of these interventions are. This knowledge helps us identify important factors in the dissemination of these interventions and the characteristics of intended users who are not reached [20]. Moreover, recent studies indicate that people react differently to motivational and persuasive strategies, which might make the need for examining user characteristics even more essential [21]. A recent review by Christensen and colleagues [22] emphasized the need for a theoretical framework to increase our understanding of attrition. They proposed using the framework adopted by the World Health Organization (WHO) [16] (ie, five dimensions of adherence: health system factors, social and economical factors, therapy-related factors, condition-related factors, and patient-related factors) and mention the possible potential of behavior theories. Furthermore, research into the reasons for use of Web-based eHealth applications can give us valuable information on what the users hope to accomplish and how the application can assist them. In addition, usability and satisfaction with an application can play an important role in the extent to which such applications are ultimately used [15,23].

We incorporated the WHO framework and behavior theories in a study of use of the Healthy Weight Assistant (HWA), a Web-based lifestyle intervention. We considered the influence of social and economic factors (demographics), condition-related factors (ie, general practitioner [GP] visits, having a chronic condition, and self-reported and self-rated dietary and PA behavior), patient-related factors or constructs identified by behavior change theories (ie, knowledge, attitude, and self-efficacy) [24-25], and reasons for use and satisfaction with the intervention.

Additionally in this study, we assessed the effectiveness of the intervention using self-reported dietary and PA behavior as primary outcome measures because the intervention was aimed at improving health behavior. We included secondary outcome measures that are known determinants of behavior change. We also chose to include measures of knowledge, attitude, and self-efficacy [24-25]. Self-rated behavior and insight into behavior were included as secondary outcome measures because one of the goals of the intervention was to improve insight into one’s own behavior.

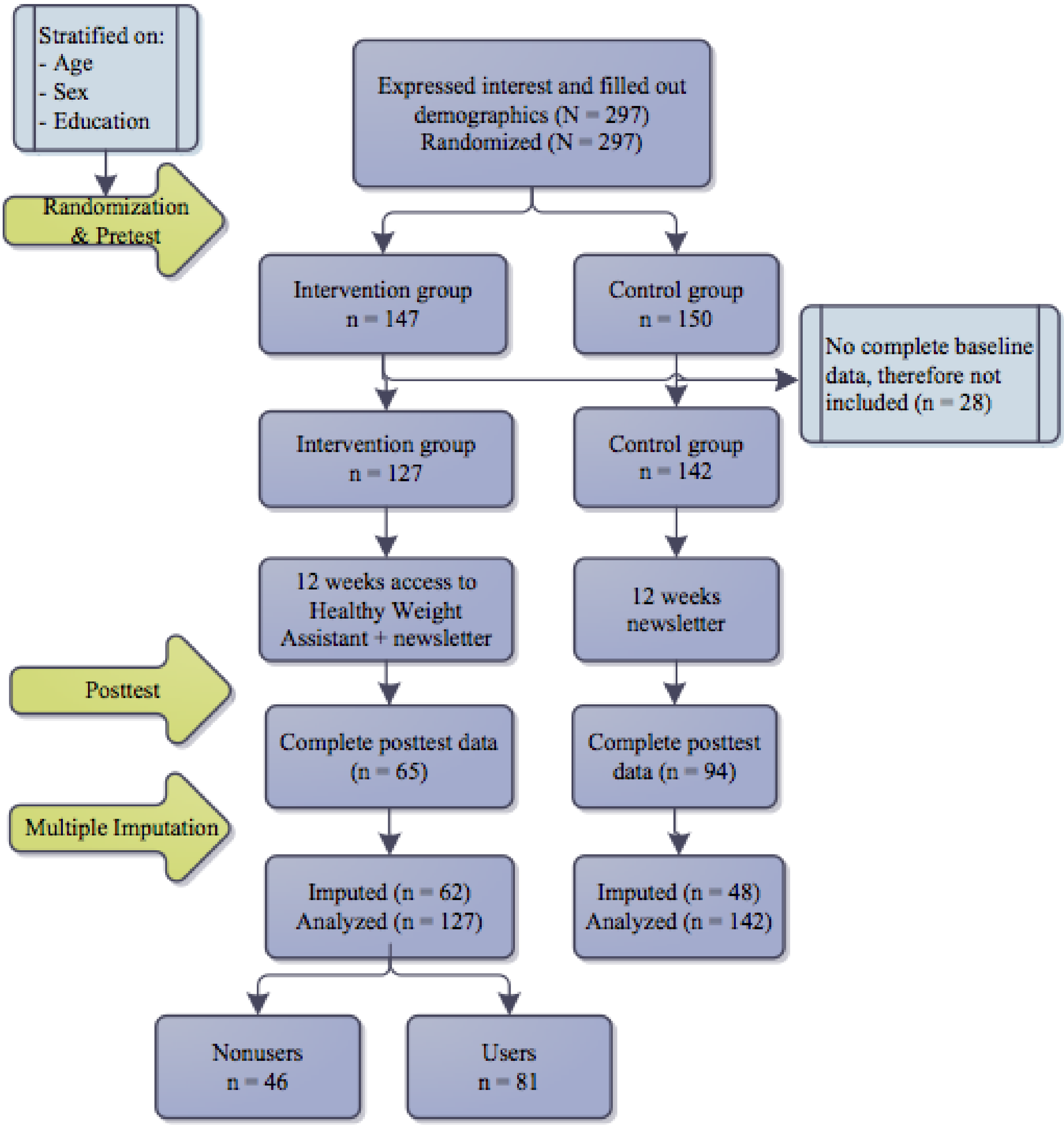
Consequently, our research questions were: What characteristics of participants are related to the use of the HWA intervention? What effects does the HWA intervention have on the primary and secondary outcome measures?

Methods

Recruitment and Design

Participants were recruited through advertisements about an online lifestyle intervention in local newspapers, supermarkets, and on health-related websites. Permission of an ethics review board for the study was not required because, according to the Dutch law, nonintrusive interventions conducted with healthy adults do not require approval from an ethics board. In total, 297 respondents expressed interest in using an online lifestyle intervention and satisfied our inclusion criteria (body mass index [BMI] 18.5 - 28.0 kg/m^2, Dutch-speaking). The inclusion criterion for BMI was chosen to reflect the target group of the intervention under investigation. The sample used in this study was a self-selected convenience sample. Enrollment took place beginning November 1, 2008, and ending December 31, 2008. All participants were randomly assigned to either the Web-based lifestyle coach or a waiting list. A total of 150 participants were allocated to the waiting list group, and 147 participants were allocated to the intervention group. Participants filled out online questionnaires before the 12-week intervention period started and again after the intervention period ended. The posttest questionnaire was available for all respondents for a period of 3 weeks beginning February 27 and ending April 16. After the intervention period, respondents in the waiting list group could use the intervention. The flowchart of the study can be found in Figure 1.

Figure 1. Flowchart of the study



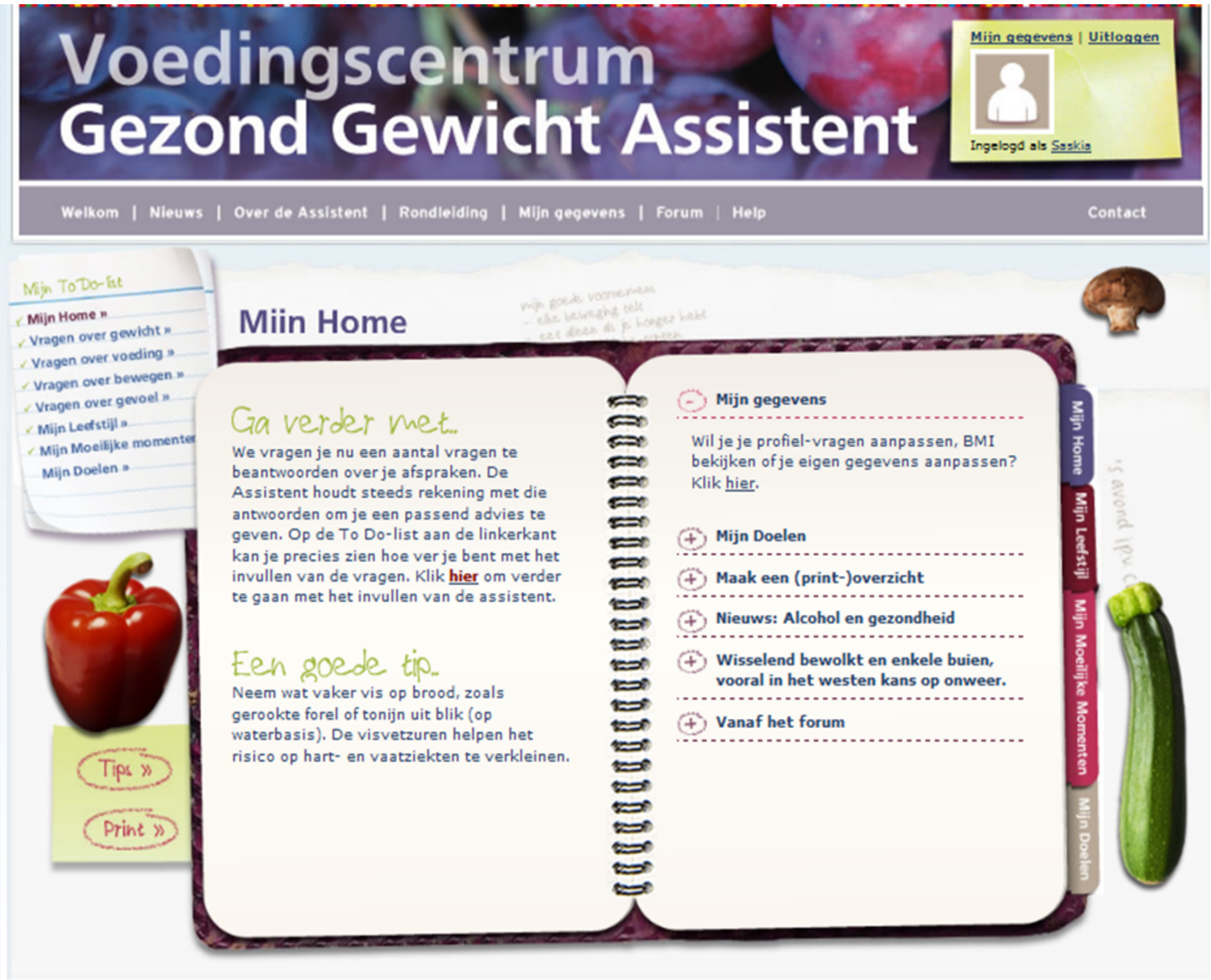
Randomization

Randomization took place 1 week before the start of the intervention period. We used block randomization with blocks of 4 participants, stratified on age, sex, and education. The randomization scheme was created by a computer application and carried out by a member of the research team. Participants who filled out demographic information were randomized. Only respondents who completed the pretest questionnaire were included; therefore, 28 respondents were excluded. Participants were not blinded to randomization outcome but received an email with information on when and how they were able to access the Healthy Weight Assistant (HWA) after filling out the pretest questionnaire.

Intervention

The Healthy Weight Assistant (HWA) is a Web-based lifestyle intervention developed by the Netherlands Nutrition Centre, which is a government-funded organization focusing on increasing the knowledge of consumers about the quality of food and encouraging consumers to eat healthily and safely. The goal of the HWA is to support people with a healthy weight and people who are slightly overweight (ie, BMI 18.5-28.0 kg/m^2) to maintain and achieve a healthy weight. The aim is not to achieve a given weight loss, but to support the achievement of healthy dietary and PA behavior. Therefore, the focus was broader than only energy balance-related behavior. The target group was selected by the Netherlands Nutrition Centre according to their BMI classification. The theoretical basis for behavior change via the HWA is the transtheoretical model [26], which entails that the participants are addressed according to the stage of change in which they find themselves when starting the application. The researchers were not the leading party in the design of the HWA but have done earlier research on the application. This previous study employed user-centered evaluation methods and has led to slight alterations in the design of the application in order to increase users' motivation to keep using the HWA and their motivation to change behavior [27].

Figure 2. The Healthy Weight Assistant



Ga verder met..

We vragen je nu een aantal vragen te beantwoorden over je afspraken. De Assistent houdt steeds rekening met die antwoorden om je een passend advies te geven. Op de To Do-list aan de linkerkant kan je precies zien hoe ver je bent met het invullen van de vragen. Klik hier om verder te gaan met het invullen van de assistent.

Een goede tip..

Neem wat vaker vis op brood, zoals gerookte forel of tonijn uit blik (op waterbasis). De visvetzuren helpen het risico op hart- en vaatziekten te verkleinen.

⊖ Mijn gegevens

Wil je je profiel-vragen aanpassen, BMI bekijken of je eigen gegevens aanpassen? Klik hier.

⊕ Mijn Doelen

⊕ Maak een (print-)overzicht

⊕ Nieuws: Alcohol en gezondheid

⊕ Wisselend bewolkt en enkele buien, vooral in het westen kans op onweer.

⊕ Vanaf het forum

The HWA consists of 4 steps, which are marked in the application by a "to-do list" and tabs in the "diary" (Figure 2). When users enter the program for the first time, they start by assessing their baseline status. In this step, users answer questions about their body weight, dietary behavior, physical activity behavior, and emotions concerning these behaviors. This results in tailored advice that can be applied in the next steps of the application. The second step is motivation. Users are asked about their motivation to change behavior, and the application assists them in making these motivations clear to themselves, thereby also focusing on clarifying their emotions related to behavior. The third step is called *difficult moments*. Users are encouraged to reflect on their difficult moments (ie, moments at which it is tempting to engage in unhealthy behavior) and to provide solutions for these moments. The HWA coaches the user throughout this step by giving automated tailored feedback based on input of the users. The final step is goal setting and monitoring achievement of goals. Users are coached to set useful and realistic goals and can opt to receive a weekly email reminder on these goals. Additionally, users can give feedback on the achievement of their own goals and access an overview of previous goals. The news section of the HWA is regularly updated, and when users exit the application, random hints are displayed. Other content is static. The HWA is designed to be used at regular intervals. The intended use is one or multiple visits within a short period of time to complete the first 3 steps. For the last step, the intended use is once a week to once a fortnight over a longer period of time. For the research period, the HWA was only available to the participants. After this period, the application was made openly accessible through a website.

Waiting List

We made use of a waiting list control group. Participants randomized in this group received an email newsletter every 3 weeks, but no access to the HWA during the intervention period. The newsletter contained general information about the study and about the University of Twente. Furthermore, it contained leisure tips, but it contained no information on healthy lifestyle. After the intervention period, participants in the waiting list group received access to the HWA. Participants in the intervention group also received the newsletter every 3 weeks.

Research Instruments

Online questionnaires were used to assess pretest and posttest values. Education was self-reported and recoded into the following three categories: low (primary and lower vocational education), moderate (secondary and middle vocational education), and high (higher vocational and university education). BMI (kg/m^2) was calculated using self-reported weight and length. Dietary behavior was measured using a 14-item self-report questionnaire of the Netherlands Nutrition Centre, based on the Netherlands classification model [28]. This questionnaire has not been validated but was used because of the applicability to the standards used by the Netherlands Nutrition Centre [29]. These standards are based on a report of the Health Council of the Netherlands, which is the basis of nutritional education in the Netherlands [30]. This questionnaire classifies respondents as *unhealthy* (not complying to the standards on all aspects), *improvable* (complying with the standards on some aspects), and *healthy* (complying with the standards on all aspects). This classification entails that respondents in the healthy category have limited room for improvement because they already comply with all of the standards. We have included a translation of this questionnaire in Multimedia Appendix 1. Physical activity behavior was measured according to the Dutch Standard for Healthy Physical Activity, using a validated 4-item self-report questionnaire [31]. This questionnaire classifies respondents into two categories, *unhealthy* (not complying with the standards) and *healthy* (complying with the standards). Again, this classification entails that respondents in the healthy category have limited room for improvement because they already comply with the standards. We have included a translation of this questionnaire in Multimedia Appendix 2. Self-efficacy for diet and PA were both measured using a 3-item questionnaire with a 5-point Likert scale ranging from 1 (very high) to 5 (very low) [32]. Knowledge was assessed using a 10-item true/false questionnaire based on the Netherlands classification model [28] for diet and a 10-item true/false questionnaire for physical activity based on the Dutch Standard for Healthy Physical Activity [33]. The total scores of these questionnaires range from 1 (very poor) to 10 (excellent). Attitude was measured using a 5-item questionnaire on health consciousness attitude and a 6-item questionnaire on health-oriented beliefs; all questions used a 5-point Likert-scale ranging from 1 (very unfavorable) to 5 (very favorable). These questionnaires were based on the research of Dutta-Bergman [34] and adapted to the Dutch situation. Self-rated behavior (henceforth self-rating) was assessed by 2 items, 1 on self-rated diet and 1 on self-rated PA, both using a scale from 1 (very poor) to 10 (excellent). Insight into behavior was calculated by comparing self-reported and self-rated diet and PA based on the classification used by Ronda et al [35]. Self-rating was recoded into categories to match the categories of self-reported behavior. Therefore, self-rated diet was recoded into three categories (1-4: unhealthy; 5-7: improvable; 8-10: healthy) and self-rated PA was recoded into two categories (1-5: unhealthy; 6-10: healthy). Respondents who did not meet the criteria for recommended healthy behavior but rated their own behavior as healthy were classified as overestimators. Respondents who did meet the criteria for healthy behavior but rated their behavior as unhealthy were classified as underestimators. The remaining respondents were considered to have had realistic insight into their behavior. Pretest and posttest questionnaires were identical except for the following additional items at posttest: the number of newsletters received and opened (waiting list group) and satisfaction with the HWA (intervention group). Satisfaction was measured using 4 items with a 5-point Likert scale ranging from 1 (very negative) to 5 (very positive) on user friendliness, usefulness, recommending to others, and willingness to continue using the HWA [36].

In addition to the online questionnaires, the HWA stored every log-on by a participant. These log files were used to attain the usage of the HWA, that is, the number of times each respondent logged on to the HWA within the intervention period.

Electronic Surveys

SurveyMonkey was used for the electronic data collection [37]. The first page of the survey consisted of an informed consent. By agreeing to participate, participants were led to the actual questionnaire. Data was protected following the security measures of SurveyMonkey [38]. Moreover, no personal identifying information apart from an email address was collected. Our survey was pretested using 5 nonparticipants comparable to the participants of the study. Feedback from the pretest was implemented in the final survey.

Our format of data collection was an “open survey” [39] posted on a website. The survey was only accessible through our research website for respondents who satisfied our inclusion criteria. The initial contact mode was through online and offline advertisements for research into an online lifestyle coach. It was mandatory for participants to fill out the questionnaire to be included in the study. We offered no incentives to participate other than the use of the lifestyle coach. The pretest questionnaire was available for 8 weeks; the posttest questionnaire was available for 3 weeks. We used randomization of items for Likert-type questions with no specific order. The number of items was 42, divided over 5 screens. All questions were mandatory except comment boxes. Respondents were able to review and, if necessary, change previous answers until they had submitted the completed questionnaire.

We were not able to record unique site visitors or survey visitors. The completion rate was 90% (269/297). To prevent multiple entries from the same person we used cookies that were stored when visiting the first page and were valid for 14 days. Also, we checked IP addresses. Entries from the same address with identical sex and birth date were checked for completeness. The most complete entry was saved, or, in case of equal completeness, the first entry was saved.

Participants

Previous research on the HWA using the same research instrument on self-reported dietary behavior yielded information on the mean and standard deviation of this primary outcome measure (mean 62.9, SD 8.43) [27]. To be able to measure a meaningful difference (3.5 points) we needed a detectable effect size of 0.4. When testing at the .05 level, and, using a power of 80%, we calculated that we needed a sample size of 200 (100 per group).

Analyses

Statistical analyses were performed using SPSS Statistics 17.0 (IBM Corporation, Somers, NY, USA). We used the multiple imputation (MI) feature of SPSS Statistics 17.0 to handle missing data of posttest nonrespondents. Demographic variables and baseline outcome measures were used as predictors in the imputation model. We used an iterative Markov chain Monte Carlo method, which is the fully conditional specification. In addition, five imputed datasets were generated on which the effectiveness analyses were performed. When possible, pooled outcomes were used for the analyses; otherwise, the five estimates were combined into a single overall estimate following the MI inference rules of Rubin [40].

Differences between users and nonusers within the intervention group were assessed using Pearson's chi-square and analysis of variance testing. Furthermore, regression analysis was used to see whether characteristics predicted use of the intervention.

Effectiveness of the intervention was assessed by intention-to-treat (ITT) using effect sizes and odds ratios. Additionally, exploratory analyses were performed on pretest and posttest scores of all participants combined and separately for the control group, the users, and the nonusers of the intervention using regression analyses and effect sizes. All reported *P* values are 2-tailed. We used no statistical measures to correct for multiple testing. Effect sizes for differences in means are presented as Cohen’s d and effect sizes for nonparametric variables are presented as *r*, calculated from the *z* scores of the Wilcoxon signed rank test [41].

Results

Response Rates

Of the 269 enrolled respondents (those who completed the pretest questionnaire), 159 respondents filled out the posttest questionnaire (response rate = 59%, 159/269). The response was significantly lower in the intervention group (51%, 65/127) than in the control group (66%, 94/142) ($P = .01$). There were baseline differences between responders (ie, respondents who filled out the posttest questionnaire) and research dropouts on outcome variables. As shown in Table 1, dropouts scored significantly lower on attitude and self-rating. In addition, within the intervention group, only 48% (30/62) of dropouts used the HWA as opposed to 78% (51/65) of responders ($\chi^2_1 = 12.424$, $P < .001$).

Table 1. Baseline differences on outcome variables between responders and dropouts

Variable	Responders (n=159)	Dropouts (n=110)	*P*
BMI (kg/m^2), mean (SD)	24.0 (2.5)	23.9 (2.5)	.83
Diet, n (%)			.18
Healthy	48 (30)	26 (24)	
Improvable	99 (62)	69 (63)	
Unhealthy	12 (8)	15 (14)	
Healthy PA, n (%)	64 (42)	41 (37)	.46
Knowledge, mean (SD)[a]	7.9 (1.1)	7.7 (1.2)	.19
Attitude, mean (SD)[b]	4.1 (0.4)	3.9 (0.5)	.001
Self-efficacy, mean (SD)[c]	2.1 (0.6)	2.2 (0.6)	.55
Self-rating, mean (SD)[d]	6.8 (1.1)	6.4 (1.5)	.02
Realistic insight, diet, n (%)	92 (60)	69 (63)	.35
Realistic insight, PA, n (%)	88 (58)	70 (64)	.60

[a] Scale from 1 (very poor) to 10 (excellent)

[b] Scale from 1 (very unfavorable) to 5 (very favorable)

[c] Scale from 1 (very high) to 5 (very low)

[d] Scale from 1 (very poor) to 10 (excellent)

Descriptive Analyses of Baseline Variables

As shown in Table 2, most of the respondents in this study were female (177/269, 66%) and in the highest education category (143/269, 53%). Mean age was 41.5 years (SD 13.5). There were no significant differences between the intervention and control group on demographic variables and reasons for use.

On outcome variables, there was one significant difference at baseline, that is, respondents in the intervention group scored significantly higher on self-efficacy than respondents in the control group. Mean scores were respectively 2.2 (SD 0.6) versus 2.1 (SD 0.6) ($F_{1,267}$ = 4.109, P = .044). The most frequently mentioned reason by respondents for wanting to use the application was to gain more insight into their own lifestyle.

Table 2. Baseline demographics and reasons for use

Variable	Total (N = 269)	Intervention (n = 127)	Control (n = 142)	*P*
Age (years), mean (SD)	41.5 (13.5)	41.2 (13.5)	41.7 (13.6)	.73
Sex, n female (%)	177 (66)	85 (67)	92 (65)	.80
Education				.71
High, n (%)	143 (53)	69 (54)	74 (52)	
Moderate, n (%)	87 (32)	42 (33)	45 (32)	
Low, n (%)	39 (15)	16 (13)	23 (16)	
Chronic condition, n (%)	48 (18)	19 (15)	29 (20)	.27
Reasons for use [a]				
Insight into lifestyle, n (%)	161 (60)	80 (63)	81 (57)	.38
Living healthier, n (%)	120 (45)	61 (48)	59 (42)	.33
Fun, n (%)	112 (42)	55 (43)	57 (40)	.62
Lose weight, n (%)	107 (40)	56 (44)	51 (36)	.21

[a] Multiple answers possible so cumulative percentages do not equal 100%

Users and Nonusers

Respondents in the waiting list (control) condition reported to have opened a mean of 3.4 (SD 1.2) out of 5 newsletters. From the log files of the HWA, we know that 81 of the 127 (64%) respondents in the intervention group used the HWA at least once, while 49% (40/81) of these used the application only once. The respondent that used the HWA most frequently used it 13 times during the intervention period of 12 weeks. The median number of times HWA was used was 1.0. Of the 127 respondents in the intervention group, 4 (3%) used the application at least the intended number of times within the intervention period (ie, once a fortnight or 6 times during the 12-week period). Satisfaction with the application was assessed within the posttest questionnaire. We used only the data provided by 50 respondents who filled out the posttest questionnaire and who had accessed the HWA at least once in the intervention period. These results are depicted in Table 3. The overall mean satisfaction score for these 50 respondents was 3.0 (SD 0.74) on the 5-point scales where 1 = very negative and 5 = very positive. A score of 3.0 lies within the neutral category.

Table 3. Satisfaction with the Healthy Weight Assistant (n = 50)

Item	Mean (SD)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)
Easy to use	3.3 (0.83)	8 (16)	22 (44)	20 (40)
Useful	2.9 (0.87)	13 (26)	25 (50)	12 (24)
Recommend to others	3.0 (0.90)	12 (24)	22 (44)	16 (32)
Keep using	2.7 (0.89)	20 (40)	22 (44)	8 (16)

Baseline differences between respondents in the intervention group who used the application (users) and the respondents in this group who did not use the HWA at least once (nonusers) are depicted in Table 4.

Table 4. Baseline differences between users and nonusers in the intervention group

Variable	Users (n=81)	Nonusers (n=46)	F or χ^2	*P*
Age (years), mean (SD)	42.6 (13.2)	38.8 (13.8)	$F_{1,125}$= 2.307	.13
Sex, n female (%)	58 (72)	27 (59)	$\chi^2_1 = 2.2$	.17
Education			$\chi^2_2 = 0.7$	.70
High, n (%)	46 (57)	23 (50)		
Moderate, n (%)	26 (32)	16 (35)		
Low, n (%)	9 (11)	7 (15)		
Chronic condition, n (%)	8 (10)	11 (24)	$\chi^2_1 = 4.5$	.04
BMI (kg/m^2), mean (SD)	24.2 (2.5)	23.7 (2.3)	$F_{1,125}$= 0.900	.35
Diet, N (%)			$\chi^2 = 8.4$	.015
Healthy, n (%)	28 (35)	6 (13)		
Improvable, n (%)	46 (57)	31 (67)		
Unhealthy, n (%)	7 (9)	9 (20)		
Healthy physical activity level, n (%)	28 (37)	19 (41)	$\chi^2_1 = 0.2$	.70
Knowledge, mean (SD)	7.9 (1.1)	7.4 (1.4)	$F_{1,125} = 4.194$	.04
Attitude, mean (SD)	4.0 (0.4)	3.9 (0.5)	$F_{1,125} = 2.665$	.11
Self-efficacy, mean (SD)	2.3 (0.6)	2.2 (0.6)	$F_{1,125} = 0.274$	.60
Self-rating, mean (SD)	6.6 (1.4)	6.5 (1.5)	$F_{1,125}$= 0.037	.85
Insight, diet			$\chi^2_2 = 8.2$	.02
Underestimation, n (%)	17 (21)	2 (4)		
Realistic insight, n (%)	52 (64)	31 (67)		
Insight, physical activity			$\chi^2_2 = 2.1$	.36
Underestimation, n (%)	1 (1)	1 (2)		
Realistic insight, n (%)	47 (58)	32 (70)		

Overall, at baseline, users were healthier (scored better on dietary behavior and had a chronic condition less often) and were more knowledgeable about healthy behavior. Furthermore, users seemed to underestimate their behavior more often than nonusers, and nonusers seemed to overestimate their behavior more often than users.

To assess whether variables of the framework proposed in the introduction could be used to predict if respondents were going to use the HWA, we performed an exploratory logistic regression using the factors from the framework (social and economic, condition-related, patient-related or constructs from behavior change theories, and reasons for use). Results of this logistic regression (Table 5) showed that one variable within the social and economic factor (ie, age) and one variable within the condition-related factor (ie, chronic health condition) significantly contributed to the model. The model showed that increased age and not having a chronic condition increased the odds of having used the application (Cox & Snell R^2 = .24, Nagelkerke R^2 = .32, Model $\chi^2_{18} = 34.15$, *P* = .01).

Table 5. Logistic regression model to predict usage of the HWA

Included		Coefficient B (Standard Error [SE])	*P*	Odds Ratio (OR) (95% Confidence Interval [CI])
Constant		-12.63 (4.013)	.002	
Factor	Variable			
Social and economic	Age	0.04 (0.018)	.02	1.04 (1.00 - 1.08)
	Internet use	0.18 (0.131)	.17	1.20 (0.93 - 1.55)
	Sex	0.50 (0.504)	.32	1.65 (0.62 - 4.44)
	Education	0.13 (0.353)	.71	1.14 (0.57 - 2.28)
Condition-related	Self-rating	−0.35 (0.379)	.36	0.71 (0.34 - 1.49)
	GP visits	1.19 (0.647)	.07	3.30 (0.93 - 11.72)
	Chronic condition	2.24 (0.749)	.003	9.40 (2.17 - 40.82)
	Diet	0.71 (0.688)	.31	2.03 (0.53 - 7.80)
	PA	0.80 (0.948)	.40	2.22 (0.35 - 14.26)
	Insight, diet	0.56 (0.667)	.40	1.76 (0.48 - 6.48)
	Insight, PA	−1.00 (0.818)	.22	0.37 (0.07 - 1.83)
Patient-related	Knowledge	0.03 (0.213)	.91	1.03 (0.68 - 1.56)
	Attitude	0.57 (0.681)	.41	1.76 (0.46 - 6.69)
	Self-efficacy	0.26 (0.458)	.57	1.30 (0.53 - 3.18)
Reasons for use	Insight into lifestyle	0.47 (0.531)	.37	1.60 (0.57 - 4.55)
	Live healthier	−0.03 (0.281)	.93	0.98 (0.56 - 1.69)
	Fun	0.13 (0.165)	.44	1.14 (0.82 - 1.57)
	Lose weight	0.16 (0.122)	.18	1.18 (0.93 - 1.50)

Furthermore, we performed a linear regression to investigate whether satisfaction with the intervention HWA predicted the number of logins (Table 6). The model showed that satisfaction did not predict frequency of use (R^2 = .05, adjusted R^2 = .04).

Table 6. Linear regression on satisfaction predicting number of log-ins to the Healthy Weight Assistant

	B (SE)	Beta
Constant	−2.61 (1.17)	
Satisfaction	0.70 (0.38)	0.23[a]

[a] *P* = .07

Effectiveness

In addition, ITT analyses were performed on all outcome variables (Table 6). We found a significant but very small effect on attitude (d = 0.08) favoring the intervention group. None of the other variables showed a significant effect of the intervention.

Complementary to the ITT analyses, we performed analyses comparing the differences of the control group with the differences of the users (results not shown). These analyses did not yield any significant effects and were comparable to the results of the ITT analyses, although the effect sizes were generally larger.

Table 7. Intention-to-treat (ITT) analyses

Variable	Intervention (n=127)		Control (n=142)		Effect Size[a] (ES) or OR (95% CI)
	Pretest	Posttest	Pretest	Posttest	
BMI, mean (SD)	24.0 (2.4)	24.1 (2.5)	23.9 (2.5)	24.0 (2.5)	ES: 0.07 (-0.10 – 0.24)
Diet					OR: 0.84 (0.44 – 1.58)
Healthy, n (%)	34 (27)	45 (35)	40 (28)	46 (32)	
Improvable, n (%)	77 (61)	73 (58)	91 (64)	89 (63)	
Unhealthy, n (%)	16 (13)	9 (7)	11 (8)	7 (5)	
Healthy PA, n (%)	49 (38.6)	58 (46)	58 (41)	69 (49)	OR: 1.10 (0.60 – 2.01)
Knowledge, mean (SD)	7.7 (1.2)	7.7 (1.3)	7.9 (1.1)	7.7 (1.3)	ES: 0.15 (−0.13 to 0.42)
Attitude, mean (SD)	4.00 (0.45)	4.03 (0.45)	4.01 (0.44)	4.02 (0.45)	ES: 0.08 (0.00 – 0.16)
Self-efficacy, mean (SD)	2.2 (0.61)	2.3 (0.70)	2.1 (0.59)	2.2 (0.64)	ES: 0.04 (−0.01 to 0.17)
Self-rating, mean (SD)	6.5 (1.4)	6.9 (1.2)	6.8 (1.2)	6.9 (1.2)	ES: 0.18 (−0.04 to 0.40)
Realistic insight, diet, n (%)	83 (65)	71 (56)	83 (59)	87 (61)	OR: 0.74 (0.35 – 1.56)
Realistic insight, PA, n (%)	79 (62)	83 (65)	84 (59)	88 (62)	OR: 0.78 (0.35 – 1.74)

[a] Effect size for ratio variables presented as Cohen’s d, that is, the number of standard deviations the intervention group (I) improved more than the control group (C) (mean improvement I – mean improvement C)/pooled SD of improvement. Effect size for ordinal variables is presented as the odds ratio.

For the group as whole (independent of randomized condition), there were significant differences between pretest and posttest scores. With respect to diet (effect size $r = -0.12$), physical activity (effect size $r = -0.09$), and self-rating (effect size $d = 0.21$) the study seemed to have had a positive influence, although the effect was small (data not shown). These differences could not be attributed to the intervention according to the ITT analyses. As mentioned in the previous paragraph, our results showed differences between users and nonusers at baseline. Therefore, we performed exploratory analyses on these groups to investigate whether choosing to use or not to use the application led to different outcomes. Tables 8 and 9 show that we found significant differences in some groups but not in others. Contrary to what we expected, only the nonusers showed a significant improvement on diet ($r = -0.23$). Examining the data more closely revealed that the control group and, to a larger extent, the users also showed improvement, although this difference was not significant. The data showed that only the users significantly improved with respect to PA behavior (effect size $r = -0.17$). The control group showed a nonsignificant improvement, while PA behavior of the nonusers deteriorated, but the change was nonsignificant. With respect to attitude, the nonusers showed a significant improvement with a medium effect size ($d = 0.28$), although the absolute difference was small. With respect to self-efficacy, the control group and the nonusers showed deterioration (effect sizes respectively $d = 0.14$ and $d = 0.33$), again with small absolute differences. Lastly, the data showed that users’ self-rated behavior was more favorable at posttest than at pretest. The size of this effect was small to medium ($d = 0.27$).

Table 8. Pretest and posttest values on outcome variables for control group, nonusers, and users

Variable	Control (n=142)		Nonusers (n=46)		Users (n=81)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
BMI, mean (SD)	23.9 (2.5)	24.0 (2.5)	23.7 (2.3)	23.9 (2.5)	24.2 (2.5)	24.2 (2.5)
Diet						
Healthy, n (%)	40 (28)	46 (32)	6 (13)	11 (24)	28 (35)	34 (42)
Improvable, n (%)	91 (64)	89 (63)	31 (68)	30 (65)	46 (57)	43 (53)
Unhealthy, n (%)	11 (8)	7 (5)	9 (20)	5 (11)	7 (9)	4 (5)
Healthy pysical activity level, n (%)	58 (41)	69 (49)	19 (41)	16 (35)	30 (37)	42 (52)
Knowledge, mean (SD)	7.9 (1.1)	7.7 (1.3)	7.4 (1.4)	7.3 (1.4)	7.9 (1.1)	7.9 (1.2)
Attitude, mean (SD)	4.0 (0.44)	4.0 (0.45)	3.9 (0.46)	4.0 (0.45)	4.0 (0.44)	4.0 (0.44)
Self-efficacy, mean (SD)	2.1 (0.59)	2.2 (0.64)	2.2 (0.62)	2.4 (0.77)	2.3 (0.61)	2.3 (0.65)
Self-rating, mean (SD)	6.8 (1.2)	6.9 (1.2)	6.5 (1.5)	6.9 (1.4)	6.6 (1.4)	6.9 (1.1)
Realistic insight, diet, n (%)	83 (59)	87 (61)	31 (67.4)	25 (54.3)	52 (64.2)	46 (56.8)
Realistic insight, PA, n (%)	84 (59)	88 (62)	32 (69.6)	27 (58.7)	47 (58.0)	56 (69.1)

Table 9. Effect size (ES) of the differences between pretest and posttest values on outcome variables for control group, nonusers, and users

Variable	Control (n = 142)		Nonusers (n = 46)		Users (n = 81)	
	ES[a]	*z (P)*[d]/ 95% CI ES[e]	ES [a]	*z (P)*[d]/ 95% CI ES[e]	ES [a]	*z (P)*[d]/95% CI ES[e]
BMI	0.02[b]	CI: −0.39 to 0.44	0.06[a]	CI: -0.64 – 0.77	0.03[a]	CI: -0.51 to 0.57
Diet	−0.09[c]	*z* = −1.45 (.15)	−0.23[b]	*z* = −2.22 (.03)	−0.13[b]	*z* = −1.62 (.11)
PA	−0.10[c]	*z* = −1.65 (.10)	−0.07[b]	Z = −0.71 (.48)	−0.17[b]	*z* = −2.12 (.03)
Knowledge	−0.15[b]	CI: −0.35 to 0.04	−0.08[a]	CI: −0.49 to 0.34	0.04[a]	CI: −0.20 to 0.29
Attitude	0.01[b]	CI: −0.06 to 0.08	0.28[a]	CI: 0.15 – 0.41	−0.05[a]	CI: −0.15 to 0.05
Self-efficacy	0.14[b]	CI: 0.03 – 0.24	0.33[a]	CI: 0.13 – 0.53	0.05[a]	CI: −0.09 to 0.19
Self-rating	0.15[b]	CI: −0.05 to 0.35	0.25[a]	CI: −0.18 to 0.68	0.27[a]	CI: 0.00 – 0.54
Insight, diet	−0.03[c]	*z* = −0.51 (.61)	−0.13[b]	*z* = −1.27 (.21)	−0.07[b]	*z* = −0.90 (.37)
Insight, PA	−0.01[c]	*z* = −0.49 (.62)	−0.13[b]	*z* = −1.21 (.23)	−0.11[b]	*z* = −1.42 (.16)

[a] Effect sizes for ratio variables are presented as Cohen's d, while effect sizes for ordinal variables are presented as *r*.

[b] Effect size (ES) presented as Cohen's d: $(mean_{post} - mean_{pre})/SD_{pooled}$

[c] Effect size presented as *r*: $z/\sqrt{(n)}$

[d] Wilcoxon signed-rank test

[e] In this column the reliability of the effect size is presented as the confidence interval for Cohen's d for ratio variables and as *z* statistic with *P* value for ordinal variables

Discussion

The results showed that the HWA was not used as often as intended. Increased age and not having a chronic condition increased the odds of having used the application at least once. Moreover, users were healthier and more knowledgeable about healthy behavior than nonusers. The ITT analyses showed no apparent effects of the intervention; however, there were differences in the effect of the intervention on users and nonusers. With respect to dietary behavior and attitude, nonusers improved more than users, while with respect to physical activity and self-rated behavior the users improved more than nonusers. On self-efficacy, the control group and the nonusers showed deterioration from baseline to posttest.

Only 64% (81 out of 127) of the participants who received access to the HWA actually used the application. This finding is not unique to this study; for example, see [6,15,20,42]. This stresses an important aspect of Web-based interventions, that is, of the respondents who agree to participate in a study on a Web-based intervention, we can expect that a substantial percentage does not use the intervention at all. In addition, we saw that the HWA is not used as often as intended in the design of the application. Of the included social and economic factors of the proposed framework, only increased age increased the odds of having used the application. This finding might seem counterintuitive, but it concurs with recent findings on the motivation to use e-consultation [43], which showed that older people were more motivated to use e-consultation than younger people. With respect to the condition-related factors, the regression analysis showed that having a chronic condition decreased the odds of using the application. An explanation might be that the HWA was not developed for people with chronic conditions and no special attention is paid to the needs of people with chronic conditions. Therefore, these people might feel that the HWA does not suit their needs. Significant differences between users and nonusers on condition-related factors showed that users were healthier. A reason for this might be that people like to be rewarded for their healthy behavior and not confronted with their unhealthy behavior.

Additionally, users more often underestimated their dietary behavior (respondents who did meet the criteria for healthy behavior but who rated their behaviour as unhealthy were classified as underestimators), while nonusers more often overestimated their behavior. This shows that the people who could have benefited most from the HWA were less likely to use the application. Of the patient-related factors or constructs from behavior change theories, only knowledge showed a significant difference between users and nonusers. Users knew more about healthy behavior, which supports the notion that the people who could have benefited most from the HWA were least likely to use the application.

There were no differences related to the reasons for use between users and nonusers, and the different reasons do not explain whether respondents used the HWA or not. However, the reasons for use might play a role in the frequency of use. The most frequently mentioned reason for wanting to use the intervention was to gain insight into one’s own behavior (60%). It might be that this goal was reached after using the HWA once, and participants might not have felt the need to use the HWA again.

Interestingly, the intervention was specifically not made to help people lose weight, but this goal was mentioned by 40% of respondents. Respondents seemed to want a quick and short-term effect (to gain insight) and might not have been willing to use the intervention frequently to work on a long-term goal (eg, a healthier lifestyle). Satisfaction with the HWA was not associated with the frequency of use. However, overall, participants were not very satisfied with the HWA, which might have contributed to the relative low usage rates. To summarize, one of the social and economic factors (ie, age), condition related factors (ie, chronic condition, self-reported behavior, and insight into behavior), and one of the patient-related factors (ie, knowledge) were related to use of the system. Satisfaction and reasons for use provided more in-depth information related to the causes of the lack of adherence to the intervention.

At baseline, the intervention and control groups showed a significant difference in attitude. The absolute value of the difference was small, however, and we don’t consider it to be a meaningful difference. Therefore, we can argue that the groups were comparable at baseline. We found no meaningful significant effects of the intervention using ITT analyses. We did find that both the waiting list group and the intervention group showed significant improvement on behavior and a significantly more favorable self-rated behavior. This well-known Hawthorne effect [44] (ie, the effect on outcome through participation in research) might be explained by the increased attention participants paid to healthy behavior due to completing questionnaires on behavior and by increased awareness of current and desired behavior. Another explanation for the improvement in all respondents might be social desirability. Thinking of the intended behavior might have influenced the responses given in the posttest questionnaire. Considering the control group, the users, and the nonusers separately showed that only nonusers significantly improved on dietary behavior. This might be due to the large differences between users and nonusers at baseline. Users were already much healthier, and both groups improved, although at posttest, the nonusers were still less healthy than the users. It seems that a ceiling effect prevented the users from improving significantly while the nonusers had much room for improvement and, for that reason, showed significant improvement. On PA, we found that only the users of the intervention improved significantly, although the effect size was not very large ($r = -0.17$). The nonusers, who chose not to use the intervention, showed a decline in behavior while the control group showed improvement. Although these differences were not significant, this does point toward a difference between choosing not to use an intervention and not being able to use an intervention. However, these differences might also reflect social desirability because of the focus on PA in the intervention. Lastly, users judged their own behavior significantly more positively after the intervention period than before. None of the other groups showed this significant improvement. Summarizing, we found no apparent effects of the HWA, but it seems that having chosen to use or not to use the intervention led to different outcomes. Combined with the differences between groups at baseline, this seems to imply that these groups are truly different and should be treated as separate entities.

In this study, we were faced with substantial dropout and nonusage rates. High dropout rates are not uncommon in this field of research and have been said to be a major challenge [45,46]. Additionally, the reduction of nonusage rates is also a major challenge [15,47]. Faced with these challenges, it is important to note that in this study the groups of dropouts and nonusers overlapped, but were not the same. Almost half of the dropouts had been users, and there were also nonusers that were responders. Consequently, it is very important keep these two concepts apart.

Our results showed that the users of the HWA were healthier than nonusers, which is an unfortunate finding not unique to this study [18]. The group for which the intervention seemed to have been most useful, namely people who had room for improvement on both diet and physical activity, were less likely to have used the HWA. This tells us that we need to try different ways to entice potential users who could benefit from the HWA to become active users. More effort should be made to tempt the nonusers of the intervention to become users. One way to do this might be to make it as easy as possible to start using the application. The moment people are interested, they should be able to use the application. In our study, there was considerable time between expressing interest and being able to use the application. Moreover, participants had to check their email, click on a link, and create a profile. All these steps require effort and could thereby decrease the odds of using the intervention. Once participants become users, the application itself can stimulate adherence. This might be done by regularly providing new content, by including reminders (through email or text messaging), or by explicitly telling participants what is expected of them in terms of usage. In our view, including these aspects would have improved the HWA.

In this study, the frameworks used to predict usage and to study effectiveness seem to have been insufficient. From the WHO framework [16], some factors, especially condition-related, seem to have explanatory power but not enough to fully explain why participants choose whether to use an intervention. This might be due to the fact that the goal of the model is adherence to treatment and not adherence to technology. Moreover, attitude, self-efficacy, and knowledge do not contribute to a better understanding of the effects of the intervention. These variables from classic behavior change theories might not discriminate enough. To gain more insight into how online interventions can support people in changing their behavior, we should try to take into account the specific barriers and opportunities of eHealth interventions and integrate them into a comprehensive conceptual framework.

Limitations

A limitation of this study is the use of self-reported behaviors. Although we used questionnaires used in previous studies, there is a chance of biased results due to social desirability or lack of insight into behavior. As a consequence, a possible change in insight into behavior might not be reflected in our results. It could be that at baseline, participants provided optimistic self-reported behavior. Due to the intervention, the users might have provided more realistic self-reported behavior at posttest. Unfortunately, this potentially positive effect of the HWA could not be tested in this study. A second limitation is related to the participants in this study. Most respondents were female and highly educated. Various studies have reported overrepresentation of this group [6,18,48,49]. Nevertheless, the question remains whether these results can be generalized to the broader target population of the HWA. Another limitation of this study is that we measured the usage of the system as the number of log-ons per participant. What participants did while logged on and for how long they were logged on, we do not know. As more and more eHealth research takes the usage of the applications into account, it might be beneficial to standardize the assessment of usage. Furthermore, a limitation of this study is related to the response rate. Our overall response rate was quite low (59%), and we found significant differences between responders and nonresponders. We accounted for this bias by using multiple imputation procedures. However, imputing 41% of the data might have yielded unreliable estimates, although research has shown that imputing up to 58% can be more reliable than listwise deletion [40,50]. In our view, this study has provided valuable insights into the users of a Web-based intervention. However, had we been able to conduct this study again, we would have changed the way we dealt with several issues. First of all, we would have included a larger number of respondents to certify a sample size large enough to account for the high dropout and nonusage rates. Second, we would have tested and adapted the application during development so that we could have chosen the outcome measures and study period to better reflect the goals and expected effects of the application. Unfortunately, this was not possible in the current study, and this stresses the importance of a close collaboration between researchers and developers of eHealth interventions.

Future Work

Usage is a major issue in research into the effects of eHealth applications. More research is needed into transforming potential users into actual users and into keeping them engaged with the application and, thereby, stimulating them to keep using the intervention. Moreover, long-term research on the use of eHealth applications is needed to provide insight into the way usage fluctuates over time. From this study, we have gained insight into differences between users and nonusers, which can be seen as a first step to decreasing attrition. The next step might be found when looking at the opportunities technology has to offer. For example, several recent studies have shown beneficial effects of adding mobile technology [51-53] and devices that provide automated tailored feedback [54]. Additionally, the field of persuasive technology might provide us with insight into how technology as a medium can persuade and motivate users to change behavior [55,56].

Acknowledgments

We are grateful for funding from the Netherlands Nutrition Centre.

Conflicts of Interest

None declared

Multimedia Appendix 1

Translated questionnaire dietary behavior

[PDF file (Adobe PDF File), 57 KB - jmir_v13i2e32_app1.pdf]

Multimedia Appendix 2

Translated questionnaire physical activity behavior

[PDF file (Adobe PDF File), 53 KB - jmir_v13i2e32_app2.pdf]

References

1. Glenny AM, O'Meara S, Melville A, Sheldon TA, Wilson C. The treatment and prevention of obesity: a systematic review of the literature. Int J Obes Relat Metab Disord 1997 Sep;21(9):715-737. [Medline: 9376884]
2. Hardeman W, Griffin S, Johnston M, Kinmonth AL, Wareham NJ. Interventions to prevent weight gain: a systematic review of psychological models and behaviour change methods. Int J Obes Relat Metab Disord 2000 Feb;24(2):131-143. [Medline: 10702762]
3. Doak CM, Visscher TL, Renders CM, Seidell JC. The prevention of overweight and obesity in children and adolescents: a review of interventions and programmes. Obes Rev 2006 Feb;7(1):111-136. [doi: 10.1111/j.1467-789X.2006.00234.x] [Medline: 16436107]
4. Müller MJ, Asbeck I, Mast M, Langnäse K, Grund A. Prevention of obesity--more than an intention. Concept and first results of the Kiel Obesity Prevention Study (KOPS). Int J Obes Relat Metab Disord 2001 May;25(Suppl 1):S66-S74 [FREE Full text] [doi: 10.1038/sj.ijo.0801703] [Medline: 11466593]
5. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
6. Rothert K, Strecher VJ, Doyle LA, Caplan WM, Joyce JS, Jimison HB, et al. Web-based weight management programs in an integrated health care setting: a randomized, controlled trial. Obesity (Silver Spring) 2006 Feb;14(2):266-272 [FREE Full text] [doi: 10.1038/oby.2006.34] [Medline: 16571852]
7. Wammes B, Oenema A, Brug J. The evaluation of a mass media campaign aimed at weight gain prevention among young Dutch adults. Obesity (Silver Spring) 2007 Nov;15(11):2780-2789 [FREE Full text] [doi: 10.1038/oby.2007.330] [Medline: 18070769]
8. Neuhauser L, Kreps GL. eHealth communication and behavior change: promise and performance. Social Semiotics 2010;20(1):9-27. [doi: 10.1080/10350330903438386]
9. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
10. Moore TJ, Alsabeeh N, Apovian CM, Murphy MC, Coffman GA, Cullum-Dugan D, et al. Weight, blood pressure, and dietary benefits after 12 months of a Web-based Nutrition Education Program (DASH for health): longitudinal observational study. J Med Internet Res 2008;10(4):e52 [FREE Full text] [doi: 10.2196/jmir.1114] [Medline: 19073541]
11. Svetkey LP, Stevens VJ, Brantley PJ, Appel LJ, Hollis JF, Loria CM, Weight Loss Maintenance Collaborative Research Group. Comparison of strategies for sustaining weight loss: the weight loss maintenance randomized controlled trial. JAMA 2008 Mar 12;299(10):1139-1148 [FREE Full text] [doi: 10.1001/jama.299.10.1139] [Medline: 18334689]
12. An JY, Hayman LL, Park YS, Dusaj TK, Ayres CG. Web-based weight management programs for children and adolescents: a systematic review of randomized controlled trial studies. ANS Adv Nurs Sci 2009;32(3):222-240. [doi: 10.1097/ANS.0b013e3181b0d6ef] [Medline: 19707091]
13. Hunter CM, Peterson AL, Alvarez LM, Poston WC, Brundige AR, Haddock CK, et al. Weight management using the internet a randomized controlled trial. Am J Prev Med 2008 Feb;34(2):119-126. [doi: 10.1016/j.amepre.2007.09.026] [Medline: 18201641]
14. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
15. 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]
16. Sabaté E. World Health Organization. 2003. Adherence to long-term therapies: evidence for action. URL: www.who.int/chp/knowledge/publications/adherence_report/en/ [accessed 2011-03-21] [WebCite Cache ID 5xM0pPQEy]
17. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. J Med Internet Res 2010;12(1):e3 [FREE Full text] [doi: 10.2196/jmir.1361] [Medline: 20147006]
18. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
19. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. J Med Internet Res 2008;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]

20. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
21. Kaptein M, Eckles D. Selecting effective means to any end: Futures and ethics of persuasion profiling. In: Proceedings of the 5th International Conference Persuasive. 2010 Presented at: Persuasive Technology; June 7 - 10, 2010; Copenhagen. [doi: 10.1007/978-3-642-13226-1_10]
22. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
23. DeLone W, McLean E. Information systems success revisited. In: Proceedings of the Thirty-fifth Hawaii International Conference on System Science. Vol 8. Los Alamitos, CA: IEEE Computer Society Press; 2002 Presented at: International Conference on System Science; January 7 - 10, 2002; Hawaii URL: http://doi.ieeecomputersociety.org/10.1109/HICSS.2002.994345
24. Noar SM, Zimmerman RS. Health Behavior Theory and cumulative knowledge regarding health behaviors: are we moving in the right direction? Health Educ Res 2005 Jun;20(3):275-290 [FREE Full text] [doi: 10.1093/her/cyg113] [Medline: 15632099]
25. Fishbein M. Developing effective behavior change interventions: some lessons learned from behavioral research. NIDA Res Monogr 1995;155:246-261. [Medline: 8594461]
26. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
27. Kelders SM, van Gemert-Pijnen JE, Werkman A, Seydel ER. Evaluation of a web-based lifestyle coach designed to maintain a healthy bodyweight. J Telemed Telecare 2010;16(1):3-7. [doi: 10.1258/jtt.2009.001003] [Medline: 20086259]
28. Netherlands Nutrition Center. 2005. Criteria for the nutritional evaluation of foods. URL: http://www.voedingscentrum.nl/resources2008/Criteriaengelssite.pdf [accessed 2010-06-18] [WebCite Cache ID 5qZzTonkw]
29. Voedingscentrum. 2009. Richtlijnen voedselkeuze. URL: http://www.voedingscentrum.nl/resources2008/RICHTLIJNENVOEDSELKEUZE20091.pdf [accessed 2011-02-10] [WebCite Cache ID 5wOUPJzrw]
30. Health Council of the Netherlands. Guidelines for a Healthy Diet 2006. URL: http://www.gezondheidsraad.nl/sites/default/files/2006@21N.pdf [accessed 2011-02-10] [WebCite Cache ID 5wOV0qNR2]
31. Douwes M, Hildebrandt VH. Vragen naar de mate van lichamelijke activiteit. Geneeskd Sport 2000;33(1):9-16.
32. Sallis JF, Pinsky RB, Grossman RM, Patterson TL, Nader PR. The development of self-efficacy scales for health-related diet and exercise behaviors. Health Educ Res 1988;3(3):283-292.
33. Kemper HGC, Ooijendijk WTM, Stiggelbout M. Consensus over de Nederlandse Norm voor Gezond Bewegen. TSG 2000;78:180-183.
34. Dutta-Bergman MJ. Health attitudes, health cognitions, and health behaviors among Internet health information seekers: population-based survey. J Med Internet Res 2004 May 28;6(2):e15 [FREE Full text] [doi: 10.2196/jmir.6.2.e15] [Medline: 15249264]
35. Ronda G, Van Assema P, Brug J. Stages of change, psychological factors and awareness of physical activity levels in The Netherlands. Health Promot Int 2001 Dec;16(4):305-314 [FREE Full text] [Medline: 11733449]
36. Tsang MW, Mok M, Kam G, Jung M, Tang A, Chan U, et al. Improvement in diabetes control with a monitoring system based on a hand-held, touch-screen electronic diary. J Telemed Telecare 2001;7(1):47-50. [Medline: 11265938]
37. Surveymonkey. URL: http://www.surveymonkey.com/ [accessed 2010-06-18] [WebCite Cache ID 5qZzM0Ftb]
38. Surveymonkey. How do you keep our data secure and where is it stored? URL: http://help.surveymonkey.com/app/answers/detail/a_id/337/kw/security [accessed 2010-06-18] [WebCite Cache ID 5qZzPBfrg]
39. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
40. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York, NY: John Wiley & Sons, Inc; 1987.
41. Rosenthal R. Meta-analytic Procedures for Social Research. Newbury Park, CA: Sage Publications, Inc; 1991.
42. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. J Med Internet Res 2008;10(5):e54 [FREE Full text] [doi: 10.2196/jmir.1016] [Medline: 19097975]
43. Nijland N, van Gemert-Pijnen JE, Boer H, Steehouder MF, Seydel ER. Increasing the use of e-consultation in primary care: results of an online survey among non-users of e-consultation. Int J Med Inform 2009 Oct;78(10):688-703. [doi: 10.1016/j.ijmedinf.2009.06.002] [Medline: 19625210]
44. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. BMC Med Res Methodol 2007;7:30 [FREE Full text] [doi: 10.1186/1471-2288-7-30] [Medline: 17608932]
45. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009;11(2):e9 [FREE Full text] [doi: 10.2196/jmir.1052] [Medline: 19403465]
46. Couper MP, Peytchev A, Strecher VJ, Rothert K, Anderson J. Following up nonrespondents to an online weight management intervention: randomized trial comparing mail versus telephone. J Med Internet Res 2007;9(2):e16 [FREE Full text] [doi: 10.2196/jmir.9.2.e16] [Medline: 17567564]

47. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
48. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. J Med Internet Res 2010;12(1):e9 [FREE Full text] [doi: 10.2196/jmir.1347] [Medline: 20350926]
49. Postel MG, de Jong CA, de Haan HA. Does e-therapy for problem drinking reach hidden populations? Am J Psychiatry 2005 Dec;162(12):2393 [FREE Full text] [doi: 10.1176/appi.ajp.162.12.2393] [Medline: 16330613]
50. King G, Honaker J, Joseph A, Scheve K. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. American Political Science Review 2001;95:49-69.
51. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
52. Tufano JT, Karras BT. Mobile eHealth interventions for obesity: a timely opportunity to leverage convergence trends. J Med Internet Res 2005;7(5):e58 [FREE Full text] [doi: 10.2196/jmir.7.5.e58] [Medline: 16403722]
53. 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]
54. Richardson CR, Brown BB, Foley S, Dial KS, Lowery JC. Feasibility of adding enhanced pedometer feedback to nutritional counseling for weight loss. J Med Internet Res 2005;7(5):e56 [FREE Full text] [doi: 10.2196/jmir.7.5.e56] [Medline: 16403720]
55. Chatterjee S, Price A. Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inform Assoc 2009;16(2):171-178. [doi: 10.1197/jamia.M2859] [Medline: 19074300]
56. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann Publishers; 2003.

Abbreviations

BMI: body mass index
CI: confidence interval
ES: effect size
GP: general practitioner
HWA: Healthy Weight Assistant
ITT: intention-to-treat
OR: odds ratio
PA: physical activity
SE: standard error
WHO: World Health Organization

Edited by G Eysenbach; submitted 30.07.10; peer-reviewed by M Wanner, M Verheijden; comments to author 15.11.10; revised version received 12.02.11; accepted 10.03.11; published 14.04.11.

Please cite as:
Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER
Effectiveness of a Web-based Intervention Aimed at Healthy Dietary and Physical Activity Behavior: A Randomized Controlled Trial About Users and Usage
J Med Internet Res 2011;13(2):e32
URL: http://www.jmir.org/2011/2/e32/
doi:10.2196/jmir.1624
PMID:21493191

Original Paper

Effects of a Multilingual Information Website Intervention on the Levels of Depression Literacy and Depression-Related Stigma in Greek-Born and Italian-Born Immigrants Living in Australia: A Randomized Controlled Trial

Litza A Kiropoulos[1,2], BSc(Hons) MClinPsych PhD; Kathleen M Griffiths[3], PhD; Grant Blashki[4], MD MBBS FRACGP

[1]The University of Melbourne, Psychological Sciences, Melbourne, Australia

[2]School of Psychology & Psychiatry and School of Primary Health Care, Monash University, Melbourne, Australia

[3]Depression & Anxiety Consumer Research Unit, ehub: emental health Research & Development, Centre for Mental Health Research, The Australian National University, ACT, Australia

[4]Nossal Institute for Global Health, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Litza A Kiropoulos, BSc(Hons) MClinPsych PhD
The University of Melbourne
Psychological Sciences
The University of Melbourne Psychology Clinic
14-20 Blackwood Street, University of Melbourne
Melbourne, 3010
Australia
Phone: 61 3 9035 4063
Fax: 61 3 9326 7616
Email: litzak@unimelb.edu.au

Abstract

Background: Little is known about the efficacy of Internet-based information interventions in increasing depression literacy or reducing depression stigma and depressive symptoms in people from non–English-speaking backgrounds.

Objective: Our objective was to investigate the effects of Multicultural Information on Depression Online (MIDonline), an Internet-based multilingual depression-specific information resource, on depression literacy, depression stigma, and depressive symptoms in Greek-born and Italian-born immigrants to Australia.

Method: In all, 202 Greek- and Italian-born immigrants aged 48 to 88 years were randomly allocated to an online depression information intervention (n =110) or a depression interview control group (n = 92). Participants allocated to the information intervention only had access to the website during the 1- to 1.5-hour intervention session. The primary outcome measures were depression literacy (depression knowledge), personal stigma (personal stigma toward people with a mental illness), perceived stigma (participants' views about the probable attitude of the general community toward people with mental illness), and depressive symptoms. Depression literacy, personal and perceived stigma, and depressive symptoms were assessed at preassessment, postassessment, and at a 1-week follow-up assessment. The trial was undertaken at Monash University, Melbourne, Australia. Randomization and allocation to trial group were carried out using a computer-generated table.

Results: For depression literacy, there was a significant difference between the MIDonline and the control group with those in the MIDonline intervention displaying higher depression literacy scores postassessment ($F_{1,178} = 144.99$, $P < .001$) and at the follow-up assessment ($F_{1,178} = 129.13$, $P < .001$) than those in the control group. In addition, those in the MIDonline intervention showed a significantly greater decrease in mean personal stigma scores postassessment ($F_{1,178} = 38.75$, $P < .001$) and at the follow-up assessment ($F_{1,176} = 11.08$, $P = .001$) than those in the control group. For perceived stigma, there was no significant difference between the MIDonline intervention and the control group at postassessment ($F_{1,178} = 0.60$, $P = .44$) and at the follow-up assessment ($F_{1,176} = 1.06$, $P = .30$). For level of depression, there was no significant difference between the MIDonline intervention and the control group at preassessment ($F_{1,201} = 0.56$, $P = .45$), postassessment ($F_{1,178} = 0.03$, $P = .86$), or at the follow-up assessment, ($F_{1,175} = 1.71$, $P = .19$). Within group effect sizes for depression literacy were −1.78 (MIDonline) and −0.07 (control);

for personal stigma, they were 0.83 (MIDonline) and 0.06 (control); for perceived stigma, they were 0.14 (MIDonline) and 0.16 (control); and for depressive symptoms, they were 0.10 (MIDonline) and 0.10 (control).

Conclusions: Current results suggested that the Internet may be a feasible and effective means for increasing depression knowledge and decreasing personal stigma in non–English-speaking immigrant populations residing in English-speaking countries. The lack of change in perceived stigma in this trial is consistent with results in other trials examining online depression stigma interventions in English-speaking groups.

Trial Registration: ISRCTN76460837; http://www.controlled-trials.com/ISRCTN76460837 (Archived by WebCite at http://www.webcitation.org/5xjxva4Uq)

(J Med Internet Res 2011;13(2):e34) doi:10.2196/jmir.1527

KEYWORDS

Depression literacy; depression-related stigma; immigrants; Internet-based interventions; depression; randomized control trial

Introduction

Low levels of depression literacy (also called depression knowledge) and high levels of stigma associated with mental disorders may be barriers to seeking help from health professionals [1-3]. Previous research has indicated that stigma associated with having a mental disorder is more prominent in non–English-speaking immigrant communities, especially among people born in Greece [4]. In 2006, Australia was home to over 110,000 Greek-born and 199,100 Italian-born people, with Melbourne having the largest Greek-born population outside of Greece [5]. In addition, the 2 most common of the 400 non-English languages spoken at home in Australia in 2006 were Italian and Greek, accounting for 1.6% and 1.3% respectively of the languages spoken at home among the Australian population [5]. However, despite non–English-speaking people comprising a substantial proportion of Australia's population, with Italian-born and Greek-born making up the third and sixth largest non–English-speaking overseas-born groups living in Australia [5], there is a lack of research into depression literacy and stigma related to depression in these groups.

Online psychoeducational interventions have been reported to be effective in increasing depression literacy and reducing personal stigma (stigma toward people with a mental illness) related to depression in English-speaking populations [6]. Online information and psychological interventions have also been found to be effective in the treatment of depressive symptoms in adults (eg, [7-9]) and in addressing stigma related to mental health [10]. However, there is a lack of culturally appropriate evidence-based educational interventions for depression for non–English-speaking immigrant communities especially in the Australian context. There are good reasons to target such interventions to middle- and older-aged immigrants of a non–English-speaking background, especially those who are Greek- and Italian-born. Compared with Anglo-Australians, people in these groups have been found to hold stronger stigmatizing attitudes toward mental illness such as schizophrenia and depression, they tend to be at higher risk of developing depression, and they have been found to underutilize professional psychological services [4,11]. In this study, we evaluated the impact of a new, open access, Web-based multilingual informational intervention, Multicultural Information on Depression online (MIDonline), on depression literacy, personal (personal stigma toward people with a mental illness) and perceived stigma (participants' views about the probable attitude of the general community toward people with a mental illness), and depressive symptoms among Greek-born and Italian-born immigrants living in Melbourne, Australia.

Methods

Trial Design

This was a single centre, cross-sectional, parallel group, randomized controlled trial with balanced randomization (ie, 1:1), and with stratification at level of country. The study was conducted at Monash University, Melbourne, Australia.

Participants

Participants were a community sample of 129 Greek-born and 73 Italian-born immigrants living in Melbourne, Australia. They were recruited between November 11, 2006, and June 6, 2009, by advertising the research project in Greek and Italian social and welfare clubs and in the print and radio media directed at Greek- and Italian-speaking residents of Melbourne. Eligibility criteria for the study included being 45 years of age or over and having been born in Greece or Italy. The total sample comprised 144 women and 58 men. The mean participant age was 65.4 years (SD 8.57, range 48 to 88 years). Monash University Committee on Ethics in Research Involving Humans granted approval for the study.

Interventions

MIDonline Intervention

The intervention comprised the consumer section of the MIDonline website [12], which provides online multilingual information about depression designed for middle- to older-aged consumers from a non–English-speaking background. The material is available in Greek, Italian, and English. The website content incorporates information about symptoms and case studies of depression, how depression is diagnosed, related disorders, causes, treatment options, how to find a bilingual mental health professional and professional psychological care, stigma related to mental illness, and multilingual translated resources. The information is provided in a culturally relevant way. For example, case studies are representative of middle- to older-aged people of both genders who are Greek- and Italian-born.

Figure 1. Screenshot of the MIDonline website intervention in English

Multicultural Information on Depression

General Information

Providing multilingual and culturally relevant information on depression for culturally and linguistically diverse (CALD) adult populations in Australia.

The information provided on this website is designed to support, not replace, the relationship that exists between a patient or site visitor and his / her physician or mental health professional.

CONTINUE.....

Figure 2. Screenshot of the MIDonline website in Italian

SINTOMI DELLA DEPRESSIONE

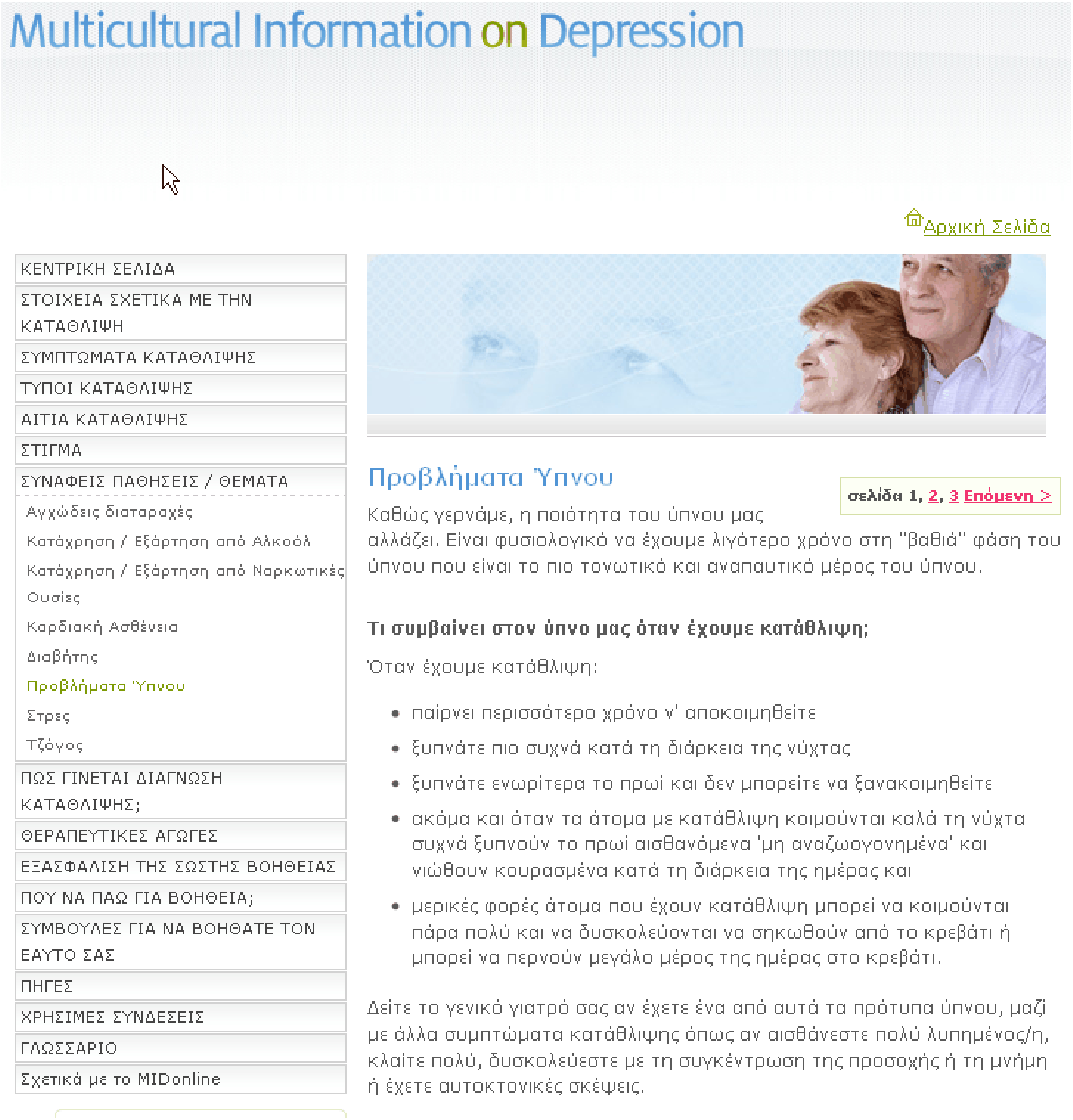
Se stai affrontando alcuni di questi sintomi quasi ogni giorno per un periodo minimo di 2 settimane e ti senti triste o ti manca maggiormente interesse o piacere forse soffri di depressione.

Non tutti i sofferenti della depressione sono afflitti da tutti questi sintomi. In alcune persone si manifestano pochi sintomi, ed in altre si manifestano molti sintomi.
La gravità dei sintomi depressivi e la loro durata è diversa da una persona all'altra.

Emozionali

- *Tristezza persistente, ansia o sentirsi vuoti ed "intorpiti"*
- *Sentimenti di disperazione, pessimismo, di basso autovalore*
- *Sensi di colpevolezza*
- *Sentirsi non amato da nessuno*

Figure 3. Screenshot of the MIDonline website in Greek



ΚΕΝΤΡΙΚΗ ΣΕΛΙΔΑ
ΣΤΟΙΧΕΙΑ ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΚΑΤΑΘΛΙΨΗ
ΣΥΜΠΤΩΜΑΤΑ ΚΑΤΑΘΛΙΨΗΣ
ΤΥΠΟΙ ΚΑΤΑΘΛΙΨΗΣ
ΑΙΤΙΑ ΚΑΤΑΘΛΙΨΗΣ
ΣΤΙΓΜΑ
ΣΥΝΑΦΕΙΣ ΠΑΘΗΣΕΙΣ / ΘΕΜΑΤΑ
  Αγχώδεις διαταραχές
  Κατάχρηση / Εξάρτηση από Αλκοόλ
  Κατάχρηση / Εξάρτηση από Ναρκωτικές Ουσίες
  Καρδιακή Ασθένεια
  Διαβήτης
  Προβλήματα Ύπνου
  Στρες
  Τζόγος
ΠΩΣ ΓΙΝΕΤΑΙ ΔΙΑΓΝΩΣΗ ΚΑΤΑΘΛΙΨΗΣ;
ΘΕΡΑΠΕΥΤΙΚΕΣ ΑΓΩΓΕΣ
ΕΞΑΣΦΑΛΙΣΗ ΤΗΣ ΣΩΣΤΗΣ ΒΟΗΘΕΙΑΣ
ΠΟΥ ΝΑ ΠΑΩ ΓΙΑ ΒΟΗΘΕΙΑ;
ΣΥΜΒΟΥΛΕΣ ΓΙΑ ΝΑ ΒΟΗΘΑΤΕ ΤΟΝ ΕΑΥΤΟ ΣΑΣ
ΠΗΓΕΣ
ΧΡΗΣΙΜΕΣ ΣΥΝΔΕΣΕΙΣ
ΓΛΩΣΣΑΡΙΟ
Σχετικά με το MIDonline

Προβλήματα Ύπνου

σελίδα 1, 2, 3 Επόμενη >

Καθώς γερνάμε, η ποιότητα του ύπνου μας αλλάζει. Είναι φυσιολογικό να έχουμε λιγότερο χρόνο στη "βαθιά" φάση του ύπνου που είναι το πιο τονωτικό και αναπαυτικό μέρος του ύπνου.

Τι συμβαίνει στον ύπνο μας όταν έχουμε κατάθλιψη;

Όταν έχουμε κατάθλιψη:

- παίρνει περισσότερο χρόνο ν' αποκοιμηθείτε
- ξυπνάτε πιο συχνά κατά τη διάρκεια της νύχτας
- ξυπνάτε ενωρίτερα το πρωί και δεν μπορείτε να ξανακοιμηθείτε
- ακόμα και όταν τα άτομα με κατάθλιψη κοιμούνται καλά τη νύχτα συχνά ξυπνούν το πρωί αισθανόμενα 'μη αναζωογονημένα' και νιώθουν κουρασμένα κατά τη διάρκεια της ημέρας και
- μερικές φορές άτομα που έχουν κατάθλιψη μπορεί να κοιμούνται πάρα πολύ και να δυσκολεύονται να σηκωθούν από το κρεβάτι ή μπορεί να περνούν μεγάλο μέρος της ημέρας στο κρεβάτι.

Δείτε το γενικό γιατρό σας αν έχετε ένα από αυτά τα πρότυπα ύπνου, μαζί με άλλα συμπτώματα κατάθλιψης όπως αν αισθάνεστε πολύ λυπημένος/η, κλαίτε πολύ, δυσκολεύεστε με τη συγκέντρωση της προσοχής ή τη μνήμη ή έχετε αυτοκτονικές σκέψεις.

Control Condition

The control intervention consisted of a semistructured interview with a bilingual interviewer who asked open-ended questions relating to the participant's beliefs about depression including the causes, symptoms, course and development, treatments, and outcomes of depression.

Procedures

Greek-born and Italian-born participants who opted to take part in the research contacted the researchers who were listed in the advertisements related to the study. The interviewers organized a suitable time for eligible participants to take part either in the MIDonline intervention or the depression interview. Each participant was assigned a bilingual interviewer who administered the preassessments, postassessments, follow-up assessments, and interventions. The investigators were not involved in conducting any of the interviews or interventions.

All participants took part in two face-to-face sessions. During the first session, participants were administered a face-to-face preintervention questionnaire followed by the MIDonline intervention or the depression control interview and then a face-to-face postintervention questionnaire. The second session included a follow-up face-to-face questionnaire administered 1 week later. Questionnaires were completed in an interview format with the bilingual researchers due to the suspected low literacy levels among participants. All participants who took part in the study completed the preintervention, postintervention, and 1-week follow-up questionnaires. The follow-up questionnaire was completed an average of 7.95 days (SD 2.34) after participants had completed the postintervention questionnaire. The project was conducted from November 11, 2006, through June 6, 2009. The intervention delivery and data collection took place in a consultation room located at Monash University.

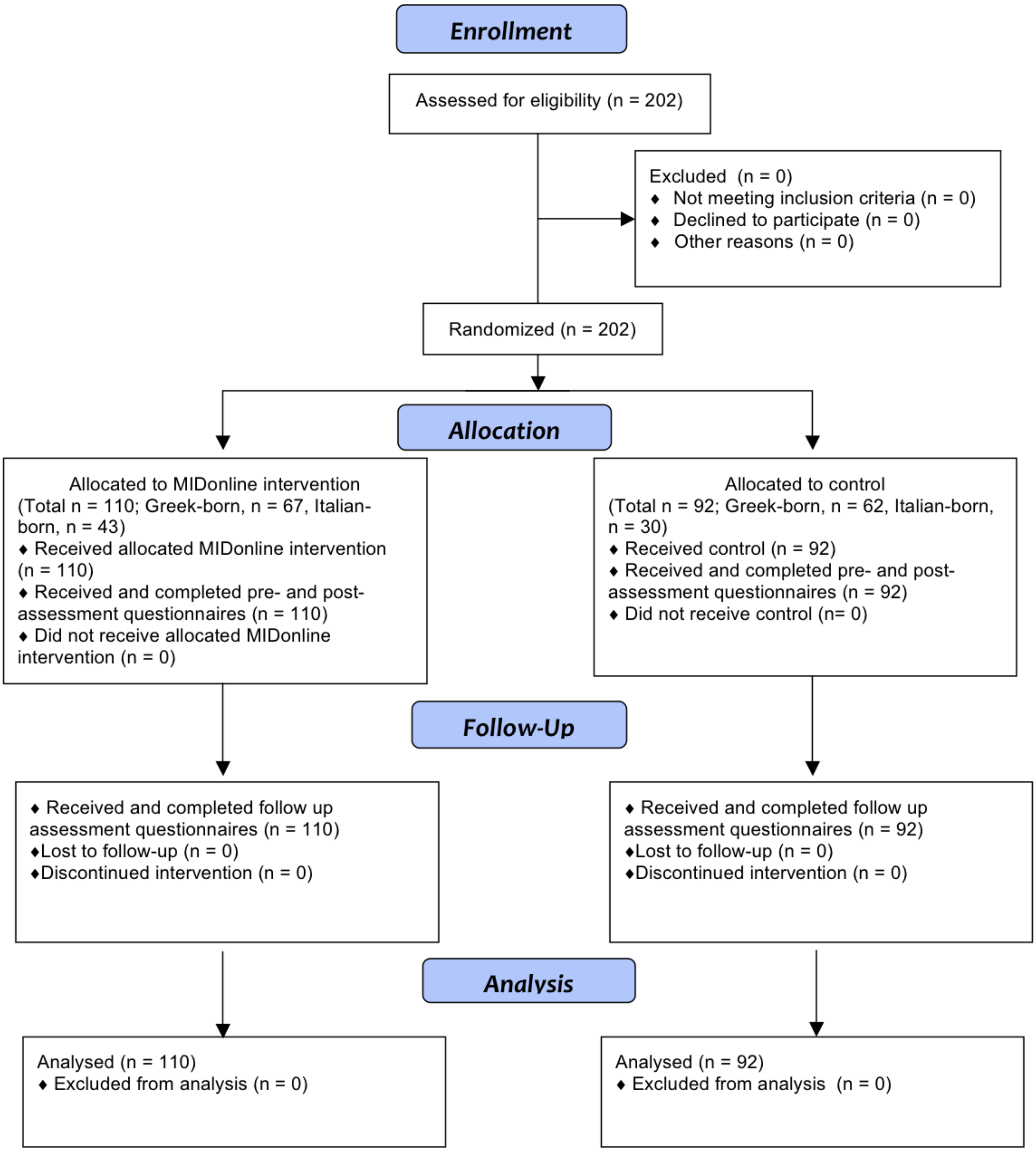
Prior to the study, all bilingual interviewers attended a training session during which they were provided with instruction and written verbatim protocols for conducting both the MIDonline intervention and the depression interview control condition.

For the MIDonline intervention, the participant and interviewer sat together in front of a computer displaying the MIDonline website. In the first 10 minutes of this session, interviewers explained that the purpose of the website was to increase knowledge about depression and instructed participants on how to navigate through the main sections of the website. These sections included information on depressive symptoms, case studies, how depression is diagnosed, treatment options, and stigma related to mental illness. Participants were then given an hour to read through the online material by themselves. Although the interviewers remained in the room during this period, they were instructed not to discuss the material with the participants. Participants only had access to the MIDonline website during the intervention on this one occasion.

Participants in the control condition took part in an interview in which interviewers asked participants a set of open-ended questions relating to the participant's beliefs about the causes, important symptoms, course and development, treatments, and outcomes of depression. Both the MIDonline intervention and the control interview took 1 to 1.5 hours each to complete.

In all, six bilingual interviewers assisted participants through the MIDonline intervention or the depression attention control interview, four of whom spoke Italian and two of whom spoke Greek. All interviewers had experience conducting bilingual interviews and were registered psychologists (except for one Greek bilingual interviewer and one Italian bilingual interviewer, each of whom was a probationary psychologist). Both the intervention and the control interview were undertaken in the participant's language of choice (either Greek or Italian). All participants provided written informed consent prior to participating in the study.

Figure 4. Flow of participants through research trial



Participation

Figure 4 displays the participant flowchart for the trial according to consolidated standards of reporting trials (CONSORT) 2010 guidelines [13].

Measures

The face-to-face questionnaires incorporated measures of sociodemographic status, personal and perceived stigma, depression severity, and clinical history. All self-report scales were translated from English into Greek and Italian for the purpose of this study by the first author (LK) and other bilingual psychologists. All item translations were reconsidered by a second bilingual psychologist and researcher. More difficult or ambiguous items were examined for meaning with lay members of the Greek and Italian communities and other mental health professionals working with these communities at the time. This method was preferred over sequential/back translation methods by an interpreter as it capitalized on the expertise of bilingual mental health professionals [4]. Validity was checked by examining the psychometric properties of the scales (factor structure and internal consistency) after data were collected and preceding any further analysis.

Sociodemographic Variables and Clinical History

The following sociodemographic variables were measured: age in years, gender, marital status (married, widowed, separated, single), birthplace and currently working (yes/no), and length of stay in Melbourne (in years). Level of education was measured on a 7-point scale (did not attend school, did not complete primary, completed primary, completed some secondary school, completed all secondary school, completed some tertiary, completed tertiary). Occupation level was measured on an 8-point scale (never did paid work, unskilled/semiskilled work, qualified tradesperson, clerical/office work, manager of a small business/shop/farm, manager of a small organization/company, professional requiring some university education, professional requiring high level of university education). Current living situation was measured with a 5-point scale where participants were asked about whether spouse, children, siblings, in-laws, parents, or other relatives were living at home with them. The Greek or Italian identity respectively with responses measured on a 4-item scale (not at all, a little, much, very much) with higher scores indicating greater ethnic identification. Items examined how Greek or Italian the participant felt, how important being Greek or Italian was to them, and to what extent they maintained the Greek or Italian way of doing things. Using the same rating scales, all participants were asked to rate the extent to which they considered themselves as having an Australian identity, how Australian the person felt, how important being Australian was to them, and to what extent they maintained the Australian way of doing things. English language proficiency was measured with a 5-item scale with responses measured on a 5-point scale (poor, fair, good, very good and excellent). Items depicted situations of different communicative difficulty and included shopping, regular banking, attending an English-speaking doctor, discussing important finances with the bank manager, and writing on formal business. Memory functioning was measured with a 6-item scale that asked participants about their memory of people's names, stories, appointments, dates, news, and grocery lists. Responses were measured on a 5-point scale (much poorer than average, a little poorer than average, average, a little better than average, much better than average). Higher scores indicated better memory functioning. Participants were asked whether they were currently receiving treatment for an emotional or psychological problem (yes/no) and what type of treatment was being received and by whom. Level of depressive symptoms was measured with the Beck Depression Inventory–II [14].

Outcome Measures

The primary outcome measures were levels of depression literacy, personal and perceived depression-related stigma, and severity of depressive symptoms.

Depression Literacy

Depression literacy was assessed using translated and adapted versions of the D-Lit scale [6] which is a 22-item true/false test of knowledge about depression. A higher score on this scale indicated greater literacy. An example of an item on the D-lit is: "People with depression may feel guilty when they are not at fault." Of the original items, 4 were replaced with the following items to better reflect the content of the MIDonline website. These items included: "Major depression is one of the leading causes of disability and loss of quality of life in the world," "Causes of depression range from genetics, our environment, personality, our biochemistry and having other medical conditions," "Nearly twice as many women as men are affected by depression," and "Both antidepressants and talk therapies like CBT can be effective in treating depression." The internal reliability analyses in the present sample indicated an alpha coefficient of .88 for the Greek version and .92 for the Italian version.

Depression Stigma

Changes in personal (reflecting participants' personal attitudes) and perceived (reflecting participants' beliefs about the attitudes of others) depression stigma were assessed using the 18-item Depression Stigma Scale [6]. Half of the items required participants to rate how strongly they personally agreed with a statement about depression (eg, "people with depression are unpredictable"). The other half of the items asked the participant to indicate what they thought most other people believed about the same issue (eg, "most people believe that people with depression are unpredictable"). Ratings were made on a 5-point Likert scale. Scores on the total scale range from 0 to 36 for the full scale and 0 to 18 for each of the two 9-item subscales, with higher scores indicating greater stigma. Cronbach alpha values for the total, personal, and perceived stigma scales were .68, .62, and .82 respectively for the Greek version. The correlation between the scores on the personal and perceived stigma scales was .13 (n = 129, $P > .05$). For the Italian version, Cronbach alpha values for the total, personal, and perceived stigma scales were .80, .76 and .72 respectively, and the correlation between the scores on the personal and perceived stigma scales was .39 (n = 73, $P = .01$).

Depressive Symptoms

Depression severity was measured with the Beck Depression Inventory–II (BDI–II) [14], which is a 21-item instrument for measuring severity of depression in adults. Responses to items covered the "past two weeks, including today." Responses on the BDI–II items range from 0 to 6 with higher values indicating higher severity. The internal reliability analyses in the present sample indicated an alpha coefficient of .90 for the Greek version and .89 for the Italian version.

Randomization

Participants were randomly assigned by the first author following a simple randomization procedure using a computerized list of random numbers to one of two intervention groups (either the MIDonline intervention (n = 110) or the control group (n = 92) using a 1:1 allocation with stratification at level of country). The sequence of numbers was concealed until the intervention was assigned. Interviewers and participants were not blinded to condition assignment. Participants contacted the bilingual interviewers who were listed on an advertisement to take part in the study.

Sample Size

To detect a significant difference in depression literacy of at least 2 points (based on data reported by Griffiths et al [6]) with a two-sided 5% significance level and a power of 80%, a total sample of 128 (64 participants in each group) was required.

Statistical Analyses

Differences between sociodemographic variables were examined using chi-square analyses. SPSS version 17 (IBM Corporation, Somers, NY, USA) was used to analyze the data.

To examine differences between intervention groups at baseline, individual analyses of variance (ANOVAs) were performed with preintervention depression literacy, personal stigma, perceived stigma, and level of depression scores.

To examine group differences at postintervention, individual analyses of covariance (ANCOVAs) were performed on the postintervention depression literacy, personal stigma, perceived stigma, and level of depression scores. The sociodemographic variables that differed between groups at baseline were used as covariates in all ANCOVA analyses. Covariates for these analyses included Australian identification, currently living alone, living arrangements, speaking English with an English-speaking doctor, and memory functioning. All covariates were dichotomized prior to use in the ANCOVA analyses. Preassessment depression literacy, personal stigma, perceived stigma, and level of depression were also used as covariates in the individual ANCOVA postassessment analyses.

To examine group differences at follow-up intervention, two sets of individual ANCOVAs were performed on follow-up intervention depression literacy, personal stigma, perceived stigma, and level of depression scores. Covariates for these analyses included all the demographic variables that differed between intervention groups at baseline. These included Australian identification, currently living alone, living arrangements, visiting an English-speaking doctor, and memory functioning. The first set of individual ANCOVAs was performed using the demographic variables that differed at baseline and the preintervention score of the variable as covariates. The second set of individual ANCOVAs was performed using the demographic variables that differed at baseline and the postintervention score of the variable as covariates.

Effect sizes from pre to post, were calculated using Cohen's d (standardized mean difference at immediate post) [15].

Results

Sample Characteristics

Table 1 summarizes the characteristics of the two birthplace groups in each condition. Overall, there were no statistically significant differences between the MIDonline and control groups on the characteristics measured. The mean age for the total sample was 65.4 years (SD 8.57). There was no significant difference in participant age for the two conditions: MIDonline mean age 65.6 years (SD 8.1), control mean age 65.2 years (SD 9.0) ($F_{1, 201}$ = 0.10, $P > .05$, n = 202). The percentages of women allocated to each condition did not differ significantly (60 out of 110 [54.5%] participants were allocated to MIDonline, and 42 out of 92 [45.5%] participants were allocated to the control condition, χ^2_1, = 0.57, $P > .05$, n = 202).

There was no significant difference in the level of education across the conditions (χ^2_6 = 11.01, $P > .05$), with the majority of participants reporting that they had completed all of primary school or some secondary school. The majority of Greek-born and Italian-born participants in the MIDonline and control conditions were married, living with their spouse, had been mainly unskilled or semiskilled workers, and were currently retired. Overall, very few participants were recent arrivals to Australia, with the majority having lived in Melbourne for an average of 43.8 years (SD 9.0). There were no differences found for length of stay for the two conditions; the mean for participants in the MIDonline group was 43.9 years (SD 9.89), and the mean for participants in the control group was 43.6 years (SD 8.34) ($F_{1, 201}$ = 0.06, $P > .05$, n = 201).

The majority of the sample rated their English proficiency as "good" for very simple situations like shopping (χ^2_3 = 5.14, P = .16, n = 202), but "poor/fair" for more difficult situations such as doing their regular banking (χ^2_3 = 3.50, P = .32, n = 202), visiting an English-speaking doctor (χ^2_3 = 7.76, P = .05, n = 202), discussing their finances (χ^2_3 = 2.51, P = .47, n = 202), and important business (χ^2_3 = 5.55, P = .13, n = 202). The majority of the sample reported having "average" memory functioning. In addition, the majority of the sample was not receiving treatment for a psychological or emotional problem. However, those who were receiving treatment were doing so from a general practitioner, psychologist, or psychiatrist, and treatment included taking medication.

For the Greek-born, there was no significant difference in the endorsement of Greek or Australian identification across conditions. Overall, the majority of Greek-born participants identified "very much" with being Greek and identified "much" with being Australian. However, for the Italian-born, there was a significant difference in the endorsement of Italian and Australian identification across conditions. Whereas the majority of Italian-born participants in the MIDonline condition identified "much" with being Italian and "a little" with being Australian, the majority in the control condition identified "very much" with being Italian and "very much" with being Australian.

Table 1. Demographic characteristics of respondents for the MIDonline and control intervention conditions

Variables	MIDonline (n = 110) n (%)	Control (n = 92) n (%)	*P* Value
Gender			.45
Male	34 (30.9%)	24 (26.1%)	
Female	76 (69.1%)	68 (73.9%)	
Married			.21
Yes	35 (31.8%)	22 (23.9%)	
No	75 (68.2%)	70 (76.1%)	
Greek identification			.24
Not at all/a little	6 (9.0%)	7 (11.9%)	
Much	17 (25.4%)	22 (37.3%)	
Very much	44 (65.7%)	30 (50.8%)	
Italian identification			.07
Not at all/a little	7 (16.3%)	8 (26.7%)	
Much	23 (53.5%)	8 (26.7%)	
Very much	13 (30.2%)	14 (46.7%)	
Australian identification			.03
Not at all/a little	52 (47.3%)	31 (34.8%)	
Much	45 (40.9%)	35 (39.3%)	
Very much	13 (11.8%)	23 (25.8%)	
Currently living alone			.01
Yes	22 (20%)	7 (7.6%)	
No	88 (80%)	85 (92.4%)	
Living arrangements			.03
with spouse	23 (22.5%)	7 (8.4%)	
with children	55 (53.9%)	50 (60.2%)	
with other relatives	24 (23.5%)	26 (31.3%)	
Level of education			.14
No/incomplete primary	17 (15.5%)	14 (15.2%)	
Completed primary	39 (35.5%)	46 (50%)	
Some secondary school	30 (27.3%)	19 (20.7%)	
All secondary school	15 (13.6%)	5 (5.4%)	
Some/completed tertiary	9 (8.2%)	8 (8.7%)	
Occupation			.60
Never worked	4 (3.7%)	6 (6.6%)	
Unskilled	65 (60.7%)	52 (57.1%)	
Tradesperson/clerical	35 (32.7%)	28 (30.8%)	
Manager/professional	3 (2.8%)	5 (5.5%)	
Working now			.11
Yes	26 (23.9%)	31 (34.1%)	
No	83 (76.1%)	60 (65.9%)	
Memory Functioning Index			<.001

Variables	MIDonline (n = 110) n (%)	Control (n = 92) n (%)	*P* Value
Above average	6 (5.5%)	24 (26.1%)	
Average	103 (94.5%)	68 (73.9%)	
Receiving psychological treatment			.26
Yes	9 (8.2%)	12 (13%)	
No	101 (94.5%)	80 (87%)	
Treatment from whom?			.24
No one	101 (91.8%)	81 (88%)	
Psychologist	7 (6.4%)	5 (5.4%)	
GP	2 (1.8%)	3 (3.3%)	
Psychiatrist	0 (0%)	3 (3.3%)	
What type of treatment?			.08
None	101 (91.8%)	81 (88%)	
Counseling	3 (2.7%)	0 (0%)	
Medication	6 (5.5%)	11 (12%)	

Test-Retest Reliability of the Translated Scales

The test-retest reliability measures (Pearson’s correlation coefficients) based on pretest and follow-up test data for the combined control and MIDonline conditions for the Greek translations of all scales were .91 (n = 128) for the depression literacy scale, .80 (n = 128) for the personal stigma subscale, and .83 (n = 128) for the perceived stigma subscale ($P < .001$ in each case). For the Italian translations of all scales the test-retest reliability measures based on pretest and follow-up test data for the combined control and MIDonline conditions were .88 (n = 71) for the depression literacy scale, .78 (n = 71) for the personal stigma subscale, and .65 (n = 71) for the perceived stigma subscale ($P < .001$ in each case).

Table 2. Mean depression literacy, personal and perceived stigma and level of depression scores for each intervention group over time and associated *P* values

	MIDonline (n = 110) Mean (SD)	Control (n = 92) Mean (SD)	*P* Value
Depression literacy			
Preintervention	10.61 (3.28)	8.17 (4.29)	< .001
Postintervention	17.43 (3.99)	8.03 (4.33)	< .001
Follow-up	16.84 (3.58)	8.22 (4.33)	< .001[a], .01[b]
Personal stigma			
Preintervention	18.38 (4.78)	18.44 (4.66)	.92
Postintervention	14.69 (3.64)	18.35 (4.57)	< .001
Follow-up	15.02 (3.95)	17.67 (4.73)	.001[a], .06[b]
Perceived stigma			
Preintervention	22.61 (4.39)	20.76 (5.40)	.008
Postintervention	21.65 (4.49)	21.06 (5.74)	.44
Follow-up	21.95 (4.13)	20.29 (5.55)	.30[a], .03[b]
Level of depression			
Preintervention	8.10 (7.84)	8.94 (7.82)	.45
Postintervention	7.26 (7.64)	8.13 (7.53)	.87
Follow-up	6.34 (6.60)	8.26 (7.88)	.18[a], .19[b]

[a] ANCOVA employed the preintervention measure of the variable as a covariate.

[b] ANCOVA employed the postintervention measure of the variable as a covariate.

Effects for Depression Literacy

Table 2 shows the mean depression literacy scores by intervention group over time and associated *P* values of the ANCOVAs performed. The effect of the intervention on depression literacy was examined by conducting an ANCOVA on the postassessment depression literacy score adjusting for baseline depression literacy (as well as Australian identification, currently living alone, living arrangements, speaking English to an English-speaking doctor, and memory functioning). This analysis revealed that those in the MIDonline intervention group showed a substantial increase in depression literacy scores compared with those in the control group after controlling for baseline depression literacy scores and relevant demographic variables ($F_{1,178}$ = 144.99, $P < .001$). A similar analysis was conducted on the follow-up intervention scores with the ANCOVA demonstrating superior depression literacy for the MIDonline intervention than the control group after controlling for baseline depression literacy and relevant demographic variables ($F_{1,178}$ = 129.13, $P < .001$) (see Table 2).

Finally, an additional ANCOVA was conducted to examine whether there were any differences between the MIDonline and control groups in depression literacy at the follow-up intervention stage after adjusting for the postintervention depression literacy score and the demographic covariates that differed at baseline. This analysis indicated that there was a significant but small reduction in the intervention effect on depression literacy at follow-up ($F_{1,176}$ = 6.35, $P < .01$).

Effects for Personal Stigma

There was no significant difference between the MIDonline and control groups for baseline personal stigma (see Table 2). ANCOVAs on the postintervention and follow-up personal stigma scores (adjusting for the preassessment personal stigma score and the demographic variables that differed between intervention groups at baseline) showed that MIDonline was associated with lower postintervention ($F_{1,178}$ = 38.75, $P < .001$) and follow-up ($F_{1,176}$ = 11.08, $P = .001$) personal stigma scores than the control group. However, a further ANCOVA of the follow-up personal stigma measures controlling for postintervention personal stigma levels indicated that there was a trend toward a small reduction in the effect at follow-up ($F_{1,176}$ = 3.65, $P < .06$) (see Table 2).

Effects for Perceived Stigma

The baseline preassessment perceived stigma score was significantly higher for the MIDonline group than the control group (see Table 2). An ANCOVA on postassessment ($F_{1,178}$ = .60, $P = .44$) and follow-up ($F_{1,176}$ = 1.06, $P = .30$) perceived stigma yielded no significant differences between the intervention and control group on this measure. However, an ANCOVA on the follow-up perceived stigma assessment score adjusting for the postassessment perceived stigma score showed group differences with those in the MIDonline intervention showing a small increase in perceived stigma scores relative to control in the period between postintervention and follow-up ($F_{1,176}$ = 4.91, $P < .03$).

Effects for Level of Depression

There was no significant difference between the MIDonline and control groups for baseline level of depression (see Table 2). ANCOVAs on the postintervention ($F_{1,178}$ = .03, P = .86) and follow-up ($F_{(1,175)}$ = 1.71, P = .19) personal stigma scores showed no differences on level of depression between groups.

Effect Sizes

The pre-post Cohen's d effect sizes for depression literacy were −1.78 (MIDonline) and −0.07 (control) respectively. The corresponding effect sizes for personal stigma were 0.83 (MIDonline) and 0.06 (control) and perceived stigma −0.14 (MIDonline) and 0.16 (control). The effect sizes for level of depression were 0.10 (MIDonline) and 0.10 (control).

Discussion

This study demonstrated that there were significant differences between the intervention group and the control group for depression literacy and personal stigma scores but not for perceived stigma or level of depression scores.

The finding that an educational intervention can increase depression literacy and reduce personal stigma related to depression has been reported previously [6]. However, to our knowledge, this is the first study to demonstrate such benefits among non–English-speaking immigrants to an English-speaking country. Strategies in the MIDonline website that may have contributed to the stigma reduction effect include the reinforcement of the message that depression is treatable and details of effective treatments, the use of culturally appropriate case studies, listings of bilingual mental health professionals, and information about culturally appropriate self-help strategies.

The identification of an intervention that will decrease personal stigma in this group is important given that the baseline levels of personal stigma for the two immigrant groups were higher than have previously been documented for the general, predominantly Anglo-Australian, adult population in previous research [6] and that previous research has found that personal depression stigma is higher in ethnic populations than English-speaking populations [6,16,17].

The finding that the depression Internet intervention did not decrease perceived stigma levels in Greek-born and Italian-born immigrants is consistent with a previous Internet intervention study that reported similar outcomes for an English-speaking community sample with elevated depressive symptoms [6]. The MIDonline website emphasizes that emotions can be changed by changing thoughts and behaviors, which may have led participants to perceive that others believe that depression is controllable by the depressed person and, hence, the fault of the depressed person. In addition, MIDonline described the history and current widespread stigmatizing beliefs about mental illness among Greek-born and Italian-born communities. This information may have reinforced the belief that stigmatizing attitudes toward mental illness are still widely and strongly held in the wider Greek-born and Italian-born immigrant communities. These findings suggest that no single approach is likely to reduce all aspects of mental health–related stigma [17].

The finding that the change in perceived stigma between postintervention and follow-up differed for the MIDonline and control groups is more difficult to interpret. Inspection of the scores suggests that the effect was primarily due to a reduction in perceived stigma among the control participants. It is conceivable that contact with interviewers among the control group led to a delayed reduction in the perception of stigma among that group, whereas, as noted above, the MIDonline content reinforced participants' perceptions of stigma in others. Alternatively, the finding may be a spurious result attributable to statistical error associated with multiple comparisons. Further research is needed to investigate this finding.

Low literacy has previously been shown to be a barrier to finding health information for adults from an ethnic background, particularly on the Internet and independent of Internet and computer literacy [18]. Many people from a non–English-speaking ethnic background worldwide lack access to mental health–related information that meets their linguistic, cultural, or literacy needs [19]. The current study is a step toward addressing these issues by matching text readability more closely to the literacy levels of our target audience by using language targeted to school grade level 7 or 8.

Viewing the MIDonline website did not result in any changes to level of depression. This contrasts with findings from a systematic review that found that psychoeducational interventions were associated with reduced depressive symptoms relative to control [20]. There are two possible explanations for this discrepancy. First, participants in the current study were not selected on the basis of depressive symptomatology or depression, and the average baseline symptom scores fell well below the BDI-II cutoff score of 14 for mild depression. More significantly, however, the 2-week time frame used in the BDI-II for assessing presence of symptoms was not appropriate for immediate and 1-week follow-up time frames. Further research is required to determine if informational interventions can reduce depressive symptoms among those with elevated depressive symptoms in this population.

Limitations

The study has some limitations. The samples may not have been representative of the broader Greek-born and Italian-born ethnic communities. Both samples were obtained by advertising the study in ethno-specific welfare and social groups and included individuals who volunteered to participate and who were willing to undertake an Internet intervention to increase their knowledge of depression. The fact that this methodology excluded individuals who do not frequent such clubs and organizations may limit the generalizability of the findings.

As part of this study, instrumentation was translated and, where appropriate, modified for the purpose of examining depression literacy and stigma in Greek-born and Italian-born people. The translated instruments may have been measuring different concepts than those measured by the English version of the scale.

Previous studies have found that socially desirable responding influences symptom reporting in non–English-speaking populations [21-23]. Such effects may occur for stigma measures. Future work is required to evaluate the effect of social desirability on the findings. Another limitation of the current study is that the same interviewer was used for the assessment and intervention phases of the study providing a potential source of bias. Similarly, the involvement of the first author in the randomization process had the potential to introduce allocation bias.

A final limitation of the study is that the posttesting and follow-up testing occurred soon after completion of the intervention. Longer follow-up periods are required to examine the sustainability of effects. In addition, an interviewer was present during the MIDonline intervention to assist participants in the navigation of the consumer section of the website. Hence, the current intervention involved some, albeit minimal, face-to-face interviewer guidance, and it is not known if automated distal delivery will yield similar findings in this population.

Future Research

The current study demonstrated the effectiveness of a culturally appropriate depression information intervention in increasing depression literacy and reducing personal stigma among Greek and Italian immigrant populations, who have been shown previously to hold strongly stigmatizing attitudes toward mental illness. Future work is required to examine whether such interventions are effective in reducing stigma among other non–English-speaking immigrant populations and in those with a current or past history of depression or their family members. Future studies should also investigate the information technology literacy and health-related Internet use among the older, non–English-speaking population.

In the current study, the magnitude of effect sizes for the Web-based intervention are very large, and in a public health context, these effect sizes could translate into large gains for large numbers of people at a low cost [24]. The intervention warrants further investigation as a method for delivering depression literacy and stigma reduction programs for depression in other non–English immigrant groups, particularly in an automated format without guidance.

Acknowledgments

The current study was funded by a major research grant by Beyondblue: the National Depression Initiative. There has been no involvement by the funding body in the review or approval of the manuscript for publication.

Conflicts of Interest

None declared

References

1. Goldney RD, Fisher LJ, Wilson DH, Cheok F. Mental health literacy of those with major depression and suicidal ideation: an impediment to help seeking. Suicide Life Threat Behav 2002;32(4):394-403. [Medline: 12501964]
2. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust 1997 Feb 17;166(4):182-186 [FREE Full text] [Medline: 9066546]
3. Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. Br J Psychiatry 2000 Nov;177:396-401 [FREE Full text] [Medline: 11059991]
4. Kiropoulos LA, Klimidis S, Minas H. Depression and anxiety: a comparison of older-aged Greek-born immigrants and Anglo-Australians. Aust N Z J Psychiatry 2004 Sep;38(9):714-724. [doi: 10.1111/j.1440-1614.2004.01445.x] [Medline: 15324336]
5. Australian Bureau of Statistics. 2007. Census of Population and Housing, 2006 URL: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Census+data [accessed 2011-04-06] [WebCite Cache ID 5xjzWBAFK]
6. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185:342-349 [FREE Full text] [doi: 10.1192/bjp.185.4.342] [Medline: 15458995]
7. 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]
8. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. J Med Internet Res 2010;12(5):e48 [FREE Full text] [doi: 10.2196/jmir.1370] [Medline: 21169164]
9. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010;12(5):e53 [FREE Full text] [doi: 10.2196/jmir.1436] [Medline: 21169166]
10. Finkelstein J, Lapshin O, Wasserman E. Randomized study of different anti-stigma media. Patient Educ Couns 2008 May;71(2):204-214. [doi: 10.1016/j.pec.2008.01.002] [Medline: 18289823]

11. Stuart GW, Klimidis S, Minas IH. The treated prevalence of mental disorder amongst immigrants and the Australian-born: community and primary-care rates. Int J Soc Psychiatry 1998;44(1):22-34. [Medline: 9574849]
12. MIDonline. 2006. Multicultural Information on Depression URL: http://www.midonline.com.au/ [accessed 2010-01-12] [WebCite Cache ID 5mk35f4ti]
13. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869. [Medline: 20332511]
14. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. 2nd edition. San Antonio: Psychological Corporation; 1996.
15. Cohen J. Statistical Power Analysis for the Behavioural Sciences. 2nd edition. Hillsdale, NJ: Lawrence Earlbaum; 1988.
16. Raguram R, Weiss MG, Channabasavanna SM, Devins GM. Stigma, depression, and somatization in South India. Am J Psychiatry 1996 Aug;153(8):1043-1049. [Medline: 8678173]
17. Griffiths KM, Christensen H, Jorm AF. Predictors of depression stigma. BMC Psychiatry 2008;8:25 [FREE Full text] [doi: 10.1186/1471-244X-8-25] [Medline: 18423003]
18. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]
19. Neuhauser L, Kreps GL. Online cancer communication: meeting the literacy, cultural and linguistic needs of diverse audiences. Patient Educ Couns 2008 Jun;71(3):365-377. [doi: 10.1016/j.pec.2008.02.015] [Medline: 18424046]
20. Donker T, Griffiths KM, Cuijpers P, Christensen H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. BMC Med 2009;7:79 [FREE Full text] [doi: 10.1186/1741-7015-7-79] [Medline: 20015347]
21. Cappeliez P. Social desirability response set and self-report depression inventories in the elderly. Clinical Gerontologist 1990;9(2):45-52.
22. Linden W, Paulhus DL, Dobson KS. Effects of response styles on the report of psychological and somatic distress. J Consult Clin Psychol 1986 Jun;54(3):309-313. [Medline: 3722558]
23. Phillips DL, Clancy KJ. Response biases in field studies of mental illness. Am Sociol Rev 1970 Jun;35(3):503-515. [Medline: 5529672]
24. Jorm AF, Griffiths KM, Christensen H, Korten AE, Parslow RA, Rodgers B. Providing information about the effectiveness of treatment options to depressed people in the community: a randomized controlled trial of effects on mental health literacy, help-seeking and symptoms. Psychol Med 2003 Aug;33(6):1071-1079. [Medline: 12946091]

Abbreviations

ANOCOVA: analysis of covariance
ANOVA: analysis of variance
BDI–II: Beck Depression Inventory–II
CONSORT: consolidated standards of reporting trials
MIDonline: Multicultural Information on Depression online

Edited by G Eysenbach; submitted 10.03.10; peer-reviewed by J Fogel; comments to author 02.04.10; revised version received 15.02.11; accepted 21.03.11; published 19.04.11.

Please cite as:
Kiropoulos LA, Griffiths KM, Blashki G
Effects of a Multilingual Information Website Intervention on the Levels of Depression Literacy and Depression-Related Stigma in Greek-Born and Italian-Born Immigrants Living in Australia: A Randomized Controlled Trial
J Med Internet Res 2011;13(2):e34
URL: http://www.jmir.org/2011/2/e34/
doi:10.2196/jmir.1527
PMID:21504872

Original Paper

A 5-Year Follow-up of Internet-Based Cognitive Behavior Therapy for Social Anxiety Disorder

Erik Hedman[1], MSc; Tomas Furmark[2], PhD; Per Carlbring[3], PhD; Brjánn Ljótsson[1], MSc; Christian Rück[1], MD PhD; Nils Lindefors[1], MD PhD; Gerhard Andersson[1,4,5], PhD

[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2]Department of Psychology, Uppsala University, Uppsala, Sweden
[3]Department of Psychology, Umeå University, Umeå, Sweden
[4]Swedish Institute for Disability Research, Linköping University, Linköping, Sweden
[5]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Corresponding Author:
Erik Hedman, MSc
Department of Clinical Neuroscience
Karolinska Institutet
Internet Psychiatry M46
Karolinska University Hospital Huddinge
Stockholm, 141 86
Sweden
Phone: 46 8 585 857 91
Fax: 46 8 779 54 16
Email: erik.hedman.2@ki.se

Abstract

Background: Internet-based cognitive behavior therapy (CBT) has been shown to be a promising method to disseminate cognitive behavior therapy for social anxiety disorder (SAD). Several trials have demonstrated that Internet-based CBT can be effective for SAD in the shorter term. However, the long-term effects of Internet-based CBT for SAD are less well known.

Objective: Our objective was to investigate the effect of Internet-based CBT for SAD 5 years after completed treatment.

Method: We conducted a 5-year follow-up study of 80 persons with SAD who had undergone Internet-based CBT. The assessment comprised a diagnostic interview and self-report questionnaires. The main outcome measure was the Liebowitz Social Anxiety Scale-Self-Report (LSAS-SR). Additional measures of social anxiety were the Social Interaction Anxiety Scale (SIAS) and the Social Phobia Scale (SPS). Attrition rates were low: 89% (71/80) of the participants completed the diagnostic interview and 80% (64/80) responded to the questionnaires.

Results: Mixed-effect models analysis showed a significant effect of time on the three social anxiety measures, LSAS-SR, SIAS, and SPS ($F_{3,98-102} = 16.05 - 29.20$, $P < .001$) indicating improvement. From baseline to 5-year follow-up, participants' mean scores on the LSAS-SR were reduced from 71.3 (95% confidence interval [CI] 66.1-76.5) to 40.3 (95% CI 35.2 - 45.3). The effect sizes of the LSAS-SR were large (Cohen's *d* range 1.30 - 1.40, 95% CI 0.77 - 1.90). Improvements gained at the 1-year follow-up were sustained 5 years after completed treatment.

Conclusions: Internet-based CBT for SAD is a treatment that can result in large and enduring effects.

Trial registration: Clinicaltrials.gov NCT01145690; http://clinicaltrials.gov/ct2/show/NCT01145690 (Archived by WebCite at http://www.webcitation.org/5ygRxDLfK)

(J Med Internet Res 2011;13(2):e39) doi:10.2196/jmir.1776

KEYWORDS

Internet; cognitive behavior therapy; anxiety disorders; social anxiety disorder; 5-year follow-up

Introduction

Social anxiety disorder (SAD) is common [1], is associated with functional impairment [2], and often becomes chronic if left untreated [3]. In recent years, Internet-based cognitive behavior therapy (CBT) has demonstrated efficacy in several randomized controlled trials [4-9]. In general, effect sizes on measures of social anxiety in these studies have been at parity with those seen in trials investigating conventional CBT (Cohen's *d* typically ranging from 1.0-1.5) [10,11]. In essence, Internet-based CBT could be described as Internet-administered self-help therapy with online therapist contact and support. The treatment components and theoretical basis are the same as in conventional CBT. While several studies have shown that conventional CBT produces long-term improvements up to 5 years after treatment [12-15], nearly all studies on Internet-based CBT have had a follow-up period of 1 year or shorter. The one exception is a study where participants receiving Internet-based CBT not only maintained their treatment gains but also were further improved at a 2.5-year follow-up [16]. This is in line with the notion that reduced anxiety following CBT to a large extent is contingent on repeated exposure [17].

The aim of the present study was to investigate the effects of Internet-based CBT for SAD 5 years after treatment, as no previous study has investigated if the effect of Internet-based CBT persists over this long period of time. We hypothesized that treatment gains would be sustained on measures of social anxiety, depressive symptoms, general anxiety, and quality of life.

Methods

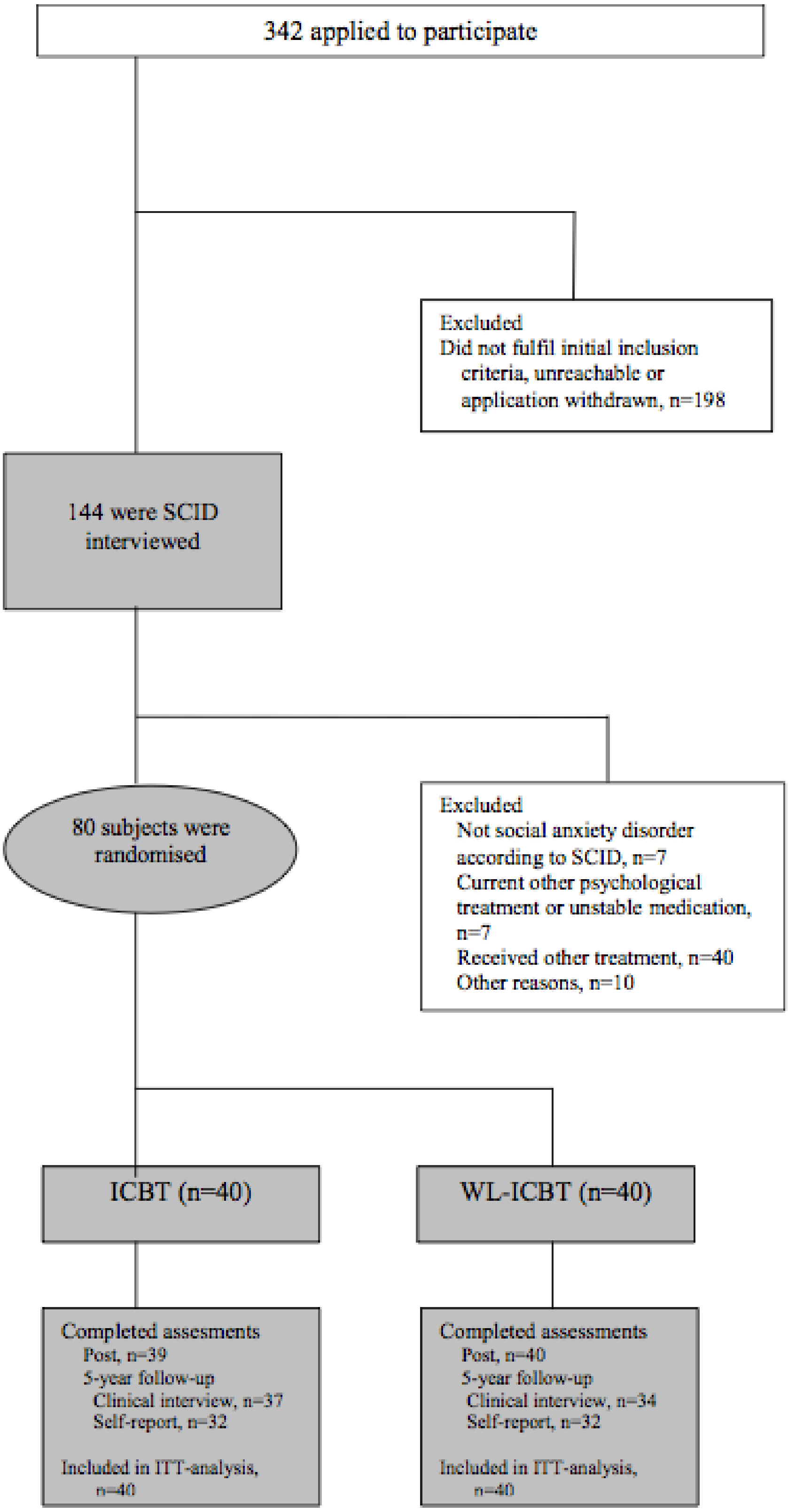
Design

This was a follow-up study assessing 80 participants who had received Internet-based CBT for SAD within the context of a randomized controlled trial (RCT) conducted in 2005. In the original RCT, participants were randomized to treatment (n = 40) or waiting list control (n = 40) with equal probability. Participants were randomized using a true random number service (http://www.random.org). Participants were randomized after inclusion in the study, ensuring that allocation status was unknown to the assessors deciding on inclusion. Following treatment and postassessment, participants in the waiting list control group were crossed over to treatment. Thus, both groups had received Internet-based CBT at 1-year follow-up. As the two groups received treatment at different time points, results are reported separately for the two groups. CBT denotes the first group, and waiting list (WL)-CBT, the latter. A detailed description of the original study is available elsewhere [6]. The trial was registered at clinicaltrials.gov (identifier NCT01145690).

Sample and Recruitment

All participants included in the original RCT were eligible to participate in this follow-up study. The main inclusion criteria were the following: participants had to have a primary diagnosis of SAD according to the Structural Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) Axis-I Disorders [18]; participants had to agree to undergo no other psychological treatment throughout the original study and keep dosage constant if on prescribed medication for anxiety or depression; and participants had to be at least 18 years old. Main exclusion criteria were not having a computer with Internet access and admitting to another serious disorder (eg, schizophrenia or substance dependence). On average, participants were 35.3 (SD 10.5) years old, and the sample comprised 70% women. Participants in the original RCT were enrolled from January 2005 through March 2005, and recruitment tool place in Uppsala, Sweden. The flow of participants throughout the study is presented in Figure 1. The follow-up study was approved by the regional ethics review board in Stockholm, Sweden, and informed consent was obtained from all participants.

Figure 1. Participant flow



Outcome Measures

The primary outcome measure was the Liebowitz Social Anxiety Scale-Self-Report (LSAS-SR) [19]. The LSAS-SR measures fear in and avoidance of 24 social situations (13 performance and 11 interaction situations) that are usually difficult for people suffering from SAD. Fear and avoidance in each situation is rated on a 4-point scale from 0 to 3. We also used the Social Interaction Anxiety Scale (SIAS) [20], the Social Phobia Scale (SPS) [20], and the Social Phobia Screening Questionnaire (SPSQ) [1] as complementary measures of social anxiety. The SPS assesses anxiety in 20 performance situations, while the

SIAS is constructed to measure anxiety in 20 social interaction situations. Each situation is rated on a 5-point scale ranging from 0 to 5. The SPSQ, designed to screen for SAD using DSMV-IV criteria, was used solely as a dichotomous indicator of SAD diagnosis.

In addition, the Montgomery-Åsberg Depression Rating Scale-Self-report (MADRS-S) [21] and the Beck Anxiety Inventory (BAI) [22] were used as secondary measures to assess depressive symptoms and general anxiety, respectively. MADRS-S comprises 9 items measuring different aspects of depressive symptoms, and each symptom is rated on a 7-point scale. The BAI assesses 21 anxiety symptoms on a 4-point scale from 0 to 3. The Quality of Life Inventory (QOLI) [23] was also administered as a secondary generic outcome measure. The QOLI measures quality of life in16 different domains (eg, work and family). For each domain, the respondent is asked to rate importance on a 3-point scale (from 0 to 2) and the degree of satisfaction on a 6-point scale (from −3 to +3). By multiplying importance by satisfaction, each domain yields a value from −6 to +6.

All measures described above have demonstrated good psychometric properties.

Clinical Assessment Interview

The SCID-I [18] was used to establish whether participants met diagnostic criteria for SAD at 5-year follow-up. Global improvement was measured by the Clinical Global Impression Improvement Scale (CGI-I) [24]. In addition, information about current and earlier psychological and pharmacological treatments was obtained. Finally, participants were asked to rate to what extent they attributed their improvement/current state to Internet-based CBT.

Treatment

The Internet-based CBT used in this study has been found efficacious in several randomized controlled trials [4,5,25]. The treatment followed a CBT model that stresses the importance of avoidance and safety behaviors as maintaining factors of SAD [26]. The most central feature of the treatment was a self-help text comprising 9 text modules delivered via the Internet, each covering a specific theme (eg, exposure and cognitive restructuring) including homework exercises.

The introductory module described basic features of SAD and facts about CBT. The topics of modules 2 to 4 were primarily the social anxiety model as presented by Clark and Wells, as well as cognitive restructuring. Modules 5 to 7 introduce safety behavior experiments, exposure exercises, and attention training. Modules 8 and 9 had a main focus on social skills and relapse prevention. The general treatment procedure was that participants read the self-help text, carried out the home work assignments, and reported to their therapist through an online message system.

Throughout the trial, all participants had access to a therapist who supervised the progress and gave feedback on homework exercises. All therapists were clinical psychologists in training during the last semester of their 5-year educational programme. In addition, participants had access to an online discussion forum where they could communicate anonymously with each other. The duration of the treatment was 9 weeks.

Procedure

The clinical assessment interview was performed by a clinical psychologist with more than 5 years experience in working with structured diagnostic assessments. The interview was conducted by telephone, which has been shown to be a reliable way of assessing psychiatric symptoms [27,28]. The LSAS-SR, SIAS, SPS, SPSQ, MADRS-S, BAI, and QOLI were administered via the Internet, a valid administration format for these instruments [29].

Statistical Analysis

Statistical analyses were conducted using PASW version 18.0 (SPSS inc, Chicago, IL). While data were analyzed on intent-to-treat basis, we did not apply last observation carried forward (LOCF) to handle missing data as that might have exaggerated the degree to which gains were sustained. Instead, we report the observed means and standard deviations as well as estimated means and standard deviations, as suggested by Gueorguieva and Krystal [30]. Estimated parameters were obtained using a mixed-models approach employing a first order autoregressive covariance structure. The following formula was used for converting standard errors to standard deviations: SD = SE (√n). As all participants received Internet-based CBT, the main analyses entailed no between-group comparisons. However, as half of the sample served as controls in the first phase of the RCT, the two groups are reported separately. We conducted mixed-effect models analysis to assess improvement over time on continuous outcome variables. Nominal data were analyzed with McNemar’s test of change. Effect sizes (Cohen’s *d*) were calculated using the observed means and pooled SDs.

Results

Attrition

Of 80 participants, 71 (89%) underwent the clinical assessment interview and 64 (80%) completed the LSAS-SR, SIAS, SPS, MADRS-S, BAI, and QOLI. There were no statistically significant differences between participants who did not provide follow-up data and those who did regarding gender ($c^2_1 = 0.39$, $P = .39$), age, and social anxiety at baseline or at 1-year follow-up ($t_{1,67\text{-}78} = 0.40 - 1.74$, $P = .68 - .09$). The reasons for not completing the 5-year follow-up are unknown.

Social Anxiety Measures

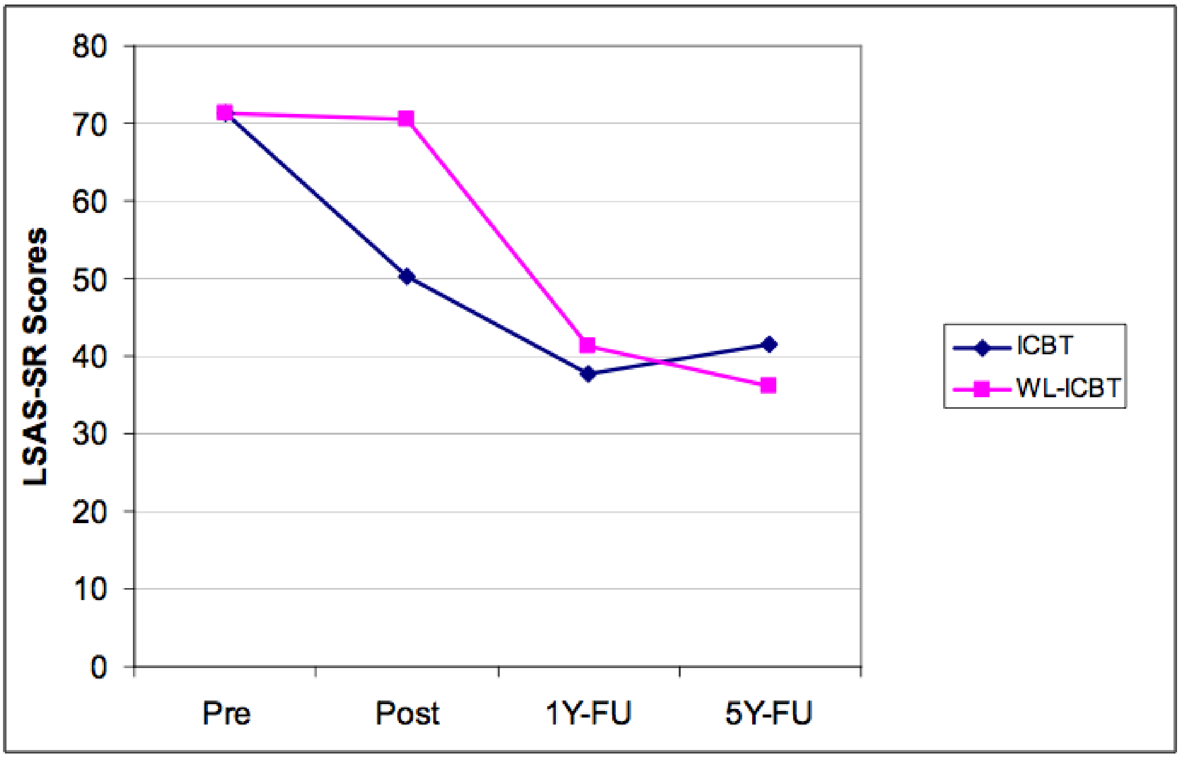
The observed and estimated means and SDs as well as effect sizes of the continuous outcome measures are presented in Table 1. Mixed-effect models analysis showed a significant effect of time on the primary outcome measure LSAS-SR, as well as on the SIAS and SPS ($F_{3,98\text{-}102} = 16.05 - 29.20$, $P < .001$). Pairwise comparisons showed that participants in both groups were significantly improved from baseline to 1- and 5-year follow-up on all social anxiety measures ($F_{1,33\text{-}38} = 15.10 - 90.05$, $P < .001$). The CBT group was further improved at 1-year follow-up compared with postassessment, and the WL-CBT group were also improved during this period ($F_{1,34\text{-}35} = 7.43 - 40.42$, $P =$

.01 - .001). There were no significant changes on the LSAS and SPS between 1- and 5-year follow-up ($F_{1,28,32}$ = 0.22, 0.93, P = .64 - .13). In the WL-CBT group but not in the CBT group, participants were further improved on the SIAS at 5-year follow-up compared with 1-year follow-up ($F_{1,29}$ = 7.85 P = .01). Figure 2 displays changes on the primary outcome measure LSAS-SR across assessment points. Note that as we used LOCF to handle missing data in the original article, there are minimal and nonsignificant discrepancies in the present report compared with the original regarding parameters at postassessment and 1-year follow-up.

Table 1. Observed and estimated means, SDs, and effect sizes (Cohen's *d*) on continuous outcome measures

Measure and Group n = 40 (CBT and WL-CBT)	Pre M (SD)	Post M (SD)	1-Year Follow-up M (SD)	Observed 5-year Follow-up M (SD)	Estimated 5-Year Follow-up M (SD)	Effect Size Within Pre 1-Year Follow-up (95%CI)	Effect Size Within Pre 5-Year Follow-up (95%CI)
LSAS-SR							
CBT	71.3 (22.5)	50.3 (21.0)	37.7 (17.7)	41.5 (23.7)	41.6 (20.9)	1.65 (1.11–2.15)	1.30 (0.77–1.79)
WL-CBT	71.3 (24.9)	70.4 (27.6)	41.3 (29.0)	36.3 (25.3)	38.9 (24.9)	1.12 (0.61-1.60)	1.40 (0.86-1.90)
SIAS							
CBT	51.0 (14.2)	38.5 (13.9)	32.8 (14.9)	36.3 (16.8)	36.1 (14.7)	1.25 (0.73-1.75)	0.95 (0.45-1.43)
WL-CBT	46.5 (17.9)	46.4 (18.7)	31.7 (18.3)	24.6 (14.7)	25.9 (15.8)	0.81 (0.34-1.27)	1.32 (0.79-1.82)
SPS							
CBT	39.2 (15.3)	25.2 (12.0)	19.0 (12.0)	22.6 (18.4)	22.6 (14.7)	1.46 (0.94-1.95)	0.98 (0.48-1.47)
WL-CBT	36.4 (17.1)	35.7 (16.4)	20.0 (14.7)	16.6 (16.4)	17.5 (15.3)	1.02 (0.53-1.49)	1.18 (0.66-1.67)
MADRS-S							
CBT	14.9 (7.8)	10.4 (6.3)	9.7 (7.0)	9.6 (7.8)	10.5 (7.9)	0.70 (0.23-1.16)	0.68 (0.20-1.15)
WL-CBT	15.7 (9.3)	16.3 (10.2)	10.9 (8.5)	7.7 (8.9)	8.3 (9.0)	0.54 (0.08-0.99)	0.88 (0.39-1.36)
BAI							
CBT	16.1 (7.4)	9.8 (5.8)	10.4 (7.2)	10.6 (10.4)	10.5 (7.9)	0.87 (0.39-1.33)	0.63 (0.15-1.10)
WL-CBT	16.2 (9.6)	15.3 (9.4)	11.8 (9.2)	8.3 (9.8)	8.7 (9.0)	0.68 (0.21-1.14)	0.81 (9.32-1.28)
QOLI							
CBT	0.8 (1.9)	1.3 (2.0)	1.7 (1.5)	1.9 (1.7)	1.7 (1.7)	0.55 (0.09-1.00)	0.63 (0.15-1.10)
WL-CBT	0.6 (1.9)	0.4 (1.6)	1.4 (1.8)	2.1 (1.8)	1.9 (1.7)	0.41 (−0.06 to 0.86)	0.77 (0.28-1.25)

Figure 2. Improvement course on the primary outcome measure LSAS-SR during the follow-up period



Abbreviations: ICBT, Internet-based Cognitive Behaviour Therapy; WL-ICBT, Waiting list followed by Internet-based Cognitive Behaviour Therapy; Pre, before treatment; Post, post-treatment; 1Y-FU, one year after treatment; 5Y-FU, five years after treatment; LSAS-SR, Liebowitz Social Anxiety Scale- Self-Report.

Depressive Symptoms, General Anxiety, and Quality of Life

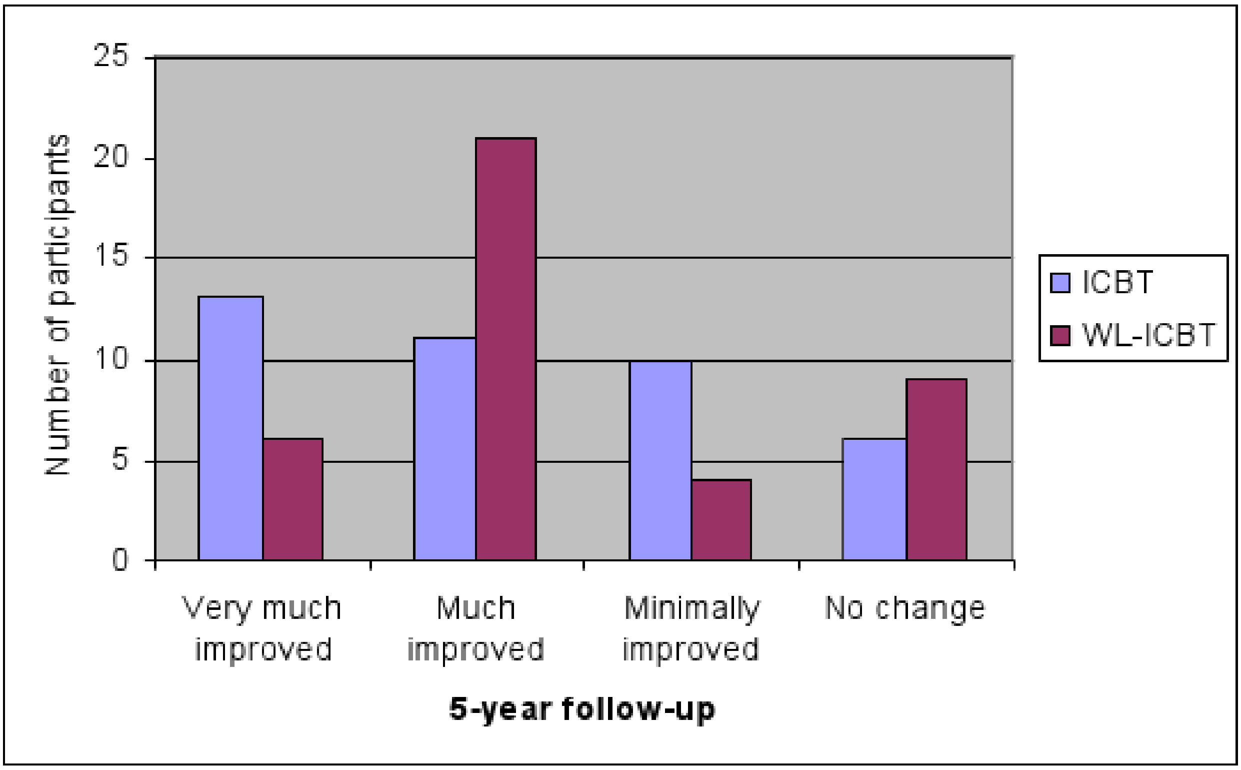
Effect sizes and observed and estimated parameters of secondary outcome measures are presented in Table 1. Mixed-effect models analysis showed a significant effect of time on the MADRS-S, BAI, and QOLI ($F_{3,97-104}$ = 4.64 - 9.78, P = .01 - .001). Pairwise comparisons showed that participants in both groups were significantly improved from baseline to 1- and 5-year follow-up on MADRS-S, BAI, and QOLI ($F_{1,32-40}$ = 4.7 - 30, P = .04 - .001). The WL-CBT was improved at 1-year follow-up compared with postassessment on these measures ($F_{1,34,35}$ = 12.12 - 13.83, $P < .001$), whereas the CBT group was not ($F_{1,35-37}$ = 0.36 - 3.09, P = .55 - .09). There were no changes on these measures from 1- to 5-year follow-up ($F_{1,28,33}$ = 0.01 - 3.80, P = .94 - .06).

Clinical Assessment Interview

Global Improvement and Diagnostic Assessment

Figure 3 displays CGI-I scores at 5-year follow-up for both groups. At this time, 60% of participants (24/40) in the CBT group and 67.5% (27/40) in the WL-CBT group were considered very much or much improved, that is, responders. At 5-year follow-up, 48% of participants (19/40) in both groups no longer met diagnostic criteria for SAD according to the clinician assessment (counting dropouts as nonresponders). McNemar's test showed that this was a statistically significant change compared with baseline ($P < .001$). According to the SPSQ, 40% (16/40) of the participants in the CBT group and 45% (18/40 in the WL-CBT group no longer met criteria for SAD (counting dropouts as nonresponders).

Figure 3. Clinical Global Impression Improvement (CGI-I) scores at 5-year follow-up (dropouts are considered non-responders)



Participants' Attribution of Improvement

Participants were asked to rate to what extent they attributed their improvement to the Internet-based CBT on a Likert-scale from 0 to 100 (0 = any improvement is completely unrelated to Internet-based CBT, 50 = any improvement is equally due to Internet-based CBT and other causes, and 100 = any improvement is completely due to Internet-based CBT). In the CBT group, the average score was 60.3 (SD 26.9) and the corresponding WL-CBT score was 61.8 (SD 25.9).

Other Psychological and Psychotropic Treatments Received Since Internet-based CBT

At 5-year follow-up, 10% (4/40) participants in the CBT group had received some form of psychological treatment (all reasons included) after Internet-based CBT. This was 11% (4/37) if counting completers only, that is, those who provided data. The corresponding percent in the WL-CBT + WL group was 17.5% (7/40). This was 21% (7/34) if counting completers only. In the CBT group, 1 of the 40 participants (2.5%), or 1 of 37 (2.7%) if counting completers only, was taking psychotropic medication, that is, selective serotonin reuptake inhibitors (SSRIs) at the time of the 5-year follow-up assessment, although 4 of 40 participants (10%), or 4 of 37 (11%) if counting completers only, had started and discontinued psychotropic medication at some point during the follow-up period (all SSRIs). In the WL-CBT group, the corresponding numbers were 3 of 40 (7.5%), or 3 of 37 (8%) if counting completers only, and 5 of 40 (12.5%), or 5 of 34 (15%) if counting completers only, respectively (all had been taking SSRIs). The status of the 11% (9/80) dropouts regarding medication is unknown.

Discussion

Main Findings

The aim of this study was to evaluate the 5-year effect of Internet-based CBT for SAD by assessing participants who had received Internet-based CBT within the context of an RCT. The results showed that improvements on measures of social anxiety at 1-year follow-up were sustained 5 years after treatment. Overall, effect sizes were large on measures of social anxiety. In addition, improvements regarding depressive symptoms, general anxiety, and quality of life were also sustained at 5-year follow-up. The results of this study indicate that participants receiving Internet-based CBT for SAD are moderately improved immediately following treatment but make further improvements within the following year. Improvements made at 1-year follow-up are, in turn, long-term enduring.

The effect sizes in this study are in line with those reported in studies investigating the long-term effects of conventional CBT for SAD [13,31]. They are also in line with results from a previous independent 2.5-year follow-up study of Internet-based CBT for SAD [16]. The major strength of this study is that attrition rates were low making the generalizability of the findings high. The low attrition rates were also reflected in the small differences in the observed and estimated estimates. Furthermore, participants attributed their improvement to Internet-based CBT to a large extent, and few had commenced other forms of psychological or psychotropic treatments after completing Internet-based CBT. Taken together, this suggests that the reduction of social anxiety observed at 5-year follow-up was largely an effect of Internet-based CBT.

Clinical Implications

There are several clinical implications of our findings. First, if Internet-based CBT for SAD has sustained effects over longer time periods, it is highly likely that it is a cost-effective treatment. We did not collect economic data in this study; however, results of a study by Titov and coworkers have demonstrated that Internet-based CBT is likely more cost-effective than group CBT due to lower costs of treatment [32]. Second, it may also be that Internet-based CBT confers benefits in another way compared with conventional therapies, since the material can be saved and used as reminders long after the treatment has ended. The effect sizes found in the present study, which are in parity with those found in trials investigating conventional CBT, suggest that Internet-based CBT has some qualities that compensate for the lack of face-to-face contact. Intriguingly, in the original trial [6], a basic patient satisfaction rating showed that 94% of the participants were satisfied with the treatment and that 91% of the participants found the feedback from the therapists to be good or excellent. This suggests that it is possible to have a good therapeutic relationship online, which has also been reported in other studies on Internet-based CBT [33].

Third, Internet-based CBT may in the future be used as a complement to conventional CBT and pharmacotherapy, as it probably can be combined with these two treatments. Internet-based CBT might enable more efficient use of health care resources, that is, as Internet-based CBT requires less therapist time, more resources can be made available for patients who need a more intensified treatment. This, in turn, could lead to a larger total proportion of treatment responders. Future research should more clearly link symptom improvement to the treatment provided and the extent to which strategies learned in treatment are used to prevent recurrence.

Limitations

The present study has several limitations, and we view the following as most important. First, common to most long-term follow-up trials, there was no randomization to a control condition with which treatment results could be compared at 5-year follow-up. However, considering the chronicity of SAD [3], we find it unlikely that improvements are due to spontaneous recovery. Furthermore, it is improbable that nonspecific treatment effects such as attention from a therapist would generate improvements that are enduring over 5 years. Second, we did not use a behavioral test to assess social anxiety, which would have been a more objective measure than the ones used. For example, Heimberg and coworkers used a test where participants were exposed to personally tailored social situations while using heart rate monitoring equipment to assess bodily symptoms of anxiety [34]. Nonetheless, we view the combination of clinician assessment and administration of questionnaires with good psychometric properties as a valid assessment method. Third, the intervals between the follow-ups were not regular, and it is not possible to infer symptom levels between the follow-ups (eg, 3 years posttreatment). As clinical assessment interviews were only conducted at pretreatment and 5-year follow-up, this uncertainty also applies to diagnostic status. However, as symptoms of SAD are not known to fluctuate spontaneously, we find it unlikely that levels of social anxiety in the present sample varied greatly between 1- and 5-year follow-up. Finally, although attrition rates were low, 11% of the participants did not attend the assessment interview. Of course, it might be that these individuals are less improved than those who participated in the 5-year follow-up assessment. However, even if those participants were nonresponders, it would have only a marginal effect on the effect size estimates. We also view the types of analyses performed, where models were created using all available data, yielded the best estimate, as last observation carried forward could have overestimated the long-term effect.

In spite of these limitations, we regard the results of the present study as important as they are the first to demonstrate that Internet-based CBT for SAD can yield large effects that are enduring over 5 years.

Acknowledgments

The funding organization (Stockholm County Council) is a public institution that had no role in the design and conduct of the study; in the collection, management, and analysis of the data; or in the preparation, review and approval of the manuscript.

Conflicts of Interest

None declared

Authors' Contributions

Author EH designed the study, performed the analyses, and drafted the paper. Authors TF, PC, BL, CR, NL, and GA designed the study, supervised the scientific work, and drafted the paper. EH had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the analysis.

References

1. Furmark T, Tillfors M, Everz P, Marteinsdottir I, Gefvert O, Fredrikson M. Social phobia in the general population: prevalence and sociodemographic profile. Soc Psychiatry Psychiatr Epidemiol 1999 Aug;34(8):416-424. [Medline: 10501711]
2. Schneier FR, Heckelman LR, Garfinkel R, Campeas R, Fallon BA, Gitow A, et al. Functional impairment in social phobia. J Clin Psychiatry 1994 Aug;55(8):322-331. [Medline: 8071299]

3. Reich J, Goldenberg I, Vasile R, Goisman R, Keller M. A prospective follow-along study of the course of social phobia. Psychiatry Res 1994 Dec;54(3):249-258. [Medline: 7792329]
4. Andersson G, Carlbring P, Holmström A, Sparthan E, Furmark T, Nilsson-Ihrfelt E, et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. J Consult Clin Psychol 2006 Aug;74(4):677-686. [doi: 10.1037/0022-006X.74.4.677] [Medline: 16881775]
5. Carlbring P, Gunnarsdóttir M, Hedensjö L, Andersson G, Ekselius L, Furmark T. Treatment of social phobia: randomised trial of internet-delivered cognitive-behavioural therapy with telephone support. Br J Psychiatry 2007 Feb;190:123-128 [FREE Full text] [doi: 10.1192/bjp.bp.105.020107] [Medline: 17267928]
6. Furmark T, Carlbring P, Hedman E, Sonnenstein A, Clevberger P, Bohman B, et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. Br J Psychiatry 2009 Nov;195(5):440-447 [FREE Full text] [doi: 10.1192/bjp.bp.108.060996] [Medline: 19880935]
7. Titov N, Andrews G, Schwencke G, Drobny J, Einstein D. Shyness 1: distance treatment of social phobia over the Internet. Aust N Z J Psychiatry 2008 Jul;42(7):585-594. [doi: 10.1080/00048670802119762] [Medline: 18612862]
8. Berger T, Hohl E, Caspar F. Internet-based treatment for social phobia: a randomized controlled trial. J Clin Psychol 2009 Oct;65(10):1021-1035. [doi: 10.1002/jclp.20603] [Medline: 19437505]
9. Titov N, Andrews G, Choi I, Schwencke G, Mahoney A. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. Aust N Z J Psychiatry 2008 Dec;42(12):1030-1040. [doi: 10.1080/00048670802512107] [Medline: 19016091]
10. Heimberg RG, Liebowitz MR, Hope DA, Schneier FR, Holt CS, Welkowitz LA, et al. Cognitive behavioral group therapy vs phenelzine therapy for social phobia: 12-week outcome. Arch Gen Psychiatry 1998 Dec;55(12):1133-1141 [FREE Full text] [Medline: 9862558]
11. Clark DM, Ehlers A, McManus F, Hackmann A, Fennell M, Campbell H, et al. Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. J Consult Clin Psychol 2003 Dec;71(6):1058-1067. [doi: 10.1037/0022-006X.71.6.1058] [Medline: 14622081]
12. Barrett PM, Duffy AL, Dadds MR, Rapee RM. Cognitive-behavioral treatment of anxiety disorders in children: long-term (6-year) follow-up. J Consult Clin Psychol 2001 Feb;69(1):135-141. [Medline: 11302272]
13. Heimberg RG, Salzman DG, Holt CS, Blendell KA. Cognitive-behavioral group treatment for social phobia: Effectiveness at five-year followup. Cognitive Ther Res 1993;17:325-339.
14. Saavedra LM, Silverman WK, Morgan-Lopez AA, Kurtines WM. Cognitive behavioral treatment for childhood anxiety disorders: long-term effects on anxiety and secondary disorders in young adulthood. J Child Psychol Psychiatry 2010 Aug;51(8):924-934. [doi: 10.1111/j.1469-7610.2010.02242.x] [Medline: 20345838]
15. Fava GA, Rafanelli C, Grandi S, Canestrari R, Morphy MA. Six-year outcome for cognitive behavioral treatment of residual symptoms in major depression. Am J Psychiatry 1998 Oct;155(10):1443-1445 [FREE Full text] [Medline: 9766780]
16. 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]
17. Michelson L, Mavissakalian M, Marchione K, Dancu C, Greenwald M. The role of self-directed in vivo exposure in cognitive, behavioral, and psychophysiological treatments of agoraphobia. Behav Ther 1986;17:91-108.
18. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002.
19. Fresco DM, Coles ME, Heimberg RG, Liebowitz MR, Hami S, Stein MB, et al. The Liebowitz Social Anxiety Scale: a comparison of the psychometric properties of self-report and clinician-administered formats. Psychol Med 2001 Aug;31(6):1025-1035. [Medline: 11513370]
20. Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther 1998 Apr;36(4):455-470. [Medline: 9670605]
21. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. Acta Psychiatr Scand 1994 Jan;89(1):21-28. [Medline: 8140903]
22. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [Medline: 3204199]
23. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the quality of life inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. Psychological Assessment 1992;4(1):92-101.
24. Guy W. Clinical Global Impressions. Rockville, MD: US Department of Health and Human services; 1976.
25. Tillfors M, Carlbring P, Furmark T, Lewenhaupt S, Spak M, Eriksson A, et al. Treating university students with social phobia and public speaking fears: Internet delivered self-help with or without live group exposure sessions. Depress Anxiety 2008;25(8):708-717. [doi: 10.1002/da.20416] [Medline: 18729147]
26. Clark DM, Wells A. A cognitive model of social phobia. In: Heimberg RG, Leibowitz M, Hope DA, Schneider FR, editors. Social Phobia: Diagnosis, Assessment and Treatment. New York, NY: Guilford Press; 1995:63-93.

27. Crippa JA, de Lima Osório F, Del-Ben CM, Filho AS, da Silva Freitas MC, Loureiro SR. Comparability between telephone and face-to-face structured clinical interview for DSM-IV in assessing social anxiety disorder. Perspect Psychiatr Care 2008 Oct;44(4):241-247. [doi: 10.1111/j.1744-6163.2008.00183.x] [Medline: 18826462]
28. Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. Am J Psychiatry 1997 Nov;154(11):1593-1598. [Medline: 9356570]
29. Hedman E, Ljótsson B, Rück C, Furmark T, Carlbring P, Lindefors N, et al. Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. Computers in Human Behavior 2010;26:736-740. [doi: 10.1016/j.chb.2010.01.010]
30. Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. Arch Gen Psychiatry 2004 Mar;61(3):310-317. [doi: 10.1001/archpsyc.61.3.310] [Medline: 14993119]
31. Hunt C, Andrews G. Long-term outcome of panic disorder and social phobia. J Anxiety Disord 1998;12(4):395-406. [Medline: 9699122]
32. 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]
33. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: a randomized controlled clinical trial. BMC Psychiatry 2007;7:13 [FREE Full text] [doi: 10.1186/1471-244X-7-13] [Medline: 17442125]
34. Heimberg RG, Dodge CS, Hope DA, Kennedy CR, Zollo LJ, Becker R. Cognitive behavioral group treatment for social phobia: Comparison with a credible placebo control. Cognitive Therapy and Research 1990;14(1):1-23. [doi: 10.1007/BF01173521]

Abbreviations

BAI: Beck Anxiety Inventory
CBT: cognitive behavior therapy
CGI-I: Clinical Global Impression Improvement Scale
CI: confidence interval
DSMV-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition
LOCF: last observation carried forward
LSAS-SR: Liebowitz Social Anxiety Scale-Self-Report
MADRS-S: Montgomery-Åsberg Depression Rating Scale-Self-Report
QOLI: Quality of Life Inventory ()
RCT: randomized controlled trial
SAD: social anxiety disorder
SIAS: Social Interaction Anxiety Scale
SPS: Social Phobia Scale
SPSQ: Social Phobia Screening Questionnaire
SSRI: selective serotonin reuptake inhibitor

Edited by G Eysenbach; submitted 18.02.11; peer-reviewed by S Rydh, J Fäldt, Y Khazaal; comments to author 09.03.11; revised version received 13.03.11; accepted 14.03.11; published 15.06.11.

Please cite as:
Hedman E, Furmark T, Carlbring P, Ljótsson B, Rück C, Lindefors N, Andersson G
A 5-Year Follow-up of Internet-Based Cognitive Behavior Therapy for Social Anxiety Disorder
J Med Internet Res 2011;13(2):e39
URL: http://www.jmir.org/2011/2/e39/
doi:10.2196/jmir.1776
PMID:21676694

Original Paper

Effectiveness of E-Self-help Interventions for Curbing Adult Problem Drinking: A Meta-analysis

Heleen Riper[1,2,3], PhD; Viola Spek[3,4], PhD; Brigitte Boon[3], PhD; Barbara Conijn[3], MSc; Jeannet Kramer[3], PhD; Katherina Martin-Abello[3], MA; Filip Smit[3,5], PhD

[1]Department of Clinical Psychology, VU University Amsterdam, Amsterdam, Netherlands
[2]EMGO+ Institute for Health and Health Care Research, VU University Medical Centre, Amsterdam, Netherlands
[3]Trimbos Institute, Netherlands Institute of Mental Health and Addiction, Utrecht, Netherlands
[4]CoRPS Center of Research on Psychology in Somatic Diseases, Tilburg University, Tilburg, Netherlands
[5]Department of Epidemiology and Biostatistics, EMGO+ Institute for Health and Health, VU Unversity Medical Centre, Amsterdam, Netherlands

Corresponding Author:
Heleen Riper, PhD
Department of Clinical Psychology
VU University Amsterdam
Van der Boechorststraat 1
Amsterdam, 1081 BT
Netherlands
Phone: 31 20 5988757
Fax: 31 20 5988758
Email: h.riper@psy.vu.nl

Abstract

Background: Self-help interventions without professional contact to curb adult problem drinking in the community are increasingly being delivered via the Internet.

Objective: The objective of this meta-analysis was to assess the overall effectiveness of these eHealth interventions.

Methods: In all, 9 randomized controlled trials (RCTs), all from high-income countries, with 9 comparison conditions and a total of 1553 participants, were identified, and their combined effectiveness in reducing alcohol consumption was evaluated by means of a meta-analysis.

Results: An overall medium effect size (g = 0.44, 95% CI 0.17-0.71, random effect model) was found for the 9 studies, all of which compared no-contact interventions to control conditions. The medium effect was maintained (g = 0.39; 95% CI 0.23-0.57, random effect model) after exclusion of two outliers. Type of control group, treatment location, type of analysis, and sample size did not have differential impacts on treatment outcome. A significant difference (P = .04) emerged between single-session personalized normative feedback interventions (g = 0.27, 95% CI 0.11-0.43) and more extended e- self-help (g = 0.61, 95% CI 0.33-0.90).

Conclusion: E-self-help interventions without professional contact are effective in curbing adult problem drinking in high-income countries. In view of the easy scalability and low dissemination costs of such interventions, we recommend exploration of whether these could broaden the scope of effective public health interventions in low- and middle-income countries as well.

(J Med Internet Res 2011;13(2):e42) doi:10.2196/jmir.1691

KEYWORDS

Meta-analysis; alcohol; problem drinking; randomized controlled trial; self-help; e-self-help; intervention; unguided self-help; low intensity interventions; Internet; adults

Introduction

The global economic and health burden of alcohol use disorders is widely recognized [1], as is the need for effective public health interventions to substantially reduce this burden [2,3].

Since the dawn of the new millennium, broad public access to the Internet and e-self-help has become a reality, and this has opened new avenues to reach out to the large, but relatively hidden, group of problem drinkers in the community [4]. Problem drinkers are defined as individuals who consume alcohol beyond the guideline for low-risk drinking. Different gradations of alcohol use disorders may underlie this excess in alcohol consumption. (See Table 1 and Textbox 1 for an overview of specified alcohol use disorders.) Ample meta-analyses have shown face-to-face screening and brief interventions (SBIs) to be effective [5], particularly in primary care [6] and college settings [7]. The wide-scale dissemination of SBIs in routine practice is hampered, however, by implementation barriers, including inadequate supporting policies and resources (time, money, and professional skills) in the health care sector [8] and by meager uptake by problem drinkers themselves [9]. E-self-help may provide a welcome extension to these SBIs. E-self-help interventions are available in both brief and more extended formats. *Single session e-personalized normative feedback* is a phrase used to describe a brief type of self-help delivered over the Internet. Personalized feedback refers to the provision of individualized observations on each drinker’s alcohol consumption patterns in comparison with the recommended low-risk drinking guidelines. Normative feedback is often an important component of these interventions, enabling problem drinkers to compare their own alcohol use (in terms of frequency, quantity, or other measures) to the level of their own cohort or peer group [10].

A more extended form of e-self-help consists of protocol-driven treatments based on principles of behavioral self-control [11], cognitive-behavioral therapy [12], motivational interviewing [13], or a combination of these. The recommended time of use of the extended self-help interventions is 6 weeks, as this is the expected time period in which changes in problematic alcohol use are appearing [14]. Potential benefits have already been illustrated in studies on e-self-help interventions that induce behavioral change in the use of substances such as alcohol or tobacco [15] or that treat mental health disorders like depression and anxiety [16-18].

The chief advantages of e-self-help interventions include their potential to reach broad groups of problem drinkers independent of time or geographical distance and at relatively low dissemination costs [19]. A recent review by Vernon and colleagues [20] has pinpointed similar reasons why users themselves find the interventions attractive, that is, they are timely, anonymous, accessible 24/7, and mostly free of charge. This is especially true of e-self-help interventions that participants can work through without involvement of a professional (defined here as *no-contact* interventions) that are offered for problem drinkers in the general population [21] or directed at students in college settings [22].

Studies investigating the effectiveness of e-self-help interventions among youth have been evaluated mostly in student settings in the United States and Australia and more recently in Europe [23-25]. In a meta-analysis conducted in 2009, Carey and colleagues [26] found a favorable impact of computer-based interventions on student alcohol consumption as compared with no-intervention controls. This favorable impact was also shown in a recent systematic review conducted in 2010 by White and colleagues [27] on the effectiveness of online programs for college and adult problem drinking. However, evaluation studies on e-self-help for student problem drinking in low- and middle-income countries (LMICs) are lacking.

Studies investigating the effectiveness of e-self-help interventions among adult problem drinkers are fewer in number, but they show promising results as well. Many can be characterized as feasibility studies with pretest-posttest designs [28], but the number of randomized controlled studies is on the rise [20]. The availability of evidence-based e-self-help interventions is growing in many high-income nations, including European countries [21,29,30], the United States [4], Canada [31], and Australia [32]. These countries have high Internet penetration rates and a strong public health focus on problem drinking.

We would argue for several reasons that e-self-help interventions could also benefit LMICs. First of all, the majority of people with alcohol use disorders in LMICs are not in treatment, and the many problem drinkers are not exposed to public health interventions at all because no appropriate strategies are in place [33]. The estimated treatment gap of 78% for people with alcohol use disorders in these countries serves to illustrate the many unmet needs [34]. Second, LMICs have meager health resources in terms of both finances and trained health professionals [33]. For countries with minimal resources and increasing problem drinking, such as India and China, low-cost e-self-help interventions might help to fill this public health gap [1,35]. Third, despite the promising results reported by a limited number of studies on face-to-face brief interventions in countries like Brazil [36], India [37], and Taiwan [38], LMICs still experience major obstacles to the full implementation of these SBIs, even more so than affluent countries [33,39]. Fourth, the high level of anonymity provided by self-help Internet interventions could be of value to problem drinkers in those LMICs, where face-to-face help for alcohol problems may be hampered by religious or cultural values that scorn alcohol use or professional help-seeking [40].

Before the scope of e-self-help interventions can be broadened in any type of country, their effectiveness needs to be evaluated beyond the individual studies that have been carried out so far on adult problem drinking. We have therefore conducted a meta-analysis of the currently available studies. As Web-based self-help interventions were preceded by CD-ROM interventions, we include studies on these as well. These CD-ROM studies used PC’s for the delivery of the intervention and applied recruitment strategies similar to the Web-based studies developed at a later stage. We hypothesized that e-self-help without professional guidance would prove effective in reducing problem drinking as compared with control groups that receive no interventions. Next, we examined whether a number of study characteristics impact the primary outcome measure of alcohol consumption. To the best of our knowledge, this paper is the first meta-analysis to report on the effectiveness of no contact e-self-help among adult problem drinkers.

Textbox 1. Alcohol use disorders

Alcohol use disorders from the lexicon of alcohol and drug terms published by the World Health Organization [41]

- Abstinence is defined as refraining from drinking alcoholic beverages.
- Moderate drinking is the consumption of alcohol that does not exceed guidelines for moderate drinking in terms of volume or quantity per occasion.
- Heavy drinking is defined as drinking in excess of the standard of moderate drinking(see moderate drinking, above).
- Hazardous use (*International Classification of Disease, Tenth Revision* [*ICD-10*] code Z72.1) is a pattern of heavy drinking and/or binge drinking that carries with it a risk of harmful consequences to the drinker. These consequences may be detrimental to physical or mental health, or have adverse social consequences to the drinker or others. Other potential consequences include worsening of existing medical conditions or psychiatric illnesses, injuries caused to self or others, due to impaired judgment after drinking, high risk sexual behavior while intoxicated, and worsening of personal or social interactions.
- Harmful drinking (*ICD-10* code F10.1) is a pattern of drinking that is causing damage to health. The damage may be either physical (eg, liver cirrhosis from chronic drinking) or mental (eg, depressive episodes secondary to drinking). Harmful patterns of use are often criticized by others and are sometimes associated with adverse social consequences of various kinds. Harmful drinking has persisted for at least 1 month or has occurred repeatedly over the past 12-month period; subject does not meet criteria for alcohol dependence.
- Alcohol dependence (*ICD-10* code F10.2) is defined as drinking that meets at least 3 of the following criteria: tolerance; withdrawal symptoms; impaired control; preoccupation with acquisition and/or use; persistent desire or unsuccessful efforts to quit; sustains social, occupational, or recreational disability; use continues despite adverse consequences.

Methods

Identification and Selection of Studies

In February 2010 we carried out systematic searches of the literature in the following bibliographical databases: MEDLINE, PsycINFO (1997 to present), Science Citation Index Expanded, Social Sciences Citation Index, Arts and Humanities Citation Index (1997 to present), CINAHL, EMBASE, the Cochrane Drug and Alcohol Group Specialized Register, the Cochrane Effective Practice and Organization of Care (EPOC) Group register, the Alcohol and Alcohol Problems Science Database, and ETOH (etoh.niaaa.nih.gov, 1972 to 2003). Searches were conducted with keywords and text words, in which words indicative of eHealth interventions (Internet, Web, online, manual, and computer) were combined with terms indicative of alcohol disorders (alcohol abuse, alcoholism, problem drinking, hazardous drinking, harmful drinking, abstinence, moderation, treatment, brief intervention, self-help, and e-self-help) and our target group (adults). Those search strategies were combined with the optimal search strategy for randomized controlled trials (RCTs) designed by the United Kingdom Cochrane Centre (Cochrane Collaboration 2008). We also scanned Dissertation Abstracts and Digital Dissertations to cross-check references relating to earlier meta-analyses and systematic reviews on eHealth interventions, brief interventions, and e-self-help interventions for problem drinking as well as unpublished literature. Reference lists of retrieved papers were screened, and papers that possibly met inclusion criteria were retrieved and studied (Figure 1). No language restrictions were applied.

Selection of Primary Studies

In this meta-analysis, we included only those studies that examined e-self-help interventions for adult problem drinkers (aged 18 or older). Studies on e-self-help interventions targeting student populations in college and university settings were excluded, as their effectiveness has been reported in other reviews and meta-analyses [22,26]. From studies that examined both problem and nonproblem drinkers (the latter not exceeding the guidelines for low-risk drinking), we entered the results for the problem drinkers only. We included only those randomized controlled trials that (1) compared e-self-help intervention groups with control groups, (2) reported data that were usable for meta-analytic procedures, and (3) assessed alcohol-drinking behavior (eg, frequency or quantity) as their primary outcome measure (Table 1).

Our initial selection from the first search was based on information derived from titles, abstracts, and keywords; if these yielded insufficient information to assess the inclusion criteria, then the full paper was retrieved. All papers excluded at this stage were independently assessed on their inclusion criteria by two independent raters (authors BB and HR) to ensure an error-free selection procedure. In addition, two independent coders, authors VS and HR, assessed the effect sizes and moderator variables of the included studies; any disagreement was resolved by discussion and consensus (Figure 1).

Methodological Quality Assessment of Primary Studies

Some 25 scales are available to assess the validity and quality of RCTs [42]. As there is no evidence that more elaborate scales give more reliable assessments of validity than simpler ones, we chose an approach similar to that suggested by Higgins and Green [42] as well as approaches applied in two reviews of brief interventions for problem drinking in primary care [6,43]. This resulted in four basic criteria that we used to assess the validity and methodological quality of the studies we analyzed: (1) random allocation to condition performed by an independent third party, (2) random allocation concealed to study participants, (3) randomization status concealed to assessors of outcomes, and (4) completeness of follow-up data.

Meta-analysis

We first examined the effects of e-self-help interventions for problematic alcohol use in comparison with control conditions. The usual approach is to calculate the standardized mean difference, also known as Cohen’s d (the mean difference divided by the pooled standard deviation) that is, $d = (m_1 - m_0) / SD$, where m_1 and m_0 are the mean scores in the experimental and control conditions.

The pooled standard deviation (SD) is defined as $SD = \sqrt{([n_1-1]S_1^2 + [n_0-1]S_0^2) / (n_1 + n_0 -2)}$ where n_1 and n_0, and S_1 and S_0 are the sample sizes and variances of the experimental and control groups as obtained from the primary study. As this effect size d is subject to small-sample bias, it can be adjusted by using a scaling factor, which is multiplied by d to arrive at Hedges' bias-corrected effect size g, where $g = d\ (1- [3/4(n_1 + n_0] -9)$.

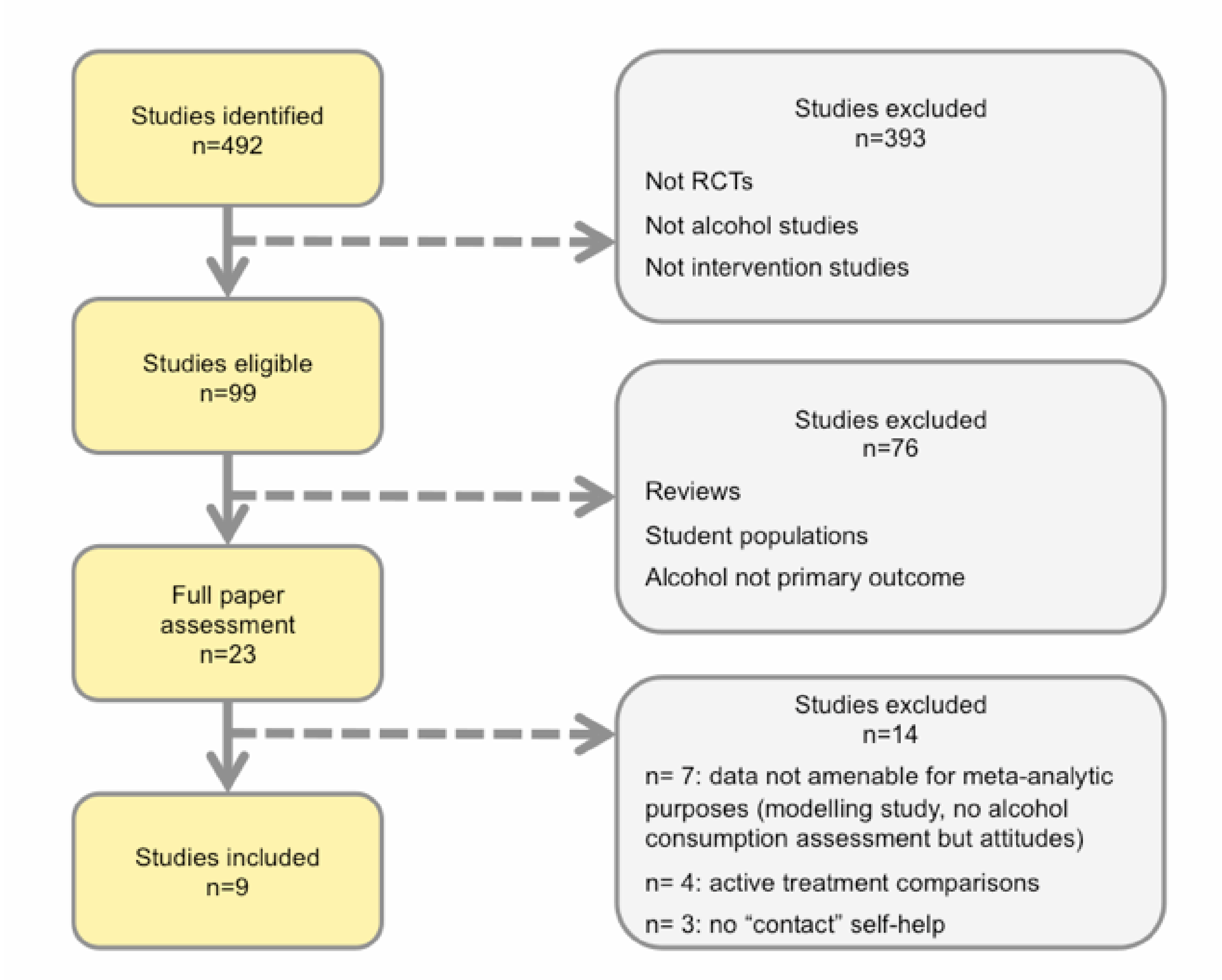
The interpretation of g is simple. An effect size of 0.5 indicates that the mean of the experimental group is half a standard deviation larger than the mean of the control group. In second-order meta-analysis, effect sizes of 0.56 to 1.20 may be interpreted as large, those from 0.33 to 0.55 as moderate, and those from 0 to 0.32 as small [44].

Our effect size calculations were performed on alcohol consumption measures only. If means and standard deviations were not reported, we used other statistics (*F* values and *P* values) to calculate effect sizes. If more than one measure was used in a single primary study, then the mean of the effect sizes was calculated so that each study was represented with only one effect size in our meta-analysis. In one study [45], two experimental conditions were compared with a control condition. In this case, the number of respondents in the control condition was divided equally over the two experimental conditions so that each respondent figured only once in the meta-analysis. In calculating pooled mean effect sizes, we used Comprehensive Meta-analysis (CMA) software, version 2.2.021 [46].

As we did not know before the analysis whether to expect heterogeneity among the studies, we used both the fixed effects model (FEM) and the random effects model (REM) to calculate pooled effect sizes. Heterogeneity was evaluated with the Q statistic and the I^2 statistic. A significant Q rejects the null hypothesis of homogeneity, indicating that the variability among the effect sizes is greater than what would likely result from sampling error alone [47]. The I^2 statistic describes the percentage of total variation across studies, that is, attributable to systematic heterogeneity rather than chance alone. An I^2 value of 25% is associated with low heterogeneity, 50% with moderate, and 75% with high heterogeneity [48].

In subgroup analyses, we tested for significant differences between the effect sizes in different categories of studies using mixed effects analyses in CMA. We analyzed the following attributes: (1) type of treatment (single session personalized normative feedback versus extended self-help); (2) venue where intervention and assessments took place (research, health care, or workplace setting versus participants' homes); (3) type of analysis (intention-to-treat versus completers only); (4) type of control condition (information, assessment-only, waiting list); and (5) small sample sizes ($n < 100$) versus large sample sizes ($n > 100$). Publication bias was tested by funnel plot and by Duval and Tweedie's trim-and-fill procedure, which yields an estimate of effect size after publication bias has been taken into account (both procedures implemented in CMA).

Figure 1. Flow chart of study selection resulting in inclusion of 9 studies and 9 comparisons



Description of Studies

A total of 492 studies were retrieved. Of these, 483 did not meet the inclusion criteria and were excluded (see Figure 1). A total of 9 trials with 9 comparisons and 1553 participants were ultimately included. Selected characteristics of those studies are summarized in Table 1. All 9 studies had been conducted in high-income countries: the United States (3), Canada (1), Netherlands (4), and Germany (1). Of these studies, 5 involved either single-session personalized normative feedback interventions while 4 involved more extended self-help interventions (see Table 1).

Table 1. Selected characteristics of studies (N = 9)

Author, Year, Country	Mode of Delivery/ Setting[a]	Target Group, Inclusion Criteria	Intervention(s)/ Dose	Recruitment	N[b]	Control	Analysis and Timing of Posttreatment Assessment	Attrition Rate (%)
Boon and Huiberts, 2006, Netherlands [49]	Internet/ research setting	Males/females, ≥ 21/14 units/week and/or ≥ 6/4 units on ≥ 1 day/week	PNF[c]/ single session	Community	191	Alcohol leaflet	Completers only/ 9 months	32
Boon et al, 2011, Netherlands [50]	Internet/ research setting	Males, ≥ 21 units/week and/or ≥ 6 units on ≥1 day/week	PNF[c]/ single session	Community	450	Alcohol leaflet	Intention-to-treat/ 1 month	8
Cunningham et al, 2009, Canada [31]	Internet/ home	Alcohol Use Disorders Identification Test (AUDIT) ≥ 11	PNF[c]/ single session	Community	72	Alcohol leaflet	Intention-to-treat/ 3 months	8
Doumas and Hannah, 2008, United States [45]	Internet/ Work-place	Males/females, ≥1 occasions with ≥ 5/4 drinks in last 2 weeks	PNF[c]/ single session	Workplace	22	Assessment only	Completers only/ 1 month	37
Hester and Delaney, 1997, United States [51]	CD-ROM/ health care setting	AUDIT ≥ 8; males/females > 120/70 units/month or ≥ 6 units on ≥ 1 day/week	BSC[d]/ 8 sessions	Community	40	Waiting list	Completers only/ 10 weeks	0
Hester et al, 2005, United States [52]	CD-ROM/ health care setting	AUDIT ≥ 8	PNF[c], BSC[d], MI[e]/single session extened	Community	61	Waiting list	Completers only/ 4 weeks	0
Kramer et al, 2009, Netherlands [53]	TV/ Internet/ manual/ home	Males/females, > 21/14 units/week and/or ≥ 6/4 units on ≥ 1 day/week past month	BSC[d], CBT[f], MI[e]/ 5 sessions	Community	181	Waiting list	Intention-to-treat/ 5 weeks	6
Neumann et al, 2006, Germany [54]	Internet/ ED[g]	AUDIT ≥ 5	PNF[c]/ single session	Emergency department	275	Assessment only	Completers only/ 6 months	37
Riper et al, 2008, Netherlands [57]	Internet/ home	Males/females, >21/14 units/week or ≥6/4 units/≥1 day for past 3 months	BSC[d], CBT[f], MI[e]/ 6 weeks	Community	261	Alcohol leaflet	Intention-to-treat/ 6 months	42

[a] At time of study; hence differences may exist between study and real life delivery

[b] Number of problem drinkers

[c] Personalized normative feedback

[d] Behavioral self-control training

[e] Motivational interviewing

[f] Cognitive behavioral therapy

[g] Emergency department

Methodological Quality of Included Studies

All studies used well-validated alcohol consumption measures and well-described, theoretically based interventions [56]. In 4 studies [49,53,50,57], treatment was allocated by an independent third party; in 2 studies [31,45] it was not, and in 3 studies [51,52,54] this was unclear. Concealment of random allocation to participants was not applicable as study participants were recruited by self-reference and informed about the study conditions. In all but 3 studies [51,52,54] all outcome measures were self-administered by participants, making it irrelevant whether the researchers assessing the outcomes were blinded to group assignment. Dropout rates differed widely from 0% to 42% (Table 1).

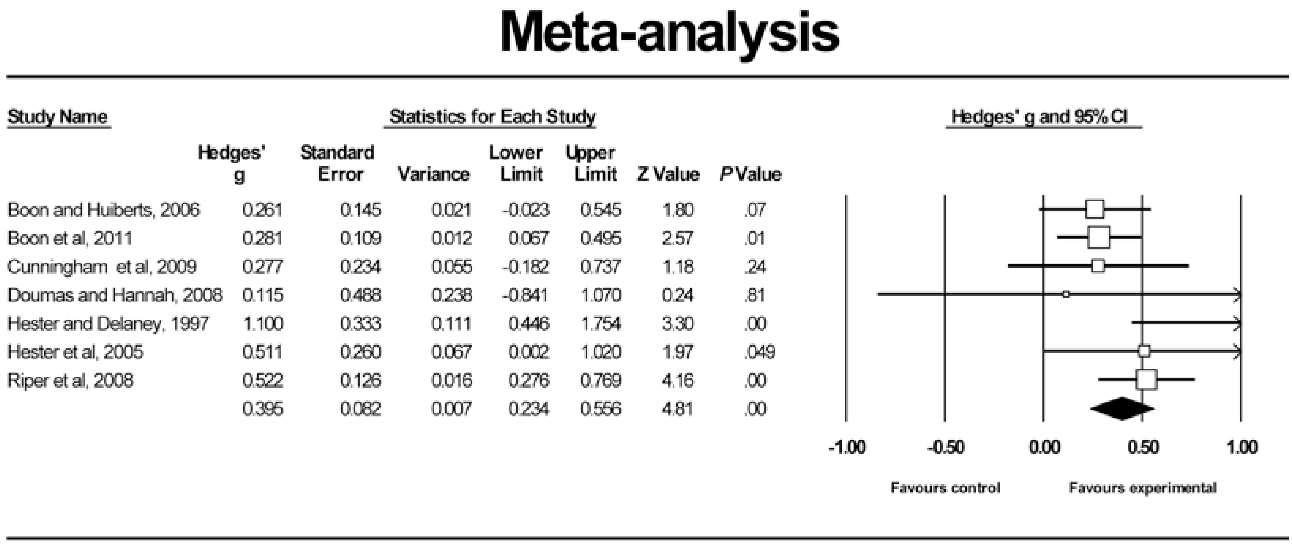
Results

Fixed effects meta-analysis of all studies resulted in a mean effect size of g = 0.39 (95% confidence interval [CI] 0.29-0.50). Random effects analysis resulted in a mean effect size of g = 0.44 (95% CI: 0.17-0.71). The hypothesis of homogeneity was rejected (Q = 42.30, I² = 81.08%, *P* < .001), thus lending preference to the random effects model. As possible sources of heterogeneity, we identified 2 outliers. The first was the Kramer study [53], with a large effect size of g = 1.11 (95% CI 0.80-1.42, random effects model). The main intervention component in this study consisted of five 25-minute televised sessions, a self-help manual, and a website. The second outlier was the Neumann study [54], with a small, negative effect size of g = −0.14 (95% CI −0.41 to 0.15, random effects model), which differed from other interventions in that recruited participants had just been treated at a hospital emergency department and were given an e-self-help intervention that did not address alcohol as such, but provided general lifestyle advice. Excluding these 2 outliers, we still obtained mean effect sizes of g = 0.39 (95% CI 0.23-0.56) in the random effects model (Q = 8.19, I² = 26.75%, *P* = .22) (see Figure 2). We also calculated the number needed to treat (NNT) as the effect size is not easy to interpret from a clinical point of view. We transformed the effect sizes based on *z* scores by using the formulae provided by Kraemer and Kupfer [58]. The results translate to an NNT of 5, indicating that about 5 problem drinkers must receive the intervention to generate 1 good treatment response. Posttreatment assessment was conducted at different points in time in different studies, ranging from 4 weeks to 9 months. Meta-regression analyses did not establish significant differences in intervention effects over time at posttreatment (beta = 0.00052, 95% CI = −0.00 to 0.009, *P* = .91).

Subgroup Analysis

We then performed subgroup analyses to examine the contrast between extended e-self-help interventions and e-single-session personalized normative feedback interventions (omitting the outliers). Based upon the mixed effects model, these yielded a mean effect size of g = 0.27 (95% CI 0.11-0.43) for the e-single-session interventions and g = 0.61 (95% CI 0.33-0.90) for the extended e-self help interventions, a significant difference of *P* = .04 (Table 2). We found no significant differences between (1) where interventions and assessments were performed (research, health centre, or workplace versus places from where participants' accessed the Internet, such as at home); (2) type of control condition (information, assessment-only, waiting-list); (3) small samples (n < 100) versus large samples (n > 100); nor (4) type of analysis (completers-only versus intention-to-treat).

Figure 2. Meta-analysis of studies (omitting outliers)



Study Name	Hedges' g	Standard Error	Variance	Lower Limit	Upper Limit	Z Value	*P* Value
Boon and Huiberts, 2006	0.261	0.145	0.021	-0.023	0.545	1.80	.07
Boon et al, 2011	0.281	0.109	0.012	0.067	0.495	2.57	.01
Cunningham et al, 2009	0.277	0.234	0.055	-0.182	0.737	1.18	.24
Doumas and Hannah, 2008	0.115	0.488	0.238	-0.841	1.070	0.24	.81
Hester and Delaney, 1997	1.100	0.333	0.111	0.446	1.754	3.30	.00
Hester et al, 2005	0.511	0.260	0.067	0.002	1.020	1.97	.049
Riper et al, 2008	0.522	0.126	0.016	0.276	0.769	4.16	.00
	0.395	0.082	0.007	0.234	0.556	4.81	.00

Sensitivity Analyses

The overall mean effect size was maintained even after exclusion of the largest study (N = 450) [50] in a random effects model (g = 0.44, 95% CI 0.24-0.64, Q = 6.87, I² = 27.23, *P* = .23).

Duval and Tweedie's trim-and-fill analysis did not detect publication bias (observed at g = 0.38, 95% CI 0.25-0.51, adjusted g = 0.35, 95% CI 0.23-0.47), nor did the funnel plot analysis detect bias. In view of these findings, we assume our post-treatment results to be robust up to 9 months [59].

Table 2. Effect sizes of e-interventions for problem drinking versus control conditions

Studies	Number of Comparisons	Hedges' g	95% CI	*P*
All studies[a]	9	0.44	0.29-0.50	
All studies, outliers excluded[a] [53,54]	7	0.39	0.23-0.57	
Type of treatment [b]				.04
e-personalised normative feedback	4	0.27	0.11-0.43	
e-self-help	3	0.61	0.33-0.90	
Type of analysis [b]				.60
Intention-to-treat	3	0.37	0.21-0.54	
Completers-only	4	0.48	0.11-0.86	
Type of venue [b]				.63
Home	2	0.47	0.25-0.69	
Research, health centre, or workplace setting	5	0.39	0.15-0.63	
Sample size [b]				.43
Small	3	0.36	0.19-0.52	
Large	4	0.52	0.14-0.91	
Type of control condition [b]				
Alchol leaflet	4	0.35	0.21-0.48	.33
Assessment only	1	0.12	−0.84 to 1.07	
Waitlist control	2	0.77	0.19-1.34	

[a] Random effect model

[b] Mixed effects model

Discussion

We found a medium effect size (g = 0.39) for eHealth interventions to reduce adult problem drinking in the general population up to 6 or 9 months posttreatment, as compared with no intervention. A significant difference (*P* = .04) emerged between e-single-session personalized normative feedback interventions (g = 0.27) and e-self-help interventions of a more extended nature (g = 0.61). This suggests that the latter may be more effective. Effects of the interventions beyond 9 months could not be assessed; one study reported 12-month follow-up results, and these suggested cost-effectiveness but a fade-out of the effect obtained at 6 months [55]; and one study was published after our search and found similar fade-out results at twelve months [60].

The medium effect size of our analysis compares favorably with the small effects reported for e-interventions in three recent meta-analyses. Rooke et al [15], Portnoy et al [61], and Riper et al [62] found small effect sizes for e-interventions for problem drinking. Some explanations may lie in differential characteristics of the interventions studied. The Riper study focused on single-session personalized normative feedback; the effect size of d = 0.23 was consistent with our present findings for this type of intervention (g = 0.27, 95% CI 0.11-0.42). Other explanations may involve characteristics of the control groups (active or nonactive). The study by Rooke et al, for example, included some active treatment comparisons, and these generally diminish effect sizes. The smaller effects in other meta-analyses may have also derived from the inclusion of both adult and college populations in some analyses, as e-self-help interventions are known to show smaller effects among younger age groups and student populations. This holds especially when these include prevention studies that include both alcohol drinking and nondrinking college students [26]. Of course, randomized controlled trials that compare different interventions in different populations are essential to assess the robustness of the various observations. The medium effect size in our analysis also compares well with effects reported for face-to-face adult brief interventions in primary care [6], for postal self-help interventions [63], and for brief interventions in non–treatment-seeking populations [5]. This further illustrates the potential of Internet interventions to extend the array of public health services to combat problem drinking.

Limitations

Because our analysis is based on a rather limited number of studies, the results can be generalized only to self-referred adult problem drinkers in high-income countries recruited via the media; this implies samples of individuals with high readiness to change [64]. Second, some studies had small samples; we dealt with this by analyzing our data with Hedges' g to adjust our estimates for small-sample bias [65]. Third, the loss to follow-up in some of the studies we reviewed was substantial.

High dropout rates are a common feature of both online and offline self-help interventions for problem drinking and for Internet interventions in general [66,67]. Although some studies applied imputation techniques to handle the loss to follow-up, the high attrition may have biased our overall results.

Implications for Clinical Practice and Future Research

The medium effect size we found for no-contact eHealth interventions could imply a major health impact at the population level, in view of the high percentages of problem drinkers that eHealth interventions might potentially reach. Naturally, eHealth interventions are subject to some constraints, as participants need computer and Internet access and a reasonable degree of literacy. While the costs of disseminating and scaling up no-contact eHealth interventions are low, the costs of developing them can be substantial [68].

Many questions still remain unanswered. We do not know yet whether e-self-help interventions that include professional contacts are more effective against problem drinking than no-contact interventions and, if so, to what *types* of problem drinkers that might apply (first-time help-seekers in the general population, primary- or secondary-care populations, hazardous or harmful drinkers, or dependent drinkers. (See Textbox 1.) This contrasts with e-self-help intervention studies focusing on common mental health disorders like depression and anxiety [69], which have established a firm evidence base for better clinical outcomes when e-self-help interventions are delivered with some form of professional guidance [18]. By contrast, Doumas and colleagues [45] have observed that a no-contact e-self-help intervention combined with 15 minutes of face-to-face motivational interviewing produced reductions in drinking that were comparable with the results of the no-contact e-self-help intervention alone. Nor could Rooke et al [15] establish any effect moderation attributable to personal guidance in alcohol e-interventions. As most participants in e-self-help interventions for problem drinking are first-time help-seekers [20,21], it could well be the case for this group that timely, anonymous access to e-interventions has a greater impact in altering drinking patterns than professional contact per se [70]. No-contact e-self-help interventions might therefore be an effective first-line choice in a stepped-care approach to problem drinking. A recently published study [71] on the posttreatment (3 months) effectiveness of intensive online treatment with active involvement of a therapist (duration of 3 months with 2 online therapist contacts a week) revealed an effect size of $d = 1.21$. This may indicate that therapist involvement will lead to higher effect sizes when compared with no-contact self-help interventions. More studies are needed, however, to assess the robustness of this observation and to assess how these more intensive treatments fit into a stepped-care approach.

From an economic point of view, no-contact e-self-help interventions could carry considerable promise as compared to other approaches like screening and brief face-to-face interventions in primary care, especially since the latter have relatively high implementation costs [72]. Studies that rigorously assess this proposition are not yet available [20], but economic evaluation studies on e-self-help for depression show favorable results as shown by a recent study of Gerhards and colleagues [73] in primary care and by Warmerdam and colleagues among depressed persons in the community [74]. The paper of Smit and colleagues in this issue shows that such economic advantages may also be expected from e-self-help intervention when implemented on a broad scale [75].

In light of all the potential benefits of e-self-help interventions in curbing problem drinking, we recommend their further evaluation and implementation as a means to bridging the gap in treating alcohol use disorders in low- and middle-income countries (LMICs). Although some may argue that this is not a viable option given the low Internet penetration in such countries, we would point out that such arguments abounded in high-income countries at the start of the millennium when e-self-help interventions for hazardous drinking were first being introduced. Meanwhile, the digital divide in Internet access between the more and less affluent countries is narrowing. In fact, one in four people worldwide now have Internet access [76]. Some e-self-help interventions have also proven well suited to be shared globally, as exemplified by the work of Munoz [77] on no-contact e-self-help for smoking cessation that are now used by people all over the world.

Conclusion

Our study has sought to synthesize the available evidence about the effectiveness of no-contact e-self-help interventions in curbing adult problem drinking. The data suggest that these are effective. Our findings also highlight the need for more evaluations of the clinical outcomes and the cost-effectiveness of online screening instruments and interventions. The paper of Smit and colleagues in this issue indeed illustrates the economic benefits of e-self-help interventions for curbing adult problem drinking [75]. Future studies should shed light on whether e-self-help interventions produce similar or better results when extended with face-to-face components and on whether they could serve as alternatives or adjuncts to face-to-face treatments in primary care settings.

Acknowledgments

This work was undertaken in the framework of the project on implementing eHealth technologies in the area of substance abuse coordinated and sponsored by the World Health Organization (WHO). The authors also thank Roseli Boerngen, PhD, and Atul Ambekar, MD, for their feedback on the manuscript. The authors alone are responsible for the views expressed in this publication.

Conflicts of Interest

None declared

References

1. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009 Jun 27;373(9682):2223-2233. [doi: 10.1016/S0140-6736(09)60746-7] [Medline: 19560604]
2. Institute of Medicine. Broadening the Base of Treatment of Alcohol Problems: Report of a Study by a Committee of the Institute of Medicine. Washington, DC: National Academy Press; 1990.
3. Cuijpers P, Riper H, Lemmers L. The effects on mortality of brief interventions for problem drinking: a meta-analysis. Addiction 2004 Jul;99(7):839-845. [doi: 10.1111/j.1360-0443.2004.00778.x] [Medline: 15200579]
4. Hester RK, Miller JH. Computer-based tools for diagnosis and treatment of alcohol problems. Alcohol Res Health 2006;29(1):36-40 [FREE Full text] [Medline: 16767852]
5. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002 Mar;97(3):279-292. [Medline: 11964101]
6. Kaner EF, Beyer F, Dickinson HO, Pienaar E, Campbell F, Schlesinger C, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev 2007(2):CD004148. [doi: 10.1002/14651858.CD004148.pub3] [Medline: 17443541]
7. Wechsler H, Lee JE, Kuo M, Seibring M, Nelson TF, Lee H. Trends in college binge drinking during a period of increased prevention efforts. Findings from 4 Harvard School of Public Health College Alcohol Study surveys: 1993-2001. J Am Coll Health 2002 Mar;50(5):203-217. [Medline: 11990979]
8. Barry KL, Blow FC, Willenbring ML, McCormick R, Brockmann LM, Visnic S. Use of Alcohol Screening and Brief Interventions in Primary Care Settings:Implementation and Barriers. Subst Abus 2004 Mar;25(1):27-36. [Medline: 15201109]
9. Cunningham JA, Breslin FC. Only one in three people with alcohol abuse or dependence ever seek treatment. Addict Behav 2004 Jan;29(1):221-223. [Medline: 14667433]
10. Chan KK, Neighbors C, Gilson M, Larimer ME, Alan Marlatt G. Epidemiological trends in drinking by age and gender: providing normative feedback to adults. Addict Behav 2007 May;32(5):967-976. [doi: 10.1016/j.addbeh.2006.07.003] [Medline: 16938410]
11. Hester RK. Self-control training. In: Hester RK, Miller WR, editors. Handbook of Alcoholism Treatment Approaches: Effective Alternatives. Needham Heigths, MA: Allyn & Bacon; 2003:152-164.
12. Kalden R, Caroll K, Donovan D, Cooney N, Monti P, Abrams D. Cognitive Behavioral Coping Skills Therapy Manual: A Clinical Research Guide for Therapists Treating Individuals With Alcohol Abuse and Dependence. In: Mattson MEM, editor. Mattson MEM, editor. NIAAA Project MATCH Monograph Series. Rockville, MD: National Institute on Alcoholism and Alcohol Abuse; 1992.
13. Miller WR, Rollnick S. Motivational Interviewing: Preparing People for Change. 2nd edition. New York, NY: Guilford; 2002.
14. Breslin FC, Sobell MB, Sobell LC, Buchan G, Cunningham JA. Toward a stepped care approach to treating problem drinkers: the predictive utility of within-treatment variables and therapist prognostic ratings. Addiction 1997 Nov;92(11):1479-1489. [Medline: 9519491]
15. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
16. 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]
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 Jun;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
18. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
19. Cunningham JA, Selby PL, Kypri K, Humphreys KN. Access to the Internet among drinkers, smokers and illicit drug users: is it a barrier to the provision of interventions on the World Wide Web? Med Inform Internet Med 2006 Mar;31(1):53-58. [doi: 10.1080/14639230600562816] [Medline: 16754367]
20. Vernon ML. A review of computer-based alcohol problem services designed for the general public. J Subst Abuse Treat 2010 Apr;38(3):203-211. [doi: 10.1016/j.jsat.2009.11.001] [Medline: 20015607]
21. Riper H, Kramer J, Conijn B, Smit F, Schippers G, Cuijpers P. Translating effective web-based self-help for problem drinking into the real world. Alcohol Clin Exp Res 2009 Aug;33(8):1401-1408. [doi: 10.1111/j.1530-0277.2009.00970.x] [Medline: 19413646]
22. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]

23. Riper H, Andersson G, Christensen H, Cuijpers P, Lange A, Eysenbach G. Theme issue on e-mental health: a growing field in internet research. J Med Internet Res 2010;12(5):e74 [FREE Full text] [doi: 10.2196/jmir.1713] [Medline: 21169177]
24. Spijkerman R, Roek MA, Vermulst A, Lemmers L, Huiberts A, Engels RC. Effectiveness of a web-based brief alcohol intervention and added value of normative feedback in reducing underage drinking: a randomized controlled trial. J Med Internet Res 2010;12(5):e65 [FREE Full text] [doi: 10.2196/jmir.1465] [Medline: 21169172]
25. Cunningham JA, Khadjesari Z, Bewick BM, Riper H. Internet-based interventions for problem drinkers: From efficacy trials to implementation. Drug Alcohol Rev 2010 Nov;29(6):617-622. [doi: 10.1111/j.1465-3362.2010.00201.x] [Medline: 20973846]
26. Carey KB, Scott-Sheldon LA, Elliott JC, Bolles JR, Carey MP. Computer-delivered interventions to reduce college student drinking: a meta-analysis. Addiction 2009 Nov;104(11):1807-1819. [doi: 10.1111/j.1360-0443.2009.02691.x] [Medline: 19744139]
27. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
28. Linke S, Brown A, Wallace P. Down your drink: a web-based intervention for people with excessive alcohol consumption. Alcohol Alcohol 2004;39(1):29-32 [FREE Full text] [Medline: 14691071]
29. Sinadinovic K, Berman AH, Hasson D, Wennberg P. Internet-based assessment and self-monitoring of problematic alcohol and drug use. Addict Behav 2010 May;35(5):464-470. [doi: 10.1016/j.addbeh.2009.12.021] [Medline: 20092953]
30. Rodríguez-Martos A, Castellano Y. Web-based screening and advice for hazardous drinkers: use of a Spanish site. Drug Alcohol Rev 2009 Jan;28(1):54-59. [doi: 10.1111/j.1465-3362.2008.00005.x] [Medline: 19320676]
31. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032. [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
32. 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]
33. Benegal V, Chand PK, Obot IS. Packages of care for alcohol use disorders in low- and middle-income countries. PLoS Med 2009 Oct;6(10):e1000170 [FREE Full text] [doi: 10.1371/journal.pmed.1000170] [Medline: 19859536]
34. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
35. Prasad R. Alcohol use on the rise in India. Lancet 2009 Jan 3;373(9657):17-18. [Medline: 19125444]
36. Simão MO, Kerr-Corrêa F, Smaira SI, Trinca LA, Floripes TM, Dalben I, et al. Prevention of "risky" drinking among students at a Brazilian university. Alcohol Alcohol 2008;43(4):470-476 [FREE Full text] [doi: 10.1093/alcalc/agn019] [Medline: 18364361]
37. Pal HR, Yadav D, Mehta S, Mohan I. A comparison of brief intervention versus simple advice for alcohol use disorders in a North India community-based sample followed for 3 months. Alcohol Alcohol 2007;42(4):328-332 [FREE Full text] [doi: 10.1093/alcalc/agm009] [Medline: 17360720]
38. Tsai YF, Tsai MC, Lin YP, Chen CY. Brief intervention for problem drinkers in a Chinese population: a randomized controlled trial in a hospital setting. Alcohol Clin Exp Res 2009 Jan;33(1):95-101. [doi: 10.1111/j.1530-0277.2008.00815.x] [Medline: 18945220]
39. Chisholm D, Lund C, Saxena S. Cost of scaling up mental healthcare in low- and middle-income countries. Br J Psychiatry 2007 Dec;191:528-535 [FREE Full text] [doi: 10.1192/bjp.bp.107.038463] [Medline: 18055957]
40. de Toledo Piza Peluso E, Blay SL. How should alcohol dependence be treated? The public view. Alcohol Alcohol 2008;43(5):600-605 [FREE Full text] [doi: 10.1093/alcalc/agn054] [Medline: 18583544]
41. 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]
42. Higgins JPT, Green SE, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 2009 URL: http://www.cochrane-handbook.org/ [accessed 2011-05-20] [WebCite Cache ID 5ysSK70Nu]
43. Whitlock EP, Polen MR, Green CA, Orleans T, Klein J, U.S. Preventive Services Task Force. Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2004 Apr 6;140(7):557-568. [Medline: 15068985]
44. Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. Am Psychol 1993 Dec;48(12):1181-1209. [Medline: 8297057]
45. Doumas DM, Hannah E. Preventing high-risk drinking in youth in the workplace: a web-based normative feedback program. J Subst Abuse Treat 2008 Apr;34(3):263-271. [doi: 10.1016/j.jsat.2007.04.006] [Medline: 17600650]
46. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive Meta-analysis Version 2. Englewood, NJ: Biostat Inc; 2005.
47. Lipsey MW, Wilson DB. Practical Meta-analysis. Thousand Oaks, CA: Sage; 2001.

48. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
49. Boon B, Huiberts A. Preventie problematisch alcoholgebruik. In: Meijer SA, Smit F, Schoemaker CG, Cuijpers P, editors. Gezond verstand. Evidence-based preventie van psychische stoornissen [Common sense. Evidence-based prevention of mental disorders] Volksgezondheid Toekomstverkenning (VTV) Themarapport. Bilthoven, Netherlands: RIVM/Trimbos Institute; 2006:107-114.
50. Boon B, Risselada A, Huiberts A, Riper H, Smit F. Curbing alcohol use in male adults through computer generated personalised advice: randomised controlled trial. J Med Internet Res 2011;13(2):e43. [doi: 10.2196/jmir.1695]
51. Hester RK, Delaney HD. Behavioral Self-Control Program for Windows: results of a controlled clinical trial. J Consult Clin Psychol 1997 Aug;65(4):686-693. [Medline: 9256570]
52. Hester RK, Squires DD, Delaney HD. The Drinker's Check-up: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers. J Subst Abuse Treat 2005 Mar;28(2):159-169. [doi: 10.1016/j.jsat.2004.12.002] [Medline: 15780546]
53. Kramer J, Riper H, Lemmers L, Conijn B, van Straten A, Smit F. Television-supported self-help for problem drinkers: a randomized pragmatic trial. Addict Behav 2009 May;34(5):451-457. [doi: 10.1016/j.addbeh.2008.12.015] [Medline: 19179014]
54. Neumann T, Neuner B, Weiss-Gerlach E, Tønnesen H, Gentilello LM, Wernecke KD, et al. The effect of computerized tailored brief advice on at-risk drinking in subcritically injured trauma patients. J Trauma 2006 Oct;61(4):805-814. [doi: 10.1097/01.ta.0000196399.29893.52] [Medline: 17033544]
55. Smit F, Riper H, Kramer J, Schippers G, Cuijpers P. Cost-effectiveness of a web-based self-help intervention for problem drinking: randomized trial. In: Riper H, editor. Riper, H, editor. Curbing Problem Drinking in the Digital Galaxy. Amsterdam: Vrije Universiteit Amsterdam; 2008:41-64.
56. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med 2010 Mar;7(3):e1000251 [FREE Full text] [doi: 10.1371/journal.pmed.1000251] [Medline: 20352064]
57. 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]
58. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry 2006 Jun 1;59(11):990-996. [doi: 10.1016/j.biopsych.2005.09.014] [Medline: 16368078]
59. Wilson DB. Meta-analyses in alcohol and other drug abuse treatment research. Addiction 2000 Nov;95(Suppl 3):S419-S438. [Medline: 11132366]
60. Cunningham JA, Wild TC, Cordingley J, Van Mierlo T, Humphreys K. Twelve-month follow-up results from a randomized controlled trial of a brief personalized feedback intervention for problem drinkers. Alcohol Alcohol 2010;45(3):258-262 [FREE Full text] [doi: 10.1093/alcalc/agq009] [Medline: 20150170]
61. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16. [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
62. Riper H, van Straten A, Keuken M, Smit F, Schippers G, Cuijpers P. Curbing problem drinking with personalized-feedback interventions: a meta-analysis. Am J Prev Med 2009 Mar;36(3):247-255. [doi: 10.1016/j.amepre.2008.10.016] [Medline: 19215850]
63. Apodaca TR, Miller WR. A meta-analysis of the effectiveness of bibliotherapy for alcohol problems. J Clin Psychol 2003 Mar;59(3):289-304. [doi: 10.1002/jclp.10130] [Medline: 12579546]
64. Rollnick S, Heather N, Gold R, Hall W. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. Br J Addict 1992 May;87(5):743-754. [Medline: 1591525]
65. Hedges LV, Olkin I. Statistical Methods for Meta-analysis. Orlando: Academic Press; 1985.
66. Miller WR, Wilbourne PL. Mesa Grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. Addiction 2002 Mar;97(3):265-277. [Medline: 11964100]
67. 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]
68. McCambridge J, O'Donnell O, Godfrey C, Khadjesari Z, Linke S, Murray E, et al. How big is the elephant in the room? Estimated and actual IT costs in an online behaviour change trial. BMC Res Notes 2010;3:172 [FREE Full text] [doi: 10.1186/1756-0500-3-172] [Medline: 20565866]
69. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
70. Bischof G, Grothues JM, Reinhardt S, Meyer C, John U, Rumpf HJ. Evaluation of a telephone-based stepped care intervention for alcohol-related disorders: a randomized controlled trial. Drug Alcohol Depend 2008 Mar 1;93(3):244-251. [doi: 10.1016/j.drugalcdep.2007.10.003] [Medline: 18054443]

71. 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]
72. Chisholm D, Rehm J, Van Ommeren M, Monteiro M. Reducing the global burden of hazardous alcohol use: a comparative cost-effectiveness analysis. J Stud Alcohol 2004 Nov;65(6):782-793. [Medline: 15700517]
73. Gerhards SA, de Graaf LE, Jacobs LE, Severens JL, Huibers MJ, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. Br J Psychiatry 2010 Apr;196:310-318 [FREE Full text] [doi: 10.1192/bjp.bp.109.065748] [Medline: 20357309]
74. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010;12(5):e53 [FREE Full text] [doi: 10.2196/jmir.1436] [Medline: 21169166]
75. Smit F, Lokkerbol J, Riper H, Majo C, Boon B, Blankers M. Modeling the cost-effectiveness of health care systems for alcohol use disorders. J Med Internet Res 2011;13. [doi: 10.2196/jmir.1694]
76. Internet World Stats. 2010. Internet Usage Statistics: World Internet Users and Population Stats URL: http://www.internetworldstats.com/stats.htm [accessed 2011-05-24] [WebCite Cache ID 5yvPkybSm]
77. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, et al. Toward evidence-based Internet interventions: A Spanish/English Web site for international smoking cessation trials. Nicotine Tob Res 2006 Feb;8(1):77-87. [doi: 10.1080/14622200500431940] [Medline: 16497602]

Edited by G Eysenbach; submitted 21.11.10; peer-reviewed by R Wiers, N Heather, R Tait, S Rooke, M Postel; comments to author 15.12.10; revised version received 12.02.11; accepted 10.03.11; published 30.06.11.

Please cite as:
Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, Smit F
Effectiveness of E-Self-help Interventions for Curbing Adult Problem Drinking: A Meta-analysis
J Med Internet Res 2011;13(2):e42
URL: http://www.jmir.org/2011/2/e42/
doi:10.2196/jmir.1691
PMID:21719411

Original Paper

Curbing Alcohol Use in Male Adults Through Computer Generated Personalized Advice: Randomized Controlled Trial

Brigitte Boon[1,2], PhD; Anneke Risselada[2,3], MSc; Annemarie Huiberts[4], PhD; Heleen Riper[5,6], PhD; Filip Smit[1,7], PhD

[1]Trimbos-institute, Department of Public Mental Health, Utrecht, Netherlands
[2]Addiction Research Institute, Rotterdam, Netherlands
[3]Municipal Health Authority, South-Holland South region, Netherlands
[4]Netherlands Institute for Health Promotion and Disease Prevention, Woerden, Netherlands
[5]Trimbos-institute, Innovation Center for Mental Health and Technology, Utrecht, Netherlands
[6]VU University, Department of Clinical Psychology, Amsterdam, Netherlands
[7]VU University Medical Centre, Department of Epidemiology and Biostatistics, Amsterdam, Netherlands

Corresponding Author:
Brigitte Boon, PhD
Trimbos-institute
Department of Public Mental Health
POBox 725
Utrecht, 3500 AS
Netherlands
Phone: 31 302971145
Fax: 31 302971111
Email: bboon@trimbos.nl

Abstract

Background: In recent years, interventions that deliver online personalized feedback on alcohol use have been developed and appear to be a feasible way to curb heavy drinking. Randomized controlled trials (RCTs) among the general adult population, however, are scarce. The present study offers an RCT of Drinktest.nl, an online personalized feedback intervention in the Netherlands.

Objective: The aim of this study was to assess the effectiveness of computer-based personalized feedback on heavy alcohol use in male adults.

Methods: Randomization stratified by age and educational level was used to assign participants to either the intervention consisting of online personalized feedback or an information-only control condition. Participants were told as a cover story that they would evaluate newly developed health education materials. Participants were males (n = 450), aged 18 to 65 years, presenting with either heavy alcohol use (> 20 units of alcohol weekly) and/or binge drinking (> 5 units of alcohol at a single occasion at least 1 day per week) in the past 6 months. They were selected with a screener from a sampling frame of 25,000 households. The primary outcome measure was the percentage of the participants that had successfully reduced their drinking levels to below the Dutch guideline threshold for at-risk drinking.

Results: Intention-to-treat analysis showed that in the experimental condition, 42% (97/230) of the participants were successful in reducing their drinking levels to below the threshold at the 1-month follow-up as compared with 31% (67/220) in the control group (odds ratio [OR] = 1.7, number needed to treat [NNT] = 8.6), which was statistically significant ($\chi^2_1 = 6.67$, $P = .01$). At the 6-month follow-up, the success rates were 46% (105/230) and 37% (82/220) in the experimental and control conditions, respectively (OR = 1.4, NNT = 11.9), but no longer statistically significant ($\chi^2_1 = 3.25$, $P = .07$).

Conclusions: Personalized online feedback on alcohol consumption appears to be an effective and easy way to change unhealthy drinking patterns in adult men, at least in the short-term.

Trial registration: International Standard Randomized Controlled Trial Number: NTR836; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=836 (Archived by WebCite at http://www.webcitation.org/5ytnEz2vp)

(J Med Internet Res 2011;13(2):e43) doi:10.2196/jmir.1695

KEYWORDS

Web-based personalized feedback; alcohol; Internet; heavy drinking; problem drinking; adult

Introduction

The present study aims to examine the effect of Drinktest (www.drinktest.nl), an online personalized feedback intervention targeted at heavy drinking adults in the Netherlands. It is important to inform heavy drinkers about the possible consequences of their drinking behavior. After all, heavy alcohol consumption is highly prevalent worldwide [1] and is associated with a significant disease burden [2] and a range of health-related adverse consequences in later life, such as liver cirrhosis, cancer of the esophagus and the stomach, and possibly the onset of depressive disorder [3]. Heavy drinking is not only associated with morbidity, but also with excess mortality [4]. For these reasons alone, there is wisdom in reducing heavy alcohol use. This is likely to also reduce the risk of a host of other problems, such as crime, domestic violence, and traffic accidents. In addition, it may result in cost savings not only in health care (eg, fewer hospital stays), but it may also be advantageous for the economy when people are less often absent from their work and are more efficient while at work [3,5].

Unfortunately, it is not easy to reach heavy drinkers with face-to-face interventions. There often is a shortage of health care professionals who can deliver the interventions, even in resource rich countries. Moreover, heavy drinkers may be reluctant to discuss their drinking behavior [6]. As a consequence, a substantial 80% of heavy drinkers do not engage in any formal treatment [7].

Offering interventions online may help to solve this problem. Online interventions designed to decrease alcohol consumption have proven to be feasible instruments to reach heavy drinkers and are generally well received [8,9]. People can engage in the intervention whenever they choose and in the privacy of their home without fear of stigmatization. Moreover, online self-help interventions require no therapist time. Evidence regarding the effectiveness of online interventions is slowly building up. Results are promising [8,10].

Most of the evidence regarding the effectiveness of online alcohol interventions is collected in studies aimed at college students. In a recent meta-analysis, Carey and colleagues [11] found that computer-delivered interventions produced significant improvement on both quantity and frequency of drinking in college samples, that the online interventions were preferred to no intervention, and that their effects were comparable to those of alternative alcohol-related interventions.

However, for population segments other than students, the results found in literature are not yet conclusive. A recent meta-analysis focusing on online alcohol and tobacco interventions in the general population suggested positive outcomes, that is, an overall effect size (Cohen’s d) of 0.22 (95% confidence interval [CI] 0.14 - 0.29) for the alcohol interventions [12]. Close examination of the findings, however, shows that the meta-analysis contained only 3 original studies directed at decreasing alcohol use in the general adult population. All other studies were either aimed at student populations and/or at reducing tobacco use.

A number of studies did report positive effects of online self-help alcohol modules in the general adult population. A recent meta-analysis [13] reported an overall medium effect size (g = 0.40, 95% CI 0.29 - 0.50) for 9 randomized controlled trials (RCTs), including the present study, on online self-help interventions targeted at reducing alcohol intake. Most of the modules in this meta-analysis are, however, fairly time consuming, ranging from one 90-minute session [14] to a 10-week program, and some require involvement of a therapist.

The present study examines the effectiveness of Drinktest, a single 10-minute online session in which tailored feedback is delivered, with no therapist involved. (See Multimedia Appendix 1 for a screenshot of Drinktest.nl.) The aforementioned meta-analysis on online self-help reported an overall effect size of g = 0.27 (95% CI 0.11 - 0.43) if only single-session personalized feedback interventions were included in the analysis [13]. Another meta-analysis by Riper and colleagues [15] identified 14 randomized controlled trials of personalized feedback, both online and offline, aimed at reducing alcohol intake. Jointly, these had a standardized mean difference (Cohen’s d) of 0.22 (95% CI 0.16 - 0.29). These effects are appreciable, especially when taking into account that no therapist’s time is involved. Personalized feedback is assumed to be more effective than general information due to two characteristics: (1) the information is perceived as more personal and, hence, more relevant, and (2) therefore, the recipient of the information pays more attention to the key message [16].

Besides Drinktest, only one other single-session Internet-based personalized feedback intervention (Check Your Drinking or CYD) was examined in an RCT directed at the general adult population [17]. The other single-session interventions included in the meta-analysis were offered in a work setting and an emergency department of a general hospital [13].

Drinktest was developed by the Netherlands Institute for Health Promotion and Disease Prevention (NIGZ). Drinktest offers brief personalized feedback regarding in an individual’s personal alcohol consumption patterns. The intervention consists of various components: overview of mean weekly alcohol intake, associated health risks, self-help guidelines to reduce alcohol intake, normative feedback to compare one’s own alcohol consumption to the level of one’s own cohort. A first version of Drinktest was found to effectively reduce alcohol intake in women but not in men [18]. Since problem drinkers (16.8% men versus 4.2% women [19]) and heavy drinkers (17.3% men versus 4.1% women [20]) are predominantly men, a second version of Drinktest was developed that was tailored to males. The present study thus offers evidence from a randomized controlled trial aimed at exploring the effectiveness of the revised Drinktest in adult males. It is hypothesized that more participants in the Drinktest condition will reduce their alcohol intake relative to those in the control condition in which a general psychoeducational brochure on alcohol is offered.

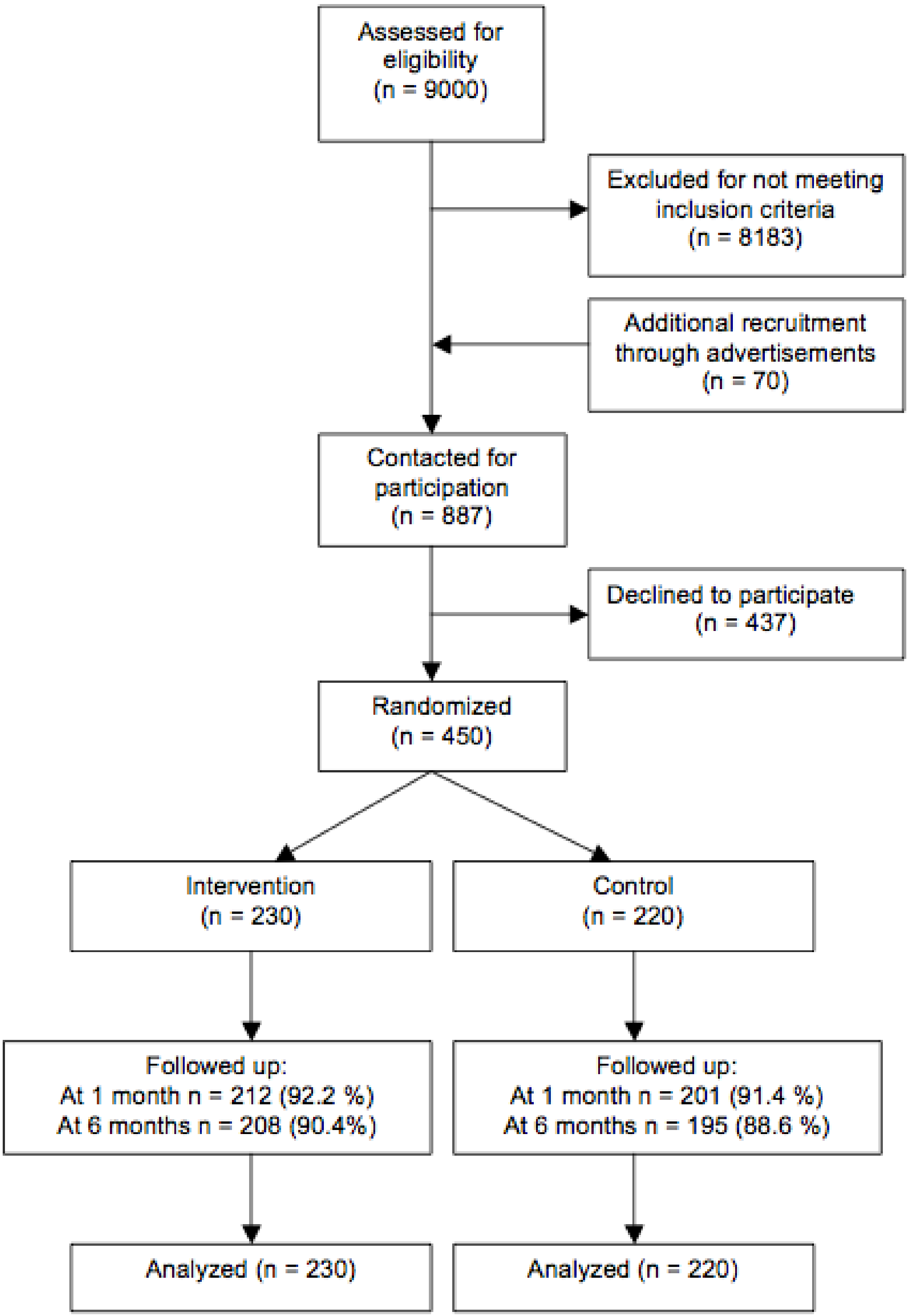
Method

Participants

A screening questionnaire was administered to all men aged 18 to 65 (n = 9000) in two nationally representative panels consisting of 25,000 households that can receive online questionnaires. Our questionnaire contained the Quantity-Frequency Variability index of alcohol intake (QFV) [21], the Dutch version [22]. All people whose alcohol consumption exceeded the threshold specified by the Dutch guideline for low-risk drinking were invited to take part in the study [23]. People exceeding this threshold qualify as heavy drinkers [24], that is, men who consumed more than 20 units of alcohol per week (heavy drinking) and/or more than 5 units of alcohol on a single occasion on at least 1 day per week (binge drinking), where 1 unit of alcohol is equal to 10 grams of ethanol. Men were not included in the study if they had received any professional help for alcohol-related problems or any medication to reduce alcohol consumption in the 12 months preceding the study.

In total, 817 men fulfilled the inclusion criteria and were willing to consider participation in the study. Additional participants were recruited through advertisements in national newspapers, to which 70 eligible men responded. All 887 men were contacted by telephone and asked to participate. After indicating their understanding that the research included a visit to the university, a total of 450 out of the 887 (50.7%) men contacted agreed to participate and gave informed consent. After 1 month, 413 participants were successfully followed-up. Of the 37 out of the 450 (8.2%) lost to follow-up, 2 had moved away and 35 did not respond. After 6 months, 403 participants were followed up successfully. Of the 47 out of the original 450 (10.4%) lost to follow-up, 4 had moved away, 41 did not respond, and 2 had died. Figure 1 presents the flow of participants through the study.

Figure 1. Flow chart



Procedure

At screening, participants were told a cover story to reduce the risk of response bias stemming from social desirability. Participants were told that they would judge newly developed educational materials addressing one of three possible life style topics: alcohol, smoking, or exercise. They were then told that they were randomly assigned to the alcohol group and that they were invited to evaluate the materials irrespective of their actual alcohol intake. They also received information on the procedure of the study, which consisted of one visit to the university and three written questionnaires at 0, 1, and 6 months to be filled in at home. Those responders who were eligible and willing to participate were contacted by telephone to explain to them again what participation would entail and to schedule the appointment at the behavioral laboratory. It was not revealed to the participants that their inclusion in the study was based on the degree of their alcohol intake. They were then randomized to either the computer-based personalized feedback (experimental condition) or the control group.

Prior to the appointment, participants received an informed consent form and a baseline questionnaire at their home address and were asked to bring both to their appointment. Written and signed consent was thus ensured. The baseline questionnaire included items measuring alcohol consumption and demographic characteristics.

On arrival, participants received a short standardized instruction on how to use the computer and Internet site (experimental condition) or how to read the leaflet (control condition). Participants were seated individually in a soundproof room for 20 minutes in order to ensure exposure to the educational materials and to reduce the effect of possible extra-experimental factors. Materials were offered to participants identical to real-life setting in order to maximize external validity of the study. Participants in the experimental condition completed the test online and were given the opportunity to make a printed copy of their personalized feedback. Participants in the control condition read the leaflet on paper in full color print and could take a copy home afterwards.

Next, participants received a short evaluation questionnaire with dummy questions to maintain the cover story and to determine participants' evaluation of the educational materials. They were then given the first part of their payment (€25).

At 1 month and at 6 months after studying the educational materials, participants received postal questionnaires. The first follow-up included questions about drinking behavior as well as determinants of alcohol consumption. The second follow-up also contained measures of alcohol consumption to assess effect maintenance. All items regarding alcohol consumption included in the measurements at different time points were identical so that comparisons could be made over time. After completing and returning the last follow-up, participants received the second payment (€25).

In order to minimize dropout rates, participants were sent a first reminder to return the follow-up questionnaire after 2 weeks, were sent a second reminder with a new questionnaire after another week, and were contacted by telephone for a third reminder if needed. After the last follow-up period had ended, participants received a standard letter explaining the true objectives of the study. The study was approved by the medical ethics committee of Erasmus Medical Centre in Rotterdam (reference number MEC-2006-343).

Intervention

Participants in the experimental condition received brief personalized feedback on alcohol use through the website www.drinktest.nl. The test is designed for adults who consume alcohol regularly or excessively and invites them to explore the possible negative consequences of their drinking behavior. The aim is 2-fold: prevention of heavy drinking and reduction of alcohol intake in heavy drinkers. In the first part of the test, respondents are asked to report their weekly alcohol consumption and number of binge drinking occasions and to indicate whether they think they consume too much alcohol and whether they intend to reduce their alcohol intake in the future. Based on this information, respondents receive the first part of the advice, which covers possible consequences of their drinking behavior. The first part of the feedback also includes a normative component in which participants can compare their alcohol consumption to that of others in the same age and gender bracket. Previous literature has revealed that including normative feedback in brief interventions aimed at reducing alcohol intake has favorable effects because people generally overestimate alcohol intake of others and underestimate their own alcohol intake [8,25]. In our study, all participants qualified as heavy drinkers; thus, all of them were invited to enter the second part of the test.

In the second part, the participants are asked questions concerning their drinking moments, drinking pattern, self-efficacy, attitude, and intention (behavioral stage according to the transtheoretical model [26]) with regard to reducing alcohol intake. Based on their answers, respondents receive personalized feedback on how to reduce alcohol intake in their specific situation. Completing the intervention takes about 10 minutes.

An early version of the intervention was designed in 2002, and its effectiveness was evaluated in a randomized controlled trial from 2005 to 2006 [18]. For the current study, a second and improved version of the website was used. Changes with regard to the former website were: (1) more tailoring to male drinking situations, since the former Drinktest was found to reduce alcohol consumption only in women [18]; and (2) the tailored advice was subdivided in smaller parts and ranked according to relevance for the participant. During the study period, March 2006 through June 2007, this new website was located in a secure area on the Internet. The first version of the website remained online, but within our study, the name of this website was never mentioned to avoid participants visiting it.

Control Condition

Participants in the control condition were given a standard brochure ("Facts About Alcohol" [23]) developed by the Netherlands Institute for Health Promotion and Disease Prevention (NIGZ) and were asked to read it carefully. The brochure contains factual information on the biological effects of alcohol, as well as on healthy and unhealthy drinking patterns.

Randomization

Randomization was conducted using a computer random number generator in the Statistical Package for the Social Sciences (SPSS), version 15.0. (SPSS Inc, Chicago, IL, USA). Randomization was stratified by age and educational level to ensure a good balance of these prognostically relevant characteristics of the participants across the experimental conditions. The condition to which participants were assigned was revealed to research assistants once recruitment was complete. All participants were blinded to assignment by providing them with a cover story (see "Procedure" section above).

Outcome Measures

The primary outcome measure was heavy drinking, defined as alcohol consumption exceeding the guidelines for low-risk drinking: an average of more than 20 alcohol units per week (excessive drinking) and/or more than 5 units on a single occasion on at least 1 day per week (binge drinking). Alcohol units per day per week were assessed with the Dutch version of the QFV [21,22]. Binge drinking was measured by asking respondents how often they drink more than 5 units of alcohol on a single occasion. Answering possibilities ranged from 0, *never* to 5, *every day*. In line with the Dutch guideline, the cutoff point to assess binge drinking was set on a score of 2, indicating a frequency of at least one session of binge drinking per week.

Power

The trial was powered to detect changes in alcohol consumption comparable to those found in a previous trial [18]. In the previous trial, a decrease in mean weekly alcohol consumption of 5.72 units (SD 15.9) was found in the experimental condition. In the control condition, participants decreased their weekly consumption with on average 1.05 units (SD 15.1). Using an alpha level of .05, 2-sided, and a power (1-beta) of .80, a total sample size of 348 men was required. To compensate for loss to follow-up, we recruited an additional 20% of the sample size. To detect changes in drinking patterns in completers-only analysis, a minimum number of 435 participants was thus needed at baseline. In our study 450 men participated.

Analysis

To check whether randomization had resulted in two comparable groups, logistic regression analysis was used with condition as the dependent variable and a set of possible confounders (among them age and level of education) as predictors. Following the CONSORT (consolidated standards of reporting trials) statement all our analyses were conducted in agreement with the intention-to-treat (ITT) principle while imputation was used to deal with loss to follow-up. Imputation of missing values was done using the expectation-maximization algorithm of Little and Rubin [27]. This is a general method of finding the maximum-likelihood estimate of the parameters of an underlying distribution from a given data set when the data has missing values [28]. The hypothesis of the study was tested using the chi-square test. We also computed the odds ratio (OR), and the number needed to treat (NNT) as the inverse of the risk difference (RD). The OR’s were obtained under a bivariate logistic regression model of response on condition and placed in their 95% confidence intervals. The RD was obtained under a linear probability model.

Furthermore, we repeated all the analyses described above for completers only and when imputation was carried out using the last observation carried forward method. Finally, in order to assess who benefited most from the intervention at the 6-month follow-up, interaction terms were computed by calculating the products of the intervention dummy (intervention versus control) with four dichotomous variables: (1) age (18-44 vs > 44), (2) education (high vs low, ie, academic or college degree versus lower levels), (3) weekly alcohol units at baseline (< 28 versus ≥ 28), and (4) binge drinking at baseline (at least once per week versus less frequently). We then entered these interaction terms together with the corresponding main effects into the logistic regression model.

All tests were conducted at alpha = .05 (two-sided) except for the check on randomization, which was done at alpha = .10 to ensure that also relatively small baseline differences between groups in terms of age and level of education would be detected. Data were analyzed using SPSS version 15.0.

Results

Participants’ Characteristics

Table 1 presents participants’ characteristics and primary outcome measures at baseline. Mean age of respondents was 40.4 (SD 15.1). Overall, most men had a high level of education, but this did not differ between the two conditions. Almost half of all respondents (214/450, 47.8%) indicated they were living with a partner, and the majority of the men reported being employed (253/450, 56.5%). Mean weekly alcohol consumption at baseline was equal across both groups, with 31 units for the experimental condition and 32 units for the brochure condition. No significant differences between conditions were found on demographic characteristics and drinking patterns (see Table 1).

Table 1. Participant characteristics and primary outcome measures at baseline

	Experimental Condition (n = 230)	Control Condition (n = 220)	Test Result	*P* Value
Age			$t_{448} = -0.17$	.87
Mean (SD)	40.6 (15.2)	40.3 (15.1)		
Education [a]			$\chi^2_2 = 1.2$	.54
Low, n (%)	36 (15.7)	43 (19.5)		
Medium, n (%)	70 (30.4)	62 (28.2)		
High, n (%)	124 (53.9)	115 (52.3)		
Living arrangement			$\chi^2_1 = 0.6$	.44
Living with partner, n (%)[b]	113 (49.6)	101 (45.9)		
Employment status			$\chi^2_1 = 0.02$	.89
Paid employment, n (%)[b]	128 (56.1)	125 (56.8)		
Weekly alcohol intake in standard units			$t_{448} = 0.53$	.60
Mean (SD)[c]	30.9 (19.2)	31.7 (14.3)		
Frequency of binge drinking			$t_{448} = 0.70$	.49
Mean (SD)[d]	2.1 (1.3)	2.2 (1.3)		

[a] Low = elementary or high school, medium = occupational certificate, high = university or college degree

[b] Includes 2 missing values

[c] A standard unit of alcohol contains 10 grams of ethanol

[d] Frequency of binge drinking defined as frequency of consuming more than 5 units of alcohol on at least one single occasion per week

Loss to follow-up after 1 month was 8.2% (37/450) and was evenly distributed across both conditions (n = 18 in the experimental condition and n = 19 in the control condition) ($\chi^2_1 = 0.98$, $P = .75$). Participants (n = 8) who returned their questionnaire after 6 months but not after 1 month were regarded as completers, and after 6 months, the total loss to follow-up was 10.4% (47/450) and again equally distributed across conditions (n = 22 for the experimental condition and n = 25 for the control condition) ($\chi^2_1 = 0.4$, $P = .53$).

Table 2 reports the participants’ characteristics at baseline for the completers and those lost to follow-up. Participants who were found to be lost to follow-up were significantly younger and more often single than completers.

Table 2. Comparison of the participants' characteristics at baseline between those successfully followed up and those lost during follow-up

	Followed up (n = 403)	Lost to Follow-up (n = 47)	Test Result	*P* Value
Age			$t_{448} = -3.0$	.003
Mean (SD)	41.2 (15.0)	34.2 (14.8)		
Education[a], n (%)			$\chi^2_2 = 0.87$	.65
Low, n (%)	70 (17.4)	9 (19.1)		
Medium, n (%)	116 (28.8)	16 (34.0)		
High, n (%)	217 (53.8)	22 (46.8)		
Living arrangement			$\chi^2_1 = 5.29$	.02
Living with partner, n (%)[b]	199 (49.6)	15 (31.9)		
Employment status			$\chi^2_1 = 0.63$	.43
Paid employment, n (%)[b]	229 (57.1)	24 (51.1)		
Weekly alcohol intake in standard units			t_{448}) = 1.7	.09
Mean (SD)[c]	30.8 (17.0)	35.2 (17.9)		
Frequency of binge drinking			$t_{448} = 1.1$	.29
Mean (SD)[d]	2.1 (1.3)	2.3 (1.2)		

[a] Low = elementary or high school, medium = occupational certificate, high = university or college degree

[b] Includes 2 missing values

[c] A standard unit of alcohol contains 10 grams of ethanol

[d] Frequency of binge drinking defined as frequency of consuming more than 5 units of alcohol on at least one single occasion per week

Outcomes

Table 3 reports the intervention effects on the primary outcome: the percentage of men who decreased their alcohol consumption to below the Dutch guideline for low-risk drinking (to score below this guideline, a person should not exceed the limits for both weekly alcohol intake, ie, > 20 alcohol units per week, as well as for binge drinking, ie, > 5 units on a single occasion on at least one day per week). At 1 month after studying the materials, significantly more participants in the experimental condition managed to cut down on their drinking to within the guideline norms than those in the leaflet condition (97/230 or 42.2% vs 67/220 or 30.5%) (OR = 1.7, 95% CI 1.13-2.46, NNT = 8.6, $\chi^2_1 = 6.7$, $P = .01$). These significant results were replicated under the completers-only and last observation carried forward imputation analyses (see Table 3).

Table 3. Change in success rates of adherence to the low-risk drinking guideline at the 1-month follow-up

	Experimental Condition		Control Condition						
	n	% Success	n	% Success	OR	95% CI	NNT[a]	χ^2	*P* Value
1-month follow-up									
Total sample EM[b] imputation	230	42.2	220	30.5	1.7	1.13-2.46	8.6	$\chi^2_1 = 6.7$	.01
Completers-only	207	41.5	195	29.2	1.7	1.14-2.60	8.1	$\chi^2_1 = 6.6$	.01
Total sample LOCF[c] imputation	230	37.4	220	25.9	1.7	1.14-2.56	8.7	$\chi^2_1 = 6.8$	.01
6-month follow-up									
Total sample EM[b] imputation	230	45.7	220	37.3	1.4	0.97-2.06	11.9	$\chi^2_1 = 3.3$	.07
Completers-only	195	44.1	188	36.7	1.3	0.90-2.05	13.5	$\chi^2_1 = 2.2$	.14
Total sample LOCF[c] imputation	230	37.4	220	31.4	1.3	0.88-1.93	16.7	$\chi^2_1 = 1.8$	.18

[a] NNT = numbers needed to treat

[b] EM imputation = imputation based on the expectation-maximization algorithm

[c] LOCF imputation = imputation based on last observation carried forward

At 6 months after studying the educational materials, even more participants in both conditions decreased their alcohol consumption to below the limits of heavy drinking, that is, 45.7% (105/230) in the experimental condition and 37.3% (82/220) in the control condition, but the difference between the conditions was no longer significant (χ^2_1 =3.3, *P* = .07) (Table 3). These findings were replicated under completers-only and last observation carried forward imputation analyses.

Predictors of Favorable Outcome

Analyses of the predictor-by-treatment interaction effects showed that favorable treatment response at the 6-month follow-up was not modified by any of the patient characteristics as measured at baseline, that is, age (Wald test = 3.03, df = 1, *P* = .25, OR = 0.64, 95% CI 0.30-1.37); level of education (Wald test = 0.15, df = 1, *P* = .70, OR = 1.16, 95% CI 0.54-2.49); weekly alcohol consumption at baseline (Wald test = 0.40, df = 1, *P* = .53, OR = 0.76, 95% CI 0.33-1.76); and binge drinking at baseline (Wald test = 3.857, df = 1, *P* = .05, OR = 2.35, 95% CI 1.00-5.49).

Discussion

Main Findings

The results of this study show that computer-based personalized feedback is successful in decreasing the percentage of male heavy drinkers in the short run. After 1 month, participants who received the intervention were more successful than controls in bringing down their alcohol consumption, even to within the guideline norms for low-risk drinking (42% versus 31%). However, after 6 months, the success rates were 46% versus 37% for the intervention and the control condition, respectively, and did not reach statistical significance, either under an intention-to-treat analysis or the completers-only analysis.

Our findings lend partial support to the idea that computer-based personalized feedback has a more favorable effect on the reduction of heavy drinking than a standard brochure on alcohol consumption, at least in the short-term. The initial effect of the intervention is further confirmed by the number needed to be treated (NNT), estimated at 8.6, which is comparable to, for example, the NNT found in Riper and colleagues (where NNT was equal to 8.5) in a more intensive intervention directed at the same population (ie, a general population >18 years of age) [10]. Although not all the heavy drinking men reduced their alcohol intake to within the low-risk drinking guideline, a substantial number of them did. Drinking below this guideline implies lower risks for the health-related problems and excess mortality due to heavy drinking, provided that effects are maintained over time [3,29]. Treatment response at the 6-month follow-up was not predicted by age, level of education, weekly alcohol consumption at baseline, or binge drinking level at baseline. This finding indicates that Web-based personalized feedback may be well suited to a heterogeneous group of heavy drinking men, as has previously been shown by Riper and colleagues [30].

These results appear to confirm previous findings on the effectiveness of online alcohol interventions [9,12,13]. Both Cunningham's Check Your Drinking and the Drinktest in the present study include normative feedback, a comparison of a person's drinking to others of the same age and sex, which is considered to be one of the effective components of personalized feedback [25].

Limitations and Strengths

The findings should be seen in the light of the limitations and strengths of our study. To begin, a substantial percentage of participants in both conditions decreased their alcohol consumption to below the Dutch guidelines for low-risk drinking. An explanation for this may be that compared with those not willing to participate in the study (dropout before randomization was 50%), participants may have been more motivated to change their alcohol intake. The fact that subjects thought they would evaluate educational materials irrespective of their actual alcohol intake, however, turns this into an unlikely explanation of the favorable effect. Also, the repeated alcohol questions participants were asked to fill out at different time points may have had an intervention-effect. Neither the possible high motivation of our subjects nor the intervention effect of the alcohol measures, however, explains the differences in alcohol consumption at follow-up between the intervention and control condition. Moreover, the effect size found in our study is similar to those reported in face-to-face inventions aimed at adults in the general population [13].

The overall loss to follow-up in our study was limited to 10%. This is a low percentage compared with the average loss to follow-up of 35% as found in meta-analyses on online alcohol interventions [11-13]. Moreover, dropout rates did not differ markedly between the intervention and control groups. We did see, however, that participants lost to follow-up were often younger and more often single. It is also possible that participants not followed up had higher alcohol consumption rates at the time of the follow-ups than those completing the study. Following the intention-to-treat principle, we handled respondents lost to follow-up in the analyses as stringently as possible by applying two different imputation techniques to estimate missing endpoints: one based on the expectation-maximization algorithm, the other conservatively based on last observation carried forward. In addition, we conducted completers-only analysis. Each time, we were able to replicate research findings. This attests to the robustness of our findings.

A limitation of this study is that we relied on self-reported measures. However, Del Boca and Darkes [31], in their review of the validity of self-reports of alcohol consumption concluded, “self-report measures have demonstrated reasonable levels of reliability and validity.” Also, we do not expect that any bias that may stem from self-report would be different in the experimental condition as compared with the control condition.

Due to the nature of the educational materials in our study, blinding of participants was not possible. This may have led respondents to underreport their alcohol consumption at follow-up. We tried to minimize the possibility of social desirability bias in the primary outcome by using a cover story. Respondents were not informed of the true objectives of the study until after the last follow-up questionnaires were received.

An additional drawback of the present study may be that the sample was limited to adult men, and the results may, therefore, not simply be generalized to women. However, restricting the present study to men was a deliberate choice, since in our previous study with an earlier version of the intervention we demonstrated that it was beneficial to women but not to men [18]. Moreover, heavy drinking is more prevalent in men [19,20].

Finally, it may be seen as a limitation of our study that we invited participants to visit our laboratory, which may threaten the external validity. This may, for example, have induced participants to spend all the available time actually reading the tailored advice, whereas if they were to conduct the intervention in their home environment, participants might not complete the entire intervention due to possible distractions or time constraints. Similarly, in a private setting, participants may not be able to ask for help, whereas in the laboratory situation a research assistant was present if they needed any support on the use of the computer and the Drinktest website. However, at this stage it was also a conscious choice to conduct a randomized controlled trial under laboratory conditions in order to assess the efficacy of the intervention after which an effectiveness study with a pragmatic randomized trial could be conducted. Furthermore, a recent meta-analysis [13] did not find any significant differences when comparing the effects of interventions based on where they were conducted (ie, research or health centre vs participants’ homes). Nevertheless, it cannot be ruled out that the laboratory setting may have helped to increase the effect of both the Drinktest and the brochure, but is unlikely to have induced a differential effect in this randomized trial favoring one condition over the other.

Conclusions

Personalized online feedback on alcohol appears to be an effective and fairly easy way to change unhealthy drinking patterns in adult men, at least in the short-term. Drinktest.nl yearly draws about 90,000 male visitors. Of these, 70% (63,000) report to be heavy drinkers, and 40% (25,200) of these heavy drinkers complete the test, implying that they actually receive the complete tailored feedback. Based on the NNT at the 6-month follow-up, assuming that the revised Drinktest will attract the same number and type of visitors each year and assuming that conducting Drinktest in a private setting would generate the same effects, this would imply that more than 2000 men per year (n = 2117) will successfully reduce their alcohol intake during at least 6 months as a consequence of spending 10 minutes of their time on Drinktest.nl. Offering personalized feedback on alcohol through highly accessible Internet sites may thus contribute to generating health gains at the population level in an efficient and economically affordable way. In fact, Smit and colleagues calculated that introducing evidence-based eHealth interventions such as Drinktest into the Dutch health care system would substantially improve the cost-effectiveness of the system for alcohol use disorders overall [32].

Acknowledgments

The study was funded by the Netherlands Health Research Council (ZonMw), Grant # 50-50110-98-235. We thank Lotte Ploegmakers, Anouk Vogelzang, Soenita Ganpat, Monique de Hoog, Esther Beekman, and Estelle Wienk for their help with the data collection and all participants for their contribution to this study. Also thanks to Rob Bovens, Odile Smeets, Anke Oenema, and Reinout Wiers for their helpful comments to previous versions of this manuscript as well as to Ingmar Franken for letting us use “his” behavioral lab at the Erasmus University Rotterdam.

Conflicts of Interest

None declared

Multimedia Appendix 1

Screenshot of Drinktest.nl

[PDF file (Adobe PDF), 102 KB - jmir_v13i2e43_app1.pdf]

References

1. World Health Organization. WHO Expert Committee on Problems Related to Alcohol Consumption, Second Report. Geneva, Switzerland: WHO technical report series, number 944; 2007. URL: http://www.who.int/substance_abuse/expert_committee_alcohol_trs944.pdf [accessed 2011-06-07] [WebCite Cache ID 5zGEO5E66]
2. World Health Organization. The World Health Report 2002: Reducing Risks, Promoting Healthy Life. Geneva, Switzerland: World Health Organization; 2002. URL: http://www.who.int/whr/2002/en/whr02_en.pdf [accessed 2011-06-07] [WebCite Cache ID 5zGEkEERC]
3. World Health Organization. Global Status Report on Alcohol 2004. Geneva, Switzerland: World Health Organization, Department of Mental Health and Substance Abuse; 2004. URL: http://www.who.int/substance_abuse/publications/global_status_report_2004_overview.pdf [accessed 2011-06-07] [WebCite Cache ID 5zGEtTs1y]
4. Holman CD, English DR, Milne E, Winter MG. Meta-analysis of alcohol and all-cause mortality: a validation of NHMRC recommendations. Med J Aust 1996 Feb 5;164(3):141-145. [Medline: 8628131]
5. 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]
6. Moyer A, Finney JW. Brief interventions for alcohol problems: factors that facilitate implementation. Alcohol Res Health 2004;28(1):44-50. [Medline: 19006991]
7. 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]
8. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
9. Vernon ML. A review of computer-based alcohol problem services designed for the general public. J Subst Abuse Treat 2010 Apr;38(3):203-211. [doi: 10.1016/j.jsat.2009.11.001] [Medline: 20015607]
10. 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]
11. Carey KB, Scott-Sheldon LA, Elliott JC, Bolles JR, Carey MP. Computer-delivered interventions to reduce college student drinking: a meta-analysis. Addiction 2009 Nov;104(11):1807-1819. [doi: 10.1111/j.1360-0443.2009.02691.x] [Medline: 19744139]
12. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
13. 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. [doi: 10.2196/jmir.1691]
14. Hester RK, Squires DD, Delaney HD. The Drinker's Check-up: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers. J Subst Abuse Treat 2005 Mar;28(2):159-169. [doi: 10.1016/j.jsat.2004.12.002] [Medline: 15780546]
15. Riper H, van Straten A, Keuken M, Smit F, Schippers G, Cuijpers P. Curbing problem drinking with personalized-feedback interventions: a meta-analysis. Am J Prev Med 2009 Mar;36(3):247-255. [doi: 10.1016/j.amepre.2008.10.016] [Medline: 19215850]
16. Kreuter MW, Oswald DL, Bull FC, Clark EM. Are tailored health education materials always more effective than non-tailored materials? Health Educ Res 2000 Jun;15(3):305-315 [FREE Full text] [Medline: 10977378]
17. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032. [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
18. Boon B, Huiberts A. Preventie problematisch alcoholgebruik. In: Meijer SA, Smit F, Schoemaker CG, Cuijpers P, editors. Gezond verstand. Evidence-based preventie van psychische stoornissen [Common sense. Evidence-based prevention of

mental disorders] Volksgezondheid Toekomstverkenning (VTV) Themarapport. Bilthoven, Netherlands: RIVM/Trimbos Institute; 2006:107-114.
19. Van Dijck D, Knibbe RA. De Prevalentie van Probleemdrinken in Nederland. Een Algemeen Bevolkingsonderzoek. Maastricht, Netherlands: Universiteit Maastricht; 2005.
20. Statline. Gezondheid, leefstijl, gebruik van zorg.: Centraal Bureau voor de Statistiek URL: http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=03799&D1=210-212&D2=0-17&D3=0&D4=a&VW=T [accessed 2010-11-15] [WebCite Cache ID 5zKr9DgQu]
21. Cahalan D, Cisin IH, Crossley HM. American Drinking Practices: A National Study of Drinking Behavior and Attitudes. New Brunswick, NJ: Rutgers Center of Alcohol Studies; 1969.
22. Lemmens P, Tan ES, Knibbe RA. Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. J Stud Alcohol 1992 Sep;53(5):476-486. [Medline: 1405641]
23. Van Bergen S, van Boxtel T, Canter C. Feiten over alcohol Facts about alcohol. Woerden, Netherlands: Nationaal Instituut voor Gezondheidsbevordering en Ziektepreventie (NIGZ); 2004.
24. World Health Organization. Lexicon of alcohol and drug terms published by the World Health Organization. URL: http://www.who.int/substance_abuse/terminology/who_lexicon/en/index.html [accessed 2010-11-03] [WebCite Cache ID 5txztWu0L]
25. Chan KK, Neighbors C, Gilson M, Larimer ME, Alan Marlatt G. Epidemiological trends in drinking by age and gender: providing normative feedback to adults. Addict Behav 2007 May;32(5):967-976. [doi: 10.1016/j.addbeh.2006.07.003] [Medline: 16938410]
26. Prochaska JO, DiClemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28:183-218. [Medline: 1620663]
27. Little RJA, Rubin DB. Statistical Analysis with Missing Data. New York, NY: John Wiley and Sons; 1987.
28. Tabachnick BG, Fidell LS. Using Multivariate Statistics. 5th edition. Boston, MA: Pearson/Allyn and Bacon; 2007.
29. Anderson P, Baumberg B. Alcohol in Europe: A Public Health Perspective. London, England: Institute of Alcohol Studies; 2006. URL: http://ec.europa.eu/health/ph_determinants/life_style/alcohol/documents/alcohol_europe.pdf [accessed 2011-06-09] [WebCite Cache ID 5zJIi8yot]
30. Riper H, Kramer J, Keuken M, Smit F, Schippers G, Cuijpers P. Predicting successful treatment outcome of web-based self-help for problem drinkers: secondary analysis from a randomized controlled trial. J Med Internet Res 2008;10(4):e46 [FREE Full text] [doi: 10.2196/jmir.1102] [Medline: 19033150]
31. Del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: state of the science and challenges for research. Addiction 2003 Dec;98 Suppl 2:1-12. [Medline: 14984237]
32. Smit F, Lokkerbol J, Riper H, Majo C, Boon B, Blankers M, et al. Modelling the cost-effectiveness of health care systems for alcohol use disorders. J Med Internet Res 2011. [doi: 10.2196/jmir.1694]

Abbreviations

CI: confidence interval
CONSORT: consolidated standards of reporting trials
CYD: Check Your Drinking
ITT: intention-to-treat
NIGZ: Netherlands Institute for Health Promotion and Disease Prevention
NNT: number needed to treat
OR: odds ratio
QFV: Quantity-Frequency Variability (index of alcohol intake)
RD: risk difference
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 23.11.10; peer-reviewed by J Cunningham, P Bendtsen; comments to author 08.12.10; revised version received 23.12.10; accepted 14.03.11; published 30.06.11.

Please cite as:
Boon B, Risselada A, Huiberts A, Riper H, Smit F
Curbing Alcohol Use in Male Adults Through Computer Generated Personalized Advice: Randomized Controlled Trial
J Med Internet Res 2011;13(2):e43
URL: http://www.jmir.org/2011/2/e43/
doi:10.2196/jmir.1695
PMID:21719412

Original Paper

A Smartphone Client-Server Teleradiology System for Primary Diagnosis of Acute Stroke

J Ross Mitchell[1,2,3], PhD; Pranshu Sharma[1], MD; Jayesh Modi[1], MD; Mark Simpson[1], MSc; Monroe Thomas[3], BSC; Michael D Hill[2], MD FRCPC; Mayank Goyal[2], MD FRCPC

[1]Imaging Informatics Lab, Department of Radiology, University of Calgary, Calgary, AB, Canada
[2]Department of Clinical Neurosciences, Calgary, AB, Canada
[3]Calgary Scientific Incorporated, Calgary, AB, Canada

Corresponding Author:
J Ross Mitchell, PhD
Imaging Informatics Laboratory
Department of Radiology
Heritage Medical Research Building, Room 372
3330 Hospital Drive NW
Calgary, AB, T2N 4N1
Canada
Phone: 1 403 210 6667
Fax: 1 403 210 6667
Email: rmitch@ucalgary.ca

Abstract

Background: Recent advances in the treatment of acute ischemic stroke have made rapid acquisition, visualization, and interpretation of images a key factor for positive patient outcomes. We have developed a new teleradiology system based on a client-server architecture that enables rapid access to interactive advanced 2-D and 3-D visualization on a current generation smartphone device (Apple iPhone or iPod Touch, or an Android phone) without requiring patient image data to be stored on the device. Instead, a server loads and renders the patient images, then transmits a rendered frame to the remote device.

Objective: Our objective was to determine if a new smartphone client-server teleradiology system is capable of providing accuracies and interpretation times sufficient for diagnosis of acute stroke.

Methods: This was a retrospective study. We obtained 120 recent consecutive noncontrast computed tomography (NCCT) brain scans and 70 computed tomography angiogram (CTA) head scans from the Calgary Stroke Program database. Scans were read by two neuroradiologists, one on a medical diagnostic workstation and an iPod or iPhone (hereafter referred to as an iOS device) and the other only on an iOS device. NCCT brain scans were evaluated for early signs of infarction, which includes early parenchymal ischemic changes and dense vessel sign, and to exclude acute intraparenchymal hemorrhage and stroke mimics. CTA brain scans were evaluated for any intracranial vessel occlusion. The interpretations made on an iOS device were compared with those made at a workstation. The total interpretation times were recorded for both platforms. Interrater agreement was assessed. True positives, true negatives, false positives, and false negatives were obtained, and sensitivity, specificity, and accuracy of detecting the abnormalities on the iOS device were computed.

Results: The sensitivity, specificity, and accuracy of detecting intraparenchymal hemorrhage were 100% using the iOS device with a perfect interrater agreement (kappa = 1). The sensitivity, specificity, and accuracy of detecting acute parenchymal ischemic change were 94.1%, 100%, and 98.09% respectively for reader 1 and 97.05%, 100%, and 99.04% for reader 2 with nearly perfect interrater agreement (kappa = .8). The sensitivity, specificity, and accuracy of detecting dense vessel sign were 100%, 95.4%, and 96.19% respectively for reader 1 and 72.2%, 100%, and 95.23% for reader 2 using the iOS device with a good interrater agreement (kappa = .69). The sensitivity, specificity, and accuracy of detecting vessel occlusion on CT angiography scans were 94.4%, 100%, and 98.46% respectively for both readers using the iOS device, with perfect interrater agreement (kappa = 1). No significant difference ($P < .05$) was noted in the interpretation time between the workstation and iOS device.

Conclusion: The smartphone client-server teleradiology system appears promising and may have the potential to allow urgent management decisions in acute stroke. However, this study was retrospective, involved relatively few patient studies, and only two readers. Generalizing conclusions about its clinical utility, especially in other diagnostic use cases, should not be made until additional studies are performed.

(J Med Internet Res 2011;13(2):e31) doi:10.2196/jmir.1732

KEYWORDS

Acute stroke; teleradiology; computed tomography; mhealth; mobile phone

Introduction

Recent advances in the treatment of acute ischemic stroke have made rapid acquisition, visualization, and interpretation of images a key factor for positive patient outcomes. Some teleradiology systems can accelerate image interpretation and reduce treatment delays [1-3]. However, teleradiology systems have three important features that limit their utility for acute stroke diagnosis. First, the systems may require that patient digital imaging and communications in medicine (DICOM) images be transferred to a remote device for viewing before interpretation can begin. Acute stroke imaging exams often contain several hundred megabytes (MB) of data. The delay caused by transmission of the image data may significantly reduce treatment effectiveness. In addition, if there is confidential information in the headers of the DICOM files, then meeting security requirements may cause additional delays and/or restrict the locations where the remote device may be located. Second, the devices used for remote visualization may limit or inconvenience physician mobility. This may result in additional delays while a physician on call travels to the nearest remote visualization device. Third, mobile devices may lack the computational capabilities to perform advanced visualizations that can aid the diagnostic process. In turn, this may prevent mobile devices from achieving functional equivalence to workstation systems and, therefore, make regulatory approval of mobile devices for this purpose more difficult.

Client-server based teleradiology systems have been described in the literature previously [4,5]. However, these are designed primarily to communicate with workstation class devices. There are no reports in the literature of interactive, streaming, client-server systems that can provide sufficient functionality, image quality, and frame rates to a current generation smartphone device to allow for primary diagnosis.

We have developed a new teleradiology system based on a client-server architecture to try to address these limitations. As a remote visualization device, our system can use an Apple (Apple Inc, Cupertino, CA) iPhone, iPod Touch, iPad (hereafter, referred to as iOS device), or a device running Android (Google Inc, Mountain View, CA) version 2.1 or newer equipped with a touch screen. When using our system, DICOM images are not transferred to the remote device. Instead, the server loads and renders the patient DICOM images, then transmits a rendered frame to the remote device. This process can occur within a few seconds and allows a remote physician to view the first frame and begin interpretation quickly. The server can be placed in a secure location, and all transmissions can occur using standard protocols over secure connections. In addition, since a server is performing the rendering, advanced visualization methods not possible on the remote device, such as three-dimensional volume rendering, may be used to generate frames for remote viewing.

Here, we report results from an initial feasibility study to determine if the new system is capable of providing accuracies and interpretation times sufficient for diagnosis of acute stroke from computerized tomography (CT) brain scans viewed using an iOS device.

Methods

One hundred and twenty recent consecutive noncontrast computed tomography (NCCT) brain scans and 70 computed tomography angiogram (CTA) head scans were obtained from the Calgary Stroke Program database. The Calgary Stroke Program is recognized as one of the leading programs in North America for stroke treatment and research.

NCCT Brain Scans

NCCT brain scans were obtained using sequential acquisition of data on a 64-row multidetector CT (Somatom Sensation 64, Siemens Healthcare, Germany) from the foramen magnum to the skull vertex using a 5-mm slice thickness. NCCT brain scans were evaluated for early signs of infarction, which includes early parenchymal ischemic changes and dense vessel sign, and to exclude acute intraparenchymal hemorrhage and any stroke mimics, such as tumor, infective/inflammatory disease, or any vascular malformation. Acute parenchymal ischemic changes were graded according to the Alberta Stroke Program Early CT Score (ASPECTS) scoring system [6]. ASPECTS is a 10-point scale that grades the extent of ischemic change within the territory of the middle cerebral artery.

CTA Head Scans

CTA scans of the head and neck were performed on 64-row multidetector CT (Somatom Sensation 64, Siemens Healthcare, Germany). Data were acquired from the ascending aorta to the skull vertex using a standard spiral acquisition after infusion of 100 ml of nonionic iodinated contrast. Axial slices of 1-mm thickness were obtained from the aortic arch to the skull base, while 0.6-mm axial slices were obtained from the skull base to the skull vertex. All images had a 220-mm field of view and contained 512 x 512 pixels. Reformations with 3-mm slices were performed in the axial, sagittal, and coronal planes for the head region and reviewed on a workstation. For the purpose of this study, only axial CTA head scans were used for interpretation on the iOS device. CTA head scans were evaluated for any intracranial vessel occlusion.

The scans were read by two neuroradiologists. One reader (author PS) interpreted scans first on a workstation located in a radiology reading room equipped with a medical-grade display (hereafter, referred to as the workstation) and then on the iOS device. The second reader (author JM) interpreted scans only on the iOS device. To avoid bias, a delay of 2 weeks was allowed between interpretation on the workstation and interpretation on the iOS device. The readers were blinded to the patients' clinical data. Furthermore, the patient exams were presented in different orders on the workstation and on the client-server teleradiology system.

The workstation (IMPAX 6.3.1.3815, Agfa Healthcare, Belgium) was connected to a medical-grade 21-inch liquid crystal display (MD21GS-3MP, NEC). This display has a resolution of 2048 x 1536 pixels (pixel pitch = 0.21 mm) and a luminance of 400 (candela) cd/m^2. All interpretations on the medical workstation were performed in a lighting-controlled (darkened) radiology reading room. All patient imaging exams were then anonymized and loaded on the server to be analyzed on the iOS device. The iOS device used in this study had a 3.5-inch diagonal screen having a resolution of 320 x 480 pixels (pixel pitch = 0.15 mm) and a luminance = 500 cd/m^2. All interpretations on the iOS device were performed under normal office lighting conditions in a room with overhead florescent light panels.

The time to interpret each exam was recorded by the readers themselves using a digital stopwatch. When using the iOS device, the interpretation time included the time required to launch the application, establish a connection to the server, select the study for interpretation, and perform the interpretation. When using the workstation, the recorded time included the time to select the study and perform the interpretation.

The Server Configuration

The visualization server had a 2.4 gigahertz (GHz) Intel Core 2 Quad Core central processing unit, 8 GB RAM and two NVIDIA GeForce 8800 (512 MB) graphics cards. It ran Red Hat Enterprise Linux 5 (Red Hat Inc, Raleigh, NC) and included the application ResolutionMD Enterprise (Calgary Scientific Inc, Calgary, Canada). ResolutionMD Enterprise (hereafter referred to as the server software) allows interactive 2-D and 3-D visualization of DICOM images on remote displays via secure hypertext transfer protocol (http). Visualization is initiated by a remote user, who connects to a particular uniform resource locator (URL) address on the server using a Web browser. The server will perform different actions depending upon the URL specified. For example, the visualization server might send a small Flash (Adobe Systems Inc, San Jose, California) or Silverlight (Microsoft Inc, Redmond, Washington) client program to the Web browser. This client program is executed by the browser and used to implement the remote user interface and manage communication with the server.

Initially, the user interacts with the client program to choose a DICOM series to visualize. Typically, the DICOM images are resident on a picture archiving and communication system (PACS). The server then loads the series of 2-D DICOM images from the PACS into memory and reformats them into a 3-D volume. It then performs a rendering operation on the 3-D volume to produce a 2-D frame, which is compressed using a lossy joint photographic experts group (JPEG) algorithm and encoded for transmission to the remote client for display. The JPEG standard includes a user definable "quality factor" that varies between 1 and 100, where 100 is the highest possible lossy JPEG encoding quality. In our study, the quality factor was set to 25 during interactive image presentation. During static image presentation, the quality factor was automatically set to 100. The quality factor has a variable effect on the achievable compression ratio for neurological CT images [7]. Compression ratios were not measured in our studies. However, previous research indicates that for neurological CT scans, a quality factor of 85 produces a compression ratio of 10:1, while a quality factor of 92 produces a compression ratio of 8:1 [7]. These compression ratios are well within published guidelines that suggest a maximum compression ratio of 12:1 for (static) neurological CT scans [8].

All quantitative operations (for example, adjusting the image intensity window and level) are performed by the server on the original DICOM data. The results are then transmitted interactively to the remote client, either as a new encoded frame, or as updated textual information. DICOM files are not transmitted to, or stored on, the client. Nevertheless, the server-side rendered image is stored temporarily in client-side volatile random-access memory (RAM) in both JPEG and decompressed image formats. Any confidential patient information (eg, patient name) that has been rendered into the image pixels is thus temporarily present on the client. Some textual attributes from the originating image's DICOM file may also be temporarily stored in volatile RAM on the client. These attributes may be used as navigational and informational displays within the client user interface. When the user exits the Web browser or moves to a new webpage, the connection to the server is closed, the rendering operation ends, a blank image is displayed on the client, and the volatile RAM on the client is cleared.

The server software provides a number of visualization protocols, including 2-D, 2-D side-by-side, linear and curvilinear multiplanar reformatting, and several 3-D volume rendering modes. At this level, the system has similarities with others that have been described in the literature [9]. However, unlike previous systems, our server software includes two important novel and proprietary enhancements. The first of these allows more efficient use of available bandwidth. The server software decomposes the sequence of 2-D frames into static and dynamic components. To do this, the server keeps track of the last image sent to the client before compression. When the next image is to be sent, the server compares 8 x 8 blocks within the new image with corresponding blocks of the previous image. Only blocks that contain changes in pixel data between the two images are compressed and transmitted to the client. The second enhancement allows Web-client applications with sophisticated and complex user interfaces. The system provides server- and client-side application program interfaces (APIs) and libraries for maintaining and synchronizing a hierarchical state model between the server and one or more clients, without requiring tight coupling of compatible data types, data structures, and programming languages. Efficiency is achieved by ensuring that the internal state model can be mapped onto the extended markup language (XML), which can be easily manipulated such that only the differences in the state between a client and server are ever transmitted. This allows the application state to be managed efficiently and shared and synchronized simultaneously with multiple clients. In turn, this permits simultaneous collaboration between multiple users across large distances and diverse networks.

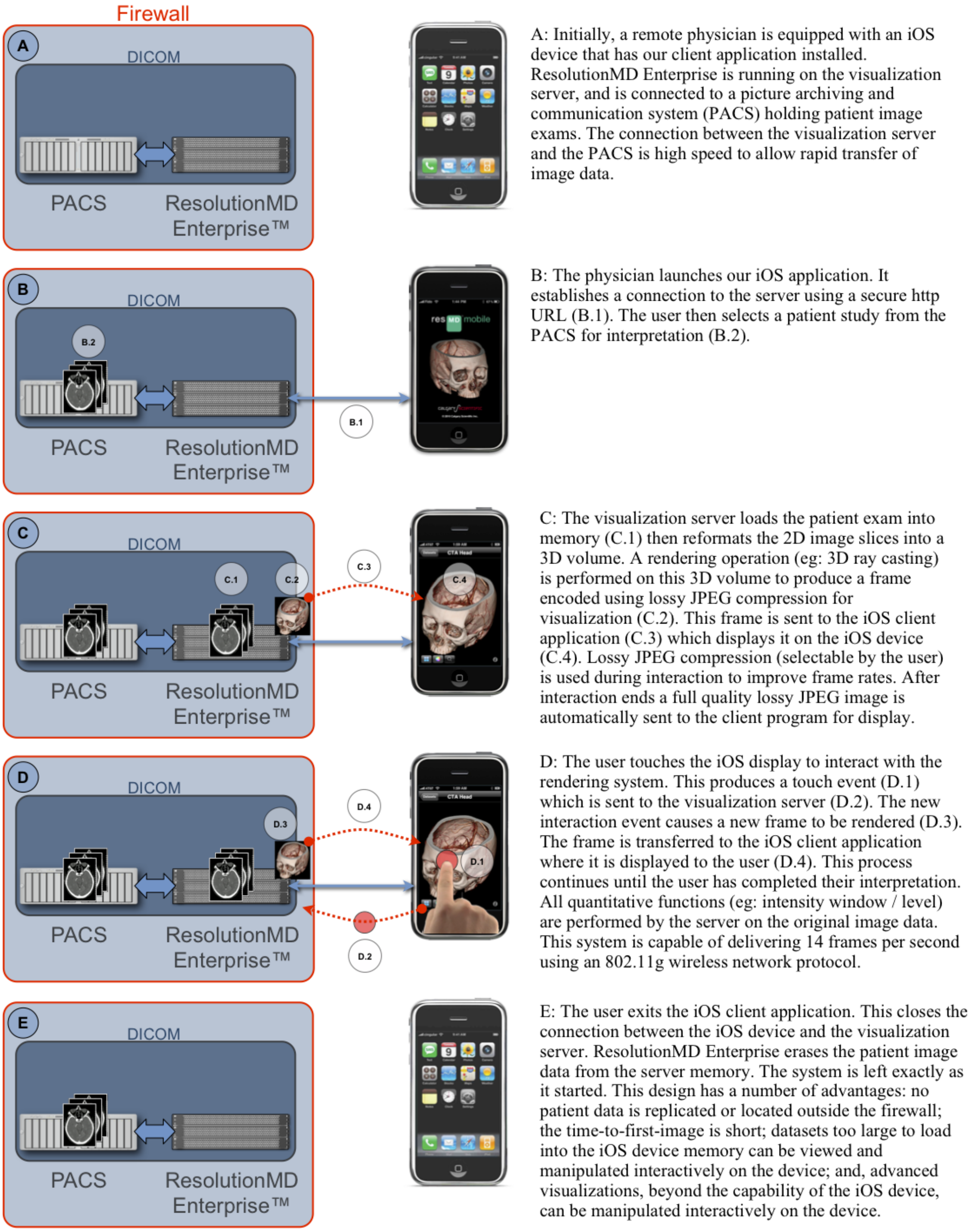
This framework also allows a single client application to connect to multiple servers, then observe and manipulate the images and states of those servers. In turn, this allows one to “mix and match” components of multiple server-side applications into a single client-side application.

The Client Configuration

The iOS device (iPod or iPhone) used in our experiments included 8 gigabytes (GB) of flash memory and an operating system version 3.1.2. The Safari Web browser application available on iOS devices is not capable of running Flash or Silverlight programs. Therefore, we developed a custom application for iOS devices to implement a remote user interface and manage communication with the server. This software was written in Objective-C 2.0 using the Xcode 3.0 integrated development environment. Our application was installed on the iOS device used in this experiment using Apple Inc’s ad-hoc distribution method. This process allows one to install and test applications without having them released on Apple Inc’s online iOS device application store. The version of our iOS device application used in these experiments only exposed some of the advanced visualization modes available in the server software. It allowed 2-D and 3-D visualizations, interactive window/level, translation, rotation, and zoom. In 3-D mode, the user could also select from a range of tissue rendering settings.

Once launched, the custom iOS device application initiates communication with a remote server running the server software (Figure 1). Users can use default server addresses provided with the iOS device application, or they can enter and save custom server addresses. The iOS device application then captures user interaction events, communicates these to the server, and manages the sequence of encoded frames for visualization. We set the default server address to point to the server used in our experiments. This eliminated the need for the user to specify an address, thereby reducing the delay before image interpretation. Communication between the visualization server and the iOS device occurred over a secure wireless network (Wi-Fi 802.11g). Communication via https to the server over third generation (3G) cellular networks is also possible on the iOS device 3GS. However, preliminary experimentation suggested that frame rates over 3G cellular networks are insufficient for practical use at this time. Consequently, cellular networks were not tested during this study. An overview of the iOS device system is provided in Figure 1.

Figure 1. A schematic overview of the client-server teleradiology system



A: Initially, a remote physician is equipped with an iOS device that has our client application installed. ResolutionMD Enterprise is running on the visualization server, and is connected to a picture archiving and communication system (PACS) holding patient image exams. The connection between the visualization server and the PACS is high speed to allow rapid transfer of image data.

B: The physician launches our iOS application. It establishes a connection to the server using a secure http URL (B.1). The user then selects a patient study from the PACS for interpretation (B.2).

C: The visualization server loads the patient exam into memory (C.1) then reformats the 2D image slices into a 3D volume. A rendering operation (eg: 3D ray casting) is performed on this 3D volume to produce a frame encoded using lossy JPEG compression for visualization (C.2). This frame is sent to the iOS client application (C.3) which displays it on the iOS device (C.4). Lossy JPEG compression (selectable by the user) is used during interaction to improve frame rates. After interaction ends a full quality lossy JPEG image is automatically sent to the client program for display.

D: The user touches the iOS display to interact with the rendering system. This produces a touch event (D.1) which is sent to the visualization server (D.2). The new interaction event causes a new frame to be rendered (D.3). The frame is transferred to the iOS client application where it is displayed to the user (D.4). This process continues until the user has completed their interpretation. All quantitative functions (eg: intensity window / level) are performed by the server on the original image data. This system is capable of delivering 14 frames per second using an 802.11g wireless network protocol.

E: The user exits the iOS client application. This closes the connection between the iOS device and the visualization server. ResolutionMD Enterprise erases the patient image data from the server memory. The system is left exactly as it started. This design has a number of advantages: no patient data is replicated or located outside the firewall; the time-to-first-image is short; datasets too large to load into the iOS device memory can be viewed and manipulated interactively on the device; and, advanced visualizations, beyond the capability of the iOS device, can be manipulated interactively on the device.

Analysis

The interpretations on an iOS device were compared to those of a workstation. Readings of reader 1 (author PS) on the workstation were considered ground truth for comparison. Any differences between the iOS device and workstation platforms in detecting stroke mimics or acute intraparenchymal hemorrhage, acute parenchymal ischemic changes, hyperdense vessel sign on NCCT brain scans and intracranial vessel occlusion on CTA brain scans were considered errors. If the difference in APSECTS scoring was more than one, it was considered a discrepancy in reading. Interrater agreement between the readers was assessed by calculating kappa using Stata 10.0 (StataCorp, College Station, Texas, USA). True positives, true negatives, false positives, and false negatives were obtained, and sensitivity, specificity, and accuracy of detecting the abnormalities on the iOS device were computed. The mean interpretation times on the workstation and iOS device were compared.

Results

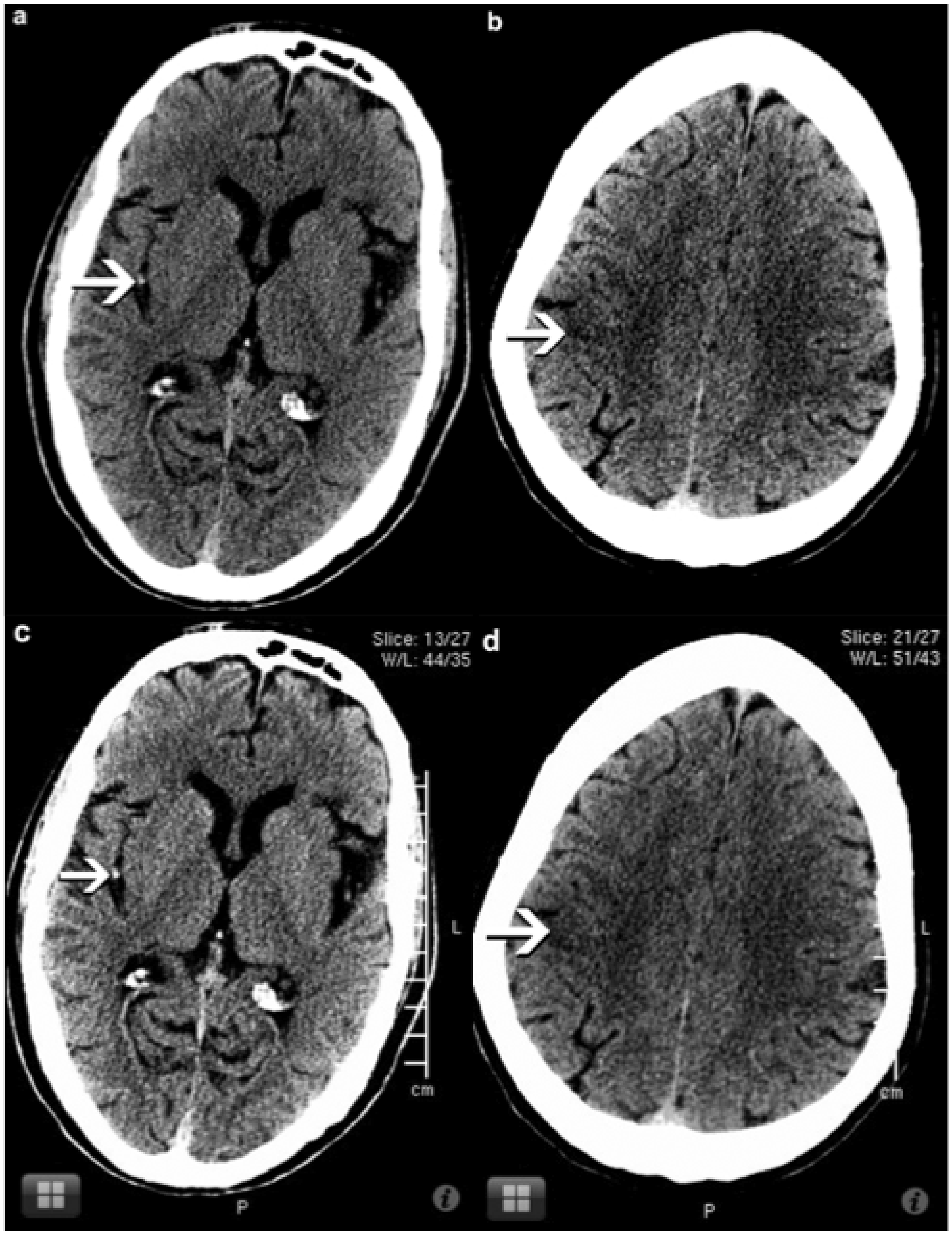
NCCT Brain Evaluation

Of the 120 NCCT brain scans, poor image quality (from patient motion during scanning) resulted in 8 patient scans being excluded leaving 112 scans for review. There were no stroke mimics identified on the workstation or the iOS device. Thus, none were falsely diagnosed as stroke on the iOS device.

Acute intracranial hemorrhage (ICH) was detected using the workstation in 7 of the 112 patients. The remaining 105 patients were then assessed for early ischemic changes. Both readers correctly diagnosed intracranial hemorrhage in all 7 patients. The sensitivity, specificity, and accuracy of detecting hemorrhage on NCCT brain scans were 100% using the iOS device with perfect interrater agreement (kappa = 1).

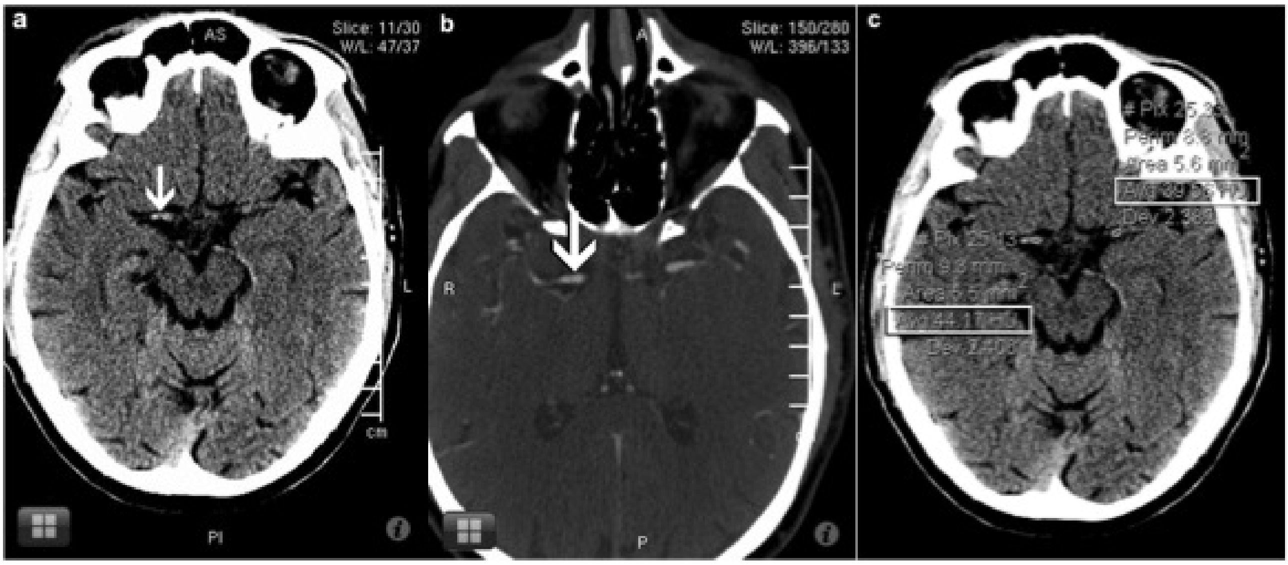
Acute parenchymal ischemic changes were seen when using the workstation in 34 of 105 patients. The sensitivity, specificity, and accuracy of detecting acute parenchymal ischemic change (Figure 2) were 94.1%, 100%, and 98.09% respectively for reader 1 and 97.05%, 100%, and 99.04% for reader 2 using the iOS device. There was nearly perfect interrater agreement (kappa = .8) between the readers.

Figure 2. In image *a*, the workstation monitor images show a dense MCA sign (arrow) and in image *b*, acute ischemic change in the ASPECTS M6 region (arrow). Corresponding iOS device images show a dense MCA sign (arrow) in image *c* and acute ischemic change in ASPECTS M6 region (arrow) in image *d*.



Dense vessel sign was detected in 18 of 105 patients with acute ischemic changes seen on the workstation. There were 4 false positives by reader 1, and 5 false negatives by reader 2 in diagnosing dense vessel sign on the iOS device. The sensitivity, specificity, and accuracy of detecting dense vessel sign (Figures 2 and 3) on NCCT brain scan were 100%, 95.4%, and 96.19% respectively for reader 1 and 72.2%, 100%, and 95.23% for reader 2 using the iOS device. There was good interrater agreement (kappa = .69) between the readers.

Figure 3. On an iOS device, an NCCT brain scan shown in image *a* was incorrectly interpreted as having a dense MCA sign (arrow). On an iOS device, a CT angiography brain scan (image *b*) demonstrates a normal patent vessel. On the workstation monitor, an NCCT brain image *c* shows nearly the same Hounsfield units (open rectangles) of both vessel segments, thus avoiding the error made on the iOS device.



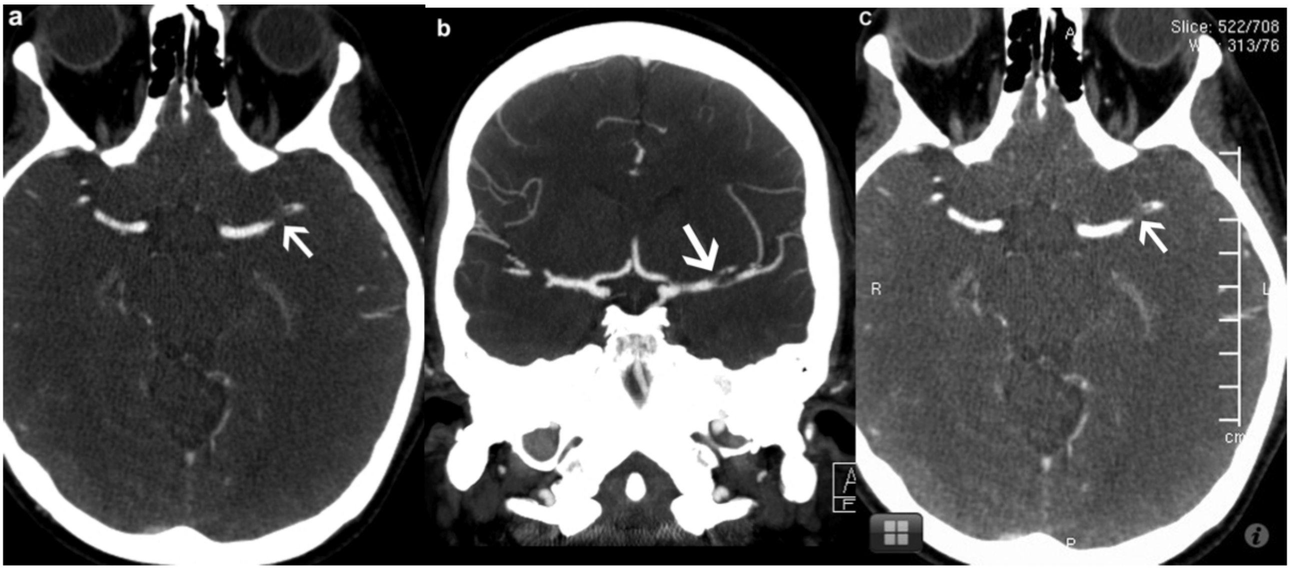
The mean time to interpret NCCT exams on the workstation was 2.1 minutes (SD 0.77), while the mean time using the iOS device was 2.7 minutes (SD 0.9) for reader 1 and 2.3 minutes (SD 1.4) for reader 2. These mean times were not statistically significantly different from each other at *P* > .05. The interpretation time ranged between 1 and 6 minutes on both platforms.

CT Angiogram Evaluation

Of the 70 CTA head scans, patient motion and/or poor contrast opacification of vessels resulted in the exclusion of 5 poor quality CTA exams, leaving 65 exams for our review. Vessel occlusion was detected in 18 of the 65 patients on the workstation.

Both readers correctly diagnosed vessel occlusion (Figure 4) in 17 of 18 patients on the iOS device. However, both readers missed an occlusion of the V4 segment of the vertebral artery using our system. All patients with patent intracranial vessels were correctly diagnosed using our system without any false-positive errors.

Figure 4. The workstation monitor axial and coronal images *a* and *b* show acute thrombus (arrow) in the proximal segment of the left MCA. Corresponding iOS device axial image *c* shows acute thrombus (arrow) in the proximal segment of the left MCA.



The sensitivity, specificity, and accuracy of detecting vessel occlusion on CT angiography scans were 94.4%, 100%, and 98.46% respectively for both readers using the iOS device. There was perfect interrater agreement (kappa = 1) between the readers.

The mean time to interpret CTA exams on the workstation was 3.5 minutes (SD 1.2), while using the iOS device the mean time was 3.63 minutes (SD 1.48) for reader 1 and 3.83 minutes (SD 0.83) for reader 2. These mean times were not statistically significantly different from each other at *P* > .05. The interpretation time ranged between 1 and 7 minutes on both platforms.

True positives, true negatives, false positives, and false negatives and sensitivity, specificity, and accuracy of detecting the abnormalities on the iOS device by both readers are shown in Tables 1 and 2.

Table 1. Sensitivity, specificity, and accuracy of stroke diagnosis using the client-server teleradiology system by reader 1

	Work-station (reader 1)	iOS Device (reader 1)	TP[a]	FP[b]	TN[c]	FN[d]	Sensitivity	Specificity	Accuracy
Hemorrhage	7/112	7/112	7	0	105	0	100%	100%	100%
Acute parenchymal change	34/105	32/105	32	0	71	2	94.11%	100%	98.09%
Dense vessel sign	18/105	22/105	18	4	83	0	100%	95.4%	96.19%
Vessel occlusion on CTA	18/65	17/65	17	0	47	1	94.4%	100%	98.46%

[a] True positive on the iOS device
[b] False positive on the iOS device
[c] True negative on the iOS device
[d] False negative on the iOS device

Table 2. Sensitivity, specificity, and accuracy of stroke diagnosis using the client-server teleradiology system by reader 2

	Work-Station (reader 1)	iOS device (reader 2)	TP[a]	FP[b]	TN[c]	FN[d]	Sensitivity	Specificity	Accuracy
Hemorrhage	7/112	7/112	7	0	105	0	100%	100%	100%
Acute parenchymal change	34/105	33/05	33	0	71	1	97.05%	100%	99.04%
Dense vessel sign	18/105	13/105	13	0	87	5	72.2%	100%	95.23%
Vessel occlusion on CTA	18/65	17/65	17	0	47	1	94.4%	100%	98.46%

[a] True positive on the iOS device
[b] False positive on the iOS device
[c] True negative on the iOS device
[d] False negative on the iOS device

Discussion

Acute ischemic stroke is the most common form of stroke, and it is also the most treatable. For each minute of acute brain ischemia, 1.9 million neurons are destroyed [10]. Treatment with thrombolysis is highly time dependent and entirely dependent upon quick imaging to make an inclusive diagnosis. Equally, expertise remains limited. Any method that can reduce the time from image acquisition to expert review and decision is beneficial. It is clear that every 15-minute delay results in a measureable reduction in the probability of good outcome after thrombolysis [11]. Hence, rapid image visualization by an expert wherever that expert is is a key factor in improving patient outcomes.

We developed a teleradiology system based on a client-server architecture that enables rapid access to radiological images on a current generation smartphone device. Through this system, a physician can securely assess remote imaging wherever a cellular or wireless network is available, which allows urgent management decisions to be made.

It is important to exclude any intracranial hemorrhage, as this is a contraindication to thrombolytic agents. A previous study has shown that acute intracranial hemorrhage is detected with high accuracy and interrater reliability (kappa = .87 - .94) on NCCT brain scans [12]. In our study, the presence of intraparenchymal hemorrhage was accurately diagnosed with perfect interrater reliability on NCCT brain scans reviewed on an iOS device. Our experience reflects that of Toomey et al [13], who found a personal digital assistant to be accurate in diagnosing acute intraparenchymal hemorrhage when compared with a workstation.

The extent of early ischemic changes in the parenchyma has been correlated with poorer clinical outcomes and increased risk of hemorrhage [14]. Previous meta-analysis has shown mean sensitivity of 66% (range 20% - 87%) and mean specificity of 87% (range 56% - 100%) with varying interrater agreement (kappa = .14 - .78) to detect an early infarction sign on NCCT brain scans [15]. Studies have been published demonstrating good sensitivity and specificity in detecting early ischemic signs using the ASPECTS scoring system [16,17]. ASPECTS is a well-validated scoring system that has good interrater reliability (kappa .71 - .89) [6,16]. In our study, there was high accuracy with nearly perfect interrater agreement (kappa = .8) in detection of early ischemic parenchymal changes on an iOS device. There were no false-positive diagnoses of early ischemic parenchymal changes by either of the readers using our system. Importantly, no large infarcts were missed by either of the readers using our system.

Hyperdense vessel sign on NCCT brain scan is a marker of intraluminal thrombus [18,19]. Studies have shown that hyperdense vessel sign is a highly specific but only moderately sensitive indicator of thromboembolic occlusion [18] with a wide range of interrater reliability (kappa = .36 - 1.00) [15]. False-positive hyperdense sign can be seen in calcified atherosclerotic vessel or a high hematocrit [20]. In this study, there were discrepancies in the detection of dense vessel signs on the iOS and the workstation with average interrater agreement (kappa = .69). There were 4 false positives by reader 1, and 5 false negatives by reader 2 in detecting hyperdense vessel sign on the iOS device. We feel that measuring the density in Hounsfield units and comparing with the opposite side vessel helps reduce false-positive and false-negative diagnoses rather than relying solely on the observed local CT image contrast. The workstation software allowed the reader to analyze and display the CT scan Hounsfield units (Figure 3). The server software does provide functions to analyze Hounsfield units. However, the iOS client software used in this study did not expose this functionality to the user. It was subsequently added to the version licensed by Health Canada.

Diseases that can mimic stroke such as tumors, infection/inflammatory diseases, and functional conditions may be sometimes difficult to distinguish from acute stroke based solely on neurological examinations. In addition, NCCT brain scans may be normal in the presence of acute ischemia if the patient is imaged very early. Thus, an objective method for confirming intracranial vessel occlusions prior to treatment is preferable. CT angiography is rapid and widely available. The location and extent of intracranial thrombus has been shown to predict functional outcome and risk of parenchymal hematoma [21]. CTA and perfusion studies provides an effective add-on to NCCT brain scans in acute stroke imaging by significantly increasing the sensitivity and reliability of acute infarct and vessel occlusion detection [22]. In our study, there was high sensitivity, specificity, and accuracy and good interrater agreement to diagnose intracranial occlusion on CTA scans reviewed on an iOS device. One patient with an occlusion of the distal segment of the vertebral artery was identified as normal on an iOS device. This false-negative error may have been avoided if our system allowed the user to view orthogonal and multiplanar reformats of the source data. Once again, the server software does provide functions for orthogonal and multiplanar reformatting. However, the iOS client software used in this study did not expose this functionality to the user. These functions were added to the new version licensed by Health Canada.

Prior to developing our iOS client software, we had a number of concerns about both the iOS device and the client-server architecture that we thought might limit clinical utility. In particular, the iOS device display is (1) much smaller than that of a workstation; (2) has only 320 x 480 pixels—insufficient to display a full 512 x 512 CT scan image; and (3) is not a medical-grade display being used in a lighting-controlled radiology reading room. In addition, we were concerned that our client-server architecture might provide insufficient interactivity for practical use. In practice, we discovered several factors that helped alleviate our concerns. First, in most NCCT brain scans, the patient's brain does not fill the entire 512 x 512 image array. In addition, readers were able to easily observe all image regions at native resolution through interactive translation, while interactive zoom allowed them to focus in on image areas of interest. Next, when interpreting images, users tend to position the iOS device display much closer to their eyes, which helps compensate, to some degree, for the smaller size of the iOS device display. In addition, the iOS device display has 25% higher pixel density and luminance than those on the medical-grade liquid crystal display (LCD) monitor of the workstation used in these experiments.

There were limitations in our study. In particular, it was a retrospective analysis performed of 173 patients' brain scans by two neuroradiologists in a research laboratory environment with high-speed network infrastructure. Only 7 of 112 (6%) hemorrhagic strokes occurred among these patients, which would cause the sensitivity of hemorrhagic diagnosis to be high regardless of how the images were interpreted. Consequently, care should be taken when drawing conclusions from our results. In the future, a larger prospective study performed by physicians on-call working under clinical constraints will be required to demonstrate the clinical utility of our system. We would also like to test the potential usefulness of our system for other acute conditions like renal colic, skeletal trauma, and acute coronary disease.

The system should provide practical frame rates over cellular or wireless networks. In our experience, a single visualization server can accommodate 10 or more simultaneous iOS device users and is capable of delivering and displaying up to 14 frames per second on an iOS device connected over a 802.11g Wi-Fi network. The frame rate was enough to provide sufficient interactivity for comfortable use. However, the frame rate on a 3G cellular network was 1 to 4 frames per second, which was insufficient for practical use. We know that fourth generation (4G) cellular networks are now installed in many metropolitan centers. We estimate that the higher bandwidth of these new cellular networks should allow 10 to 15 frames per second to be delivered to smartphones. However, currently only the iPhone 4 and a few Android-based smartphones are capable of utilizing greater network bandwidth.

Our system has been approved as a medical device under device class 2 with Health Canada. It is under review by the US Food and Drug Administration. We suspect that the regulatory agencies will have two major concerns regarding this system: that it has functional equivalence to previously licensed image interpretation platforms and that it does not unduly jeopardize patient confidentiality. Our client-server architecture may help address each of these concerns. First, protection of patient confidentiality is aided by the fact the no DICOM data is stored outside the hospital firewall on the iOS device. Second, since the server performs all rendering operations, it may be possible to expose sufficient workstation-class functionality on the iOS device to demonstrate functional equivalence. For example, the iOS device and other current generation smartphones do not have sufficient computational resources and rapid access memories to perform interactive multiplanar reformatting or advanced 3-D visualization. Yet, these functions may be required by the regulatory agencies to demonstrate functional equivalence to existing interpretation platforms. We are investigating new methods to usefully expose additional advanced visualization capabilities to the remote user on an iOS device and to optimize both server and client performance to enhance interactivity over both Wi-Fi and cellular networks.

In summary, the smartphone client-server teleradiology system appears promising and may have the potential to allow urgent management decisions in acute stroke. However, this study was retrospective, involved relatively few patient studies, and only two readers. Widespread conclusions about its clinical utility, especially in other diagnostic use cases, should not be made until additional studies are performed.

Acknowledgments

J Ross Mitchell is funded by the Alberta Informatics Circle of Research Excellence and by the Alberta Heritage Foundation for Medical Research. Michael D Hill is funded by the Heart and Stroke Foundation of Alberta and the Alberta Heritage Foundation for Medical Research.

Conflicts of Interest

J Ross Mitchell is the co-founder and chief scientist of Calgary Scientific Inc (CSI). Mayank Goyal, Michael Hill, and J Ross Mitchell are shareholders of CSI. Monroe Thomas is the Vice President of Architecture with CSI. The iOS device client software was originally developed by J Ross Mitchell and Mark Simpson as a research platform. The source code was subsequently transferred to CSI. No equity in CSI was involved in this exchange. CSI enhanced and improved the source code to create ResolutionMD Mobile. The value of CSI shares may increase if ResolutionMD Mobile becomes a successful product.

References

1. Kim DK, Yoo SK, Park JJ, Kim SH. PDA-phone-based instant transmission of radiological images over a CDMA network by combining the PACS screen with a Bluetooth-interfaced local wireless link. J Digit Imaging 2007 Jun;20(2):131-139. [doi: 10.1007/s10278-007-9002-2] [Medline: 17505870]
2. Yamada M, Watarai H, Andou T, Sakai N. Emergency image transfer system through a mobile telephone in Japan: technical note. Neurosurgery 2003 Apr;52(4):986-988. [Medline: 12657199]
3. Yamamoto LG, Williams DR. A demonstration of instant pocket wireless CT teleradiology to facilitate stat neurosurgical consultation and future telemedicine implications. Am J Emerg Med 2000 Jul;18(4):423-426. [doi: 10.1053/ajem.2000.6316] [Medline: 10919531]
4. Kotter E, Baumann T, Jäger D, Langer M. Technologies for image distribution in hospitals. Eur Radiol 2006 Jun;16(6):1270-1279. [doi: 10.1007/s00330-006-0153-1] [Medline: 16534601]
5. Pohjonen H, Ross P, Blickman JG, Kamman R. Pervasive access to images and data--the use of computing grids and mobile/wireless devices across healthcare enterprises. IEEE Trans Inf Technol Biomed 2007 Jan;11(1):81-86. [Medline: 17249406]
6. Barber PA, Demchuk AM, Zhang J, Buchan AM. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. Lancet 2000 May 13;355(9216):1670-1674. [Medline: 10905241]
7. Savcenko V, Erickson BJ, Persons KR, Campeau NG, Huston J, Wood CP, et al. An evaluation of JPEG and JPEG 2000 irreversible compression algorithms applied to neurologic computed tomography and magnetic resonance images. Joint Photographic Experts Group. J Digit Imaging 2000 May;13(2 Suppl 1):183-185. [Medline: 10847394]
8. Canadian Association of Radiologists. 2008. CAR standards for Irreversible Compression in Digital Diagnostic Imaging within Radiology URL: http://www.car.ca/uploads/standards%20guidelines/Standard_Lossy_Compression_EN.pdf [accessed 2011-03-23] [WebCite Cache ID 5xPBYQpaj]
9. Poliakov AV, Albright E, Hinshaw KP, Corina DP, Ojemann G, Martin RF, et al. Server-based approach to web visualization of integrated three-dimensional brain imaging data. J Am Med Inform Assoc 2005 Apr;12(2):140-151. [doi: 10.1197/jamia.M1671] [Medline: 15561787]
10. Saver JL. Time is brain--quantified. Stroke 2006 Jan;37(1):263-266 [FREE Full text] [doi: 10.1161/01.STR.0000196957.55928.ab] [Medline: 16339467]
11. Marler JR, Tilley BC, Lu M, Brott TG, Lyden PC, Grotta JC, et al. Early stroke treatment associated with better outcome: the NINDS rt-PA stroke study. Neurology 2000 Dec 12;55(11):1649-1655. [Medline: 11113218]
12. Kidwell CS, Chalela JA, Saver JL, Starkman S, Hill MD, Demchuk AM, et al. Comparison of MRI and CT for detection of acute intracerebral hemorrhage. JAMA 2004 Oct 20;292(15):1823-1830 [FREE Full text] [doi: 10.1001/jama.292.15.1823] [Medline: 15494579]
13. Toomey RJ, Ryan JT, McEntee MF, Evanoff MG, Chakraborty DP, McNulty JP, et al. Diagnostic efficacy of handheld devices for emergency radiologic consultation. AJR Am J Roentgenol 2010 Feb;194(2):469-474 [FREE Full text] [doi: 10.2214/AJR.09.3418] [Medline: 20093611]

14. Dzialowski I, Hill MD, Coutts SB, Demchuk AM, Kent DM, Wunderlich O, et al. Extent of early ischemic changes on computed tomography (CT) before thrombolysis: prognostic value of the Alberta Stroke Program Early CT Score in ECASS II. Stroke 2006 Apr;37(4):973-978 [FREE Full text] [doi: 10.1161/01.STR.0000206215.62441.56] [Medline: 16497977]
15. Wardlaw JM, Mielke O. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment--systematic review. Radiology 2005 May;235(2):444-453 [FREE Full text] [doi: 10.1148/radiol.2352040262] [Medline: 15858087]
16. Coutts SB, Demchuk AM, Barber PA, Hu WY, Simon JE, Buchan AM, VISION Study Group. Interobserver variation of ASPECTS in real time. Stroke 2004 May;35(5):e103-e105 [FREE Full text] [doi: 10.1161/01.STR.0000127082.19473.45] [Medline: 15073381]
17. Demchuk AM, Hill MD, Barber PA, Silver B, Patel SC, Levine SR, et al. Importance of early ischemic computed tomography changes using ASPECTS in NINDS rtPA Stroke Study. Stroke 2005 Oct;36(10):2110-2115 [FREE Full text] [doi: 10.1161/01.STR.0000181116.15426.58] [Medline: 16166579]
18. Leary MC, Kidwell CS, Villablanca JP, Starkman S, Jahan R, Duckwiler GR, et al. Validation of computed tomographic middle cerebral artery "dot"sign: an angiographic correlation study. Stroke 2003 Nov;34(11):2636-2640 [FREE Full text] [doi: 10.1161/01.STR.0000092123.00938.83] [Medline: 14593125]
19. Gács G, Fox AJ, Barnett HJ, Vinuela F. CT visualization of intracranial arterial thromboembolism. Stroke 1983 Oct;14(5):756-762 [FREE Full text] [Medline: 6658961]
20. Manelfe C, Larrue V, von Kummer R, Bozzao L, Ringleb P, Bastianello S, et al. Association of hyperdense middle cerebral artery sign with clinical outcome in patients treated with tissue plasminogen activator. Stroke 1999 Apr;30(4):769-772 [FREE Full text] [Medline: 10187877]
21. Puetz V, Dzialowski I, Hill MD, Subramaniam S, Sylaja PN, Krol A, Calgary CTA Study Group. Intracranial thrombus extent predicts clinical outcome, final infarct size and hemorrhagic transformation in ischemic stroke: the clot burden score. Int J Stroke 2008 Nov;3(4):230-236. [doi: 10.1111/j.1747-4949.2008.00221.x] [Medline: 18811738]
22. Scharf J, Brockmann MA, Daffertshofer M, Diepers M, Neumaier-Probst E, Weiss C, et al. Improvement of sensitivity and interrater reliability to detect acute stroke by dynamic perfusion computed tomography and computed tomography angiography. J Comput Assist Tomogr 2006 Feb;30(1):105-110. [Medline: 16365583]

Abbreviations

3G: third generation
4G: fourth generation
API: application program interfaces
ASPECTS: Alberta Stroke Program Early CT Score
CT: computerized tomography
CSI: Calgary Scientific Inc
CTA: computed tomography angiogram
DICOM: digital imaging and communications in medicine
GB: gigabytes
GHz: gigahertz
http: hypertext transfer protocol
ICH: intracranial hemorrhage
JPEG: joint photographic experts group
LCD: liquid crystal display
MB: megabytes
MCA: middle cerebral artery
NCCT: noncontrast computed tomography
PACS: picture archiving and communication system
RAM: random-access memory
URL: uniform resource locator
XML: extended markup language

Edited by G Eysenbach; submitted 12.01.11; peer-reviewed by H Baloch, J Cafazzo; comments to author 11.02.11; revised version received 01.03.11; accepted 10.03.11; published 06.05.11.

Please cite as:
Mitchell JR, Sharma P, Modi J, Simpson M, Thomas M, Hill MD, Goyal M
A Smartphone Client-Server Teleradiology System for Primary Diagnosis of Acute Stroke
J Med Internet Res 2011;13(2):e31
URL: http://www.jmir.org/2011/2/e31/
doi:10.2196/jmir.1732
PMID:21550961

Original Paper

Computer-Assisted Update of a Consumer Health Vocabulary Through Mining of Social Network Data

Kristina M Doing-Harris[1*], BCompSci MA MS PhD; Qing Zeng-Treitler[1*], PhD

University of Utah, Department of Biomedical Informatics, Salt Lake City, UT, United States

[*] all authors contributed equally

Corresponding Author:
Kristina M Doing-Harris, BCompSci MA MS PhD
University of Utah
Department of Biomedical Informatics
Room 5775 HSEB
26 S 2000 E
Salt Lake City, UT, 84112
United States
Phone: 1 801 518 8636
Fax: 1 801 581 4297
Email: kristina.doing-harris@utah.edu

Abstract

Background: Consumer health vocabularies (CHVs) have been developed to aid consumer health informatics applications. This purpose is best served if the vocabulary evolves with consumers' language.

Objective: Our objective was to create a computer assisted update (CAU) system that works with live corpora to identify new candidate terms for inclusion in the open access and collaborative (OAC) CHV.

Methods: The CAU system consisted of three main parts: a Web crawler and an HTML parser, a candidate term filter that utilizes natural language processing tools including term recognition methods, and a human review interface. In evaluation, the CAU system was applied to the health-related social network website PatientsLikeMe.com. The system's utility was assessed by comparing the candidate term list it generated to a list of valid terms hand extracted from the text of the crawled webpages.

Results: The CAU system identified 88,994 unique terms 1- to 7-grams ("n-grams" are n consecutive words within a sentence) in 300 crawled PatientsLikeMe.com webpages. The manual review of the crawled webpages identified 651 valid terms not yet included in the OAC CHV or the Unified Medical Language System (UMLS) Metathesaurus, a collection of vocabularies amalgamated to form an ontology of medical terms, (ie, 1 valid term per 136.7 candidate n-grams). The term filter selected 774 candidate terms, of which 237 were valid terms, that is, 1 valid term among every 3 or 4 candidates reviewed.

Conclusion: The CAU system is effective for generating a list of candidate terms for human review during CHV development.

(J Med Internet Res 2011;13(2):e37) doi:10.2196/jmir.1636

KEYWORDS

Consumer health information; vocabulary; natural language processing; automatic term recognition; data mining; social networking

Introduction

Controlled vocabularies play an important role in the development of biomedical informatics applications because data used by clinical, bibliometric, and research applications need to be coded for easy retrieval and analysis. Research and development activities have been carried out to provide standardized health vocabularies, for example, SNOMED (Systematized Nomenclature of Medicine) and LOINC (Logical Observation Identifiers Names and Codes). In the past, these vocabularies focused on the language of professionals, but lately consumer health vocabulary (CHV) [1] has been rising in prominence, and, consequently, CHV research has gained recognition.

Controlled vocabularies require maintenance and updating due to the continuing evolution of language itself [2-4]. This evolution has been seen for centuries in the regular update and revision of dictionaries [5,6]. Controlled vocabularies serving electronic applications are no exception. The demand for maintenance and updating of vocabularies is particularly high in areas related to ongoing research and development. As new findings emerge, new words are added to the vocabulary. In health care especially, there is a constant stream of new names (eg, new medications, disorders, and tests) [7].

Frequently, new health terms used by professionals migrate, in some form, into popular parlance. For example, the term *mass spectrometer* was unheard of 30 years ago, but a number of lay people now could identify it as a piece of lab equipment. Although *deoxyribonucleic acid* may be confusing, DNA is in the vocabulary of school-aged children. The media also plays a role in term migration. For example, in 2009, media coverage introduced new vocabulary words such as *pandemic*, *swine flu*, *H1N1*, *energy expenditure*, and *single-payer system* into popular speech. Similarly, the meaning and popularity of health terms change or evolve in the lay use. For example, it is common for lay people to use the term *anorexia* to refer to the concept *anorexia nervosa*, though in medical literature *anorexia* refers only to the loss of appetite.

To be effective, a CHV must keep pace with changes in the language used by consumers [1]. This paper describes a computer-assisted update (CAU) system that uses an online social network as a living corpus of health-related terms. The system parses and screens terms using the natural language processing (NLP) techniques of dictionary lookup and automatic term recognition. New candidate terms are thereby identified for inclusion in the open access and collaborative (OAC) CHV.

Background

In this background section, we will first briefly review the prior research and current practice for updating controlled health vocabularies. Next, we will discuss the automated methods used to identify valid terms from text corpora. Then, we will switch focus to provide background information on the OAC CHV research. Finally, we will describe the rationale behind using a live corpus with automated term identification for updating the OAC CHV.

Updating Controlled Health Vocabularies

Prior research has found that nearly all large controlled health vocabularies have similar core maintenance procedures [8]. Bakhshi-Raiez et al describe a framework for the maintenance of controlled health vocabularies. They refer to controlled health vocabularies as medical terminology systems (TSs). Their framework consists of four components. The primary component of their framework is "execution." This covers the core activities of the maintenance process including: collection of proposals for changes, validation of the proposals for changes, implementation of changes, verification of changes, documentation of proposals and implemented changes, and version management. The three other components, namely "process management," "change specification," and "editing tools" act in support of "execution." Bakhshi-Raiez et al conducted a survey of 37 TSs. They divided the group of TSs into quartiles based on the number of concepts included in each system. The quartile relevant to this paper is quartile IV, which included systems with more than 46,155 concepts. Quartile IV would include the OAC CHV, which has 58,319 concepts. For the execution component, almost all of the quartile IV systems satisfied the main criteria, that is, 67% included standardized change proposals, 100% validated the change proposals, 100% had maintenance teams that verified accepted proposals, 100% had structured and standardized documentation, 100% documented changes made, and 100% produced new versions with unique id's, while only 70% produced twice yearly updates. The CAU system we describe here is designed to automate the production and collection of change proposals and then assist with the validation of those proposals.

The current practice for the generation and collection of proposals for TS changes and their validation typically involves collecting proposals via email or Internet and having a team of specialists validate them. For example, there is a Web-based Semantic MediaWiki system for maintaining entries in the National Cancer Institute Metathesaurus [9]. The SPECIALIST Lexicon included in the Unified Medical Language System (UMLS) collects words from literature as well as multiple dictionaries [10]. The medical subject headings (MeSH) section staff continually revises and updates the MeSH vocabulary based on scientific literature in emerging areas of research, defines these terms within the context of the existing vocabulary, and recommends their addition to MeSH [11]. In a personal communication, Stuart Nelson, the head of MeSH, estimated that 20% of his time is devoted to updating and revision. There are also six full time MeSH analysts. Clearly, vocabulary maintenance is a labor-intensive process, one whose efficiency could be improved by the proposed CAU system. The first step to automation would be the generation and collection of proposals for changes, a step that lacks standardization in one-third of all large vocabularies' maintenance procedures [8].

Automatic Term Recognition in the Biomedical Domain

One method of automatically generating change proposals is to identify valid candidate terms in a text corpus through automatic term recognition (ATR) [12,13]. ATR studies overlap with the discipline known as named entity recognition (NER). ATR refers to systems that search for general types of terms as opposed to named entities. A term becomes a named entity when it is mapped to an ontology or dictionary of terms, which gives the term meaning in a context outside of the document in which it is found. General terms have no such wider meaning. Examples of biomedical NER systems include Termoid, MetaMap and Bio-tagger [14-16]. Examples of biomedical ATRs are Collier et al's hidden Markov model for identifying gene names and gene products, as well as Frantzi et al's "C-value" and Zeng et al's "termhood" score [17-19].

Since C-value and termhood scores are used in our study, we will briefly describe them here. The C-value equation uses part of speech-tagged data and restricts candidate terms to noun phrases. The best results are obtained with an open linguistic filter that returns noun phrases, which include multiple adjectives and nouns [18]. C-value is then calculated using the frequency of occurrence of the candidate term combined with its frequency of occurrence as part of other, longer candidate terms, along with the number of longer candidate terms and their lengths. Expanding upon the C-value, the termhood logistic regression equation (termhood score) was developed by Zeng et al to identify multi-word consumer health terms including those that are not noun phrases [19]. The features used to train the logistic-based model include parts of speech of term components, frequency of occurrence of candidate terms, as well as frequency of occurrence of said candidate terms in both larger and smaller alternative candidate terms. Zeng et al compared C-value and termhood score ratings using strings that had already been human reviewed and found that termhood score outperformed C-value on their dataset [19].

Consumer Health Vocabulary

The development of this CAU system is part of the OAC CHV research program. The OAC CHV was developed using a phased, distributed, user source-based approach [1]. To incorporate new terms, seven human review criteria were established [19]: (1) CHV terms should be syntactic constituents or phrases such as noun phrases or adjectival phrases; (2) CHV terms should have independent semantics and should not only occur as a part of longer valid terms or as a part of wild card searches; (3) CHV terms should be specific to the medical domain; (4) CHV terms should function as semantic components; (5) “n-grams” (n-grams are n consecutive words within a sentence) representing UMLS concepts are considered to be CHV terms, but CHV terms may represent non-UMLS concepts; (6) CHV terms may be eponymous forms; and (7) CHV terms may include spelling errors. These criteria guide the human review in this study; the current version of the CHV contains 152,778 entries, representing 58,319 concepts.

Live Corpora

From the beginning, research on CHV relied on text corpora containing consumer utterances. Although most of the text corpora were collected from live sources such as patient email, online forums, query logs, and social networks [16-20], they were treated as static datasets for analysis. In this study, we aimed to directly tap into the live sources. Due to the extremely fast growth of social network sites, including health-related social network sites and their public availability [21], we chose to test the CAU system on the social networking site PatientsLikeMe. Our lab has a collaborative relationship with PatientsLikeMe, which facilitated permission to use the site. The CAU system, however, could also work with other types of live sources.

PatientsLikeMe.com is an online community built to support information exchange between patients. The site provides customized disease-specific outcome and visualization tools to help patients understand and share information about their condition [22,23]. The private pages of the site are designed for patients to enter symptoms and track their disease. The public pages, on the other hand, include information provided by the site management and excerpts of information shared by users. The public pages thus contain language used by professionals as well as lay people. An example of language used by professionals would be, “ALS, or amyotrophic lateral sclerosis, is a neurodegenerative disease caused by the degeneration of motor neurons.” An example of language used by a lay person would be, “…The first thing that I thought might be your problem is malnutrition. Man, you’re losing weight crazy fast. I think you better consider getting peg tube if you desire.” By sticking to the public pages, we plan on tapping into this social networking aspect of the site without breaching the privacy of the users.

Methods

We devised the CAU system to mine Web content using a combination of NLP methods: dictionary-lookup, C-value ATR, and termhood ATR. The goal was to discover new health-related terms used by consumers but not yet included in our existing vocabulary. The best candidate term list should contain a small number of terms while providing a reasonably high yield of valid terms after the human review.

System Architecture

The CAU system architecture is shown in Figure 1. It consists of three processing stages, stage 1, stage 2, and stage 3. Stage 1 is the stage in which raw text is obtained, parsed, and n-grams, that is, groups of words, are extracted. This stage involves three substages: crawling, parsing, and n-gram extraction.

Crawling

In the crawling substage, the system crawls public pages on the Web. Crawling consists of navigating to the home page, collecting all the links to other pages in a queue, navigating to those pages in turn, and adding any links found to the end of the queue. This loop continues until the end of the queue or until a predefined number of pages has been visited. The remaining content of each page is processed by removing HTML tags, adding periods to the ends of text blocks followed by more than one new line, and saving the resultant text.

Parsing

In the parsing substage, the system uses the open-source natural language processing application Health Information Text Extractor (HITEx) [24] to identify parts of speech, noun phrases, and named entities. HITEx is an NLP system, which contains an OpenNLP parser and uses MetaMap for NER.

N-gram Extraction

In the n-gram extraction substage, the system extracts n-grams (1- to 7-grams) with overlap. Overlap means a word can be included in more than one n-gram. The n-grams are filtered to retain n-grams identified by HITEx as noun phrases, n-grams, which contain a verb (ie, potential verb phrases), and n-grams that contain the word *symptom*. N-grams that include numbers or symbols are excluded at this point. This linguistic filtering strategy is based on Frantzi et al’s finding that the C-value ATR produced better results with an open linguistic filter [18]. A stop list (ie, a list of terms that should be ignored) was created from a list of the 1000 most common phrases in English (eg, *a little*, *a few*, *we like it very much*) [25]. Terms found on the stop list are excluded and frequency information is gathered at this point.

Stage 2 is the stage in which further NLP techniques are used to identify candidate terms on the n-grams list. This stage consists of three substages, two dictionary-type look up stages, the UMLS/CHV filter substage and the VA medical record term filter substage, and one ATR stage, the ATR filter substage.

UMLS/CHV Filter

Given our interest in discovering new terms, the n-grams are looked up in the current CHV list and the UMLS Metathesaurus. To insure the most up to date version of UMLS (2010AA) was used, n-grams were checked using the UMLS Web service. Those n-grams that were not present in UMLS or CHV were denoted non-CHV.

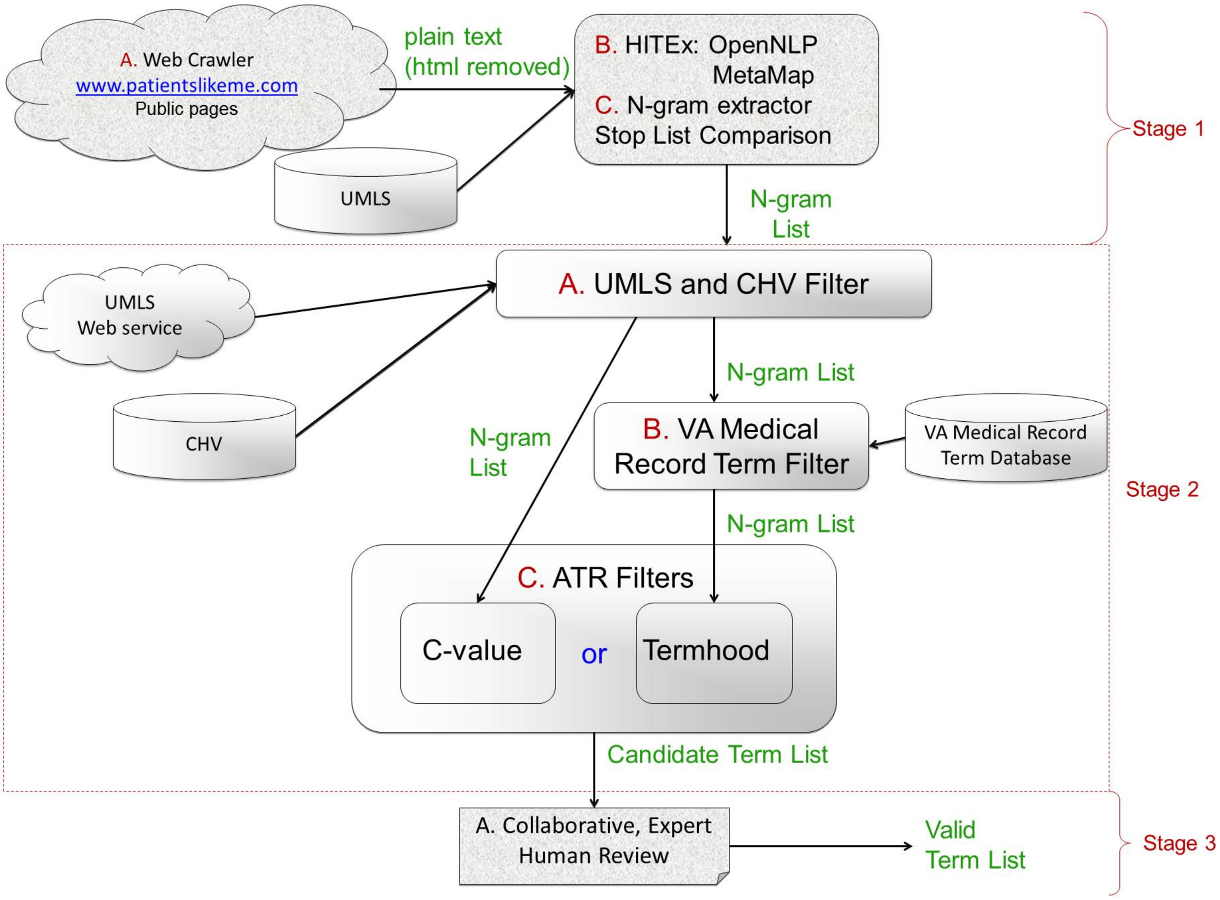
VA Medical Record Term Filter

To filter the nonmedical terms from the non-CHV n-grams, we looked them up in a database of 70,000 medical records of patients obtained from the US Department of Veteran's Affairs of patients with amyotrophic lateral sclerosis (ALS), Parkinson's, and multiple sclerosis (MS) dated from January 1, 1998, through December 31, 2008. These records contain a broad spectrum of medical topics and note types. They are not limited to neurology or the three diseases. These records were obtained by another group in our department with internal review board (IRB) approval. IRB approval was given for a member of that group to compare terms to this database for us, returning a yes/no answer. All terms, which returned *yes* were entered into a database for future comparisons. We will refer to this database as the VA medical record term database.

ATR Filters

Calculated are two ATR scores, that is, termhood and C-value. The termhood score was calculated using the logistic regression equation described in Zeng et al [19]. The C-value is calculated using the equation described in Frantzi et al [18].

Figure 1. CAU System Diagram



The third stage is the human (expert) review stage. In this stage, candidate terms are submitted for collaborative expert review. To aid this process, we created an interactive website for the OAC CHV. Approved reviewers can access the site and recommend URLs for inclusion in the crawl, review candidate terms, review recent candidate term comments, and review CHV preferred names. While reviewing candidate terms, the reviewer can vote for or against a term's inclusion in the OAC CHV, see all terms' frequencies and votes, and get three examples of a term in the context of the webpage on which it was found. Reviewers can also comment on a candidate term without registering a vote. Each reviewer can vote for a term only once. The public may comment on the candidate terms on the CHV Wiki by browsing the term list, choosing candidate terms, and clicking the term on which they would like to comment.

Evaluation

The system was evaluated by crawling the PatientsLikeMe.com website, examining the candidate terms identified and calculating valid term yield (ie, percentage of total candidate terms which are valid terms). For the purposes of this paper, the final stage of collaborative human review was replaced by the creation of a valid term list, which functions as the gold standard for this study. The valid term list was manually extracted from the webpages by the first author and filtered to exclude terms already represented in the UMLS/CHV. Reviewing the first 300 pages encountered in the crawl produced a valid term list containing 651 terms, which we considered sufficiently large for the purposes of this experiment. Therefore, we restricted the processing by the system to those pages.

To assess the accuracy of the valid term list, a panel of expert reviewers (two physicians and two allied medical personnel) reviewed 100 random non-CHV terms found in our VA medical record term database from the initial parse of the webpages. Each expert's agreement with the gold standard valid term list was assessed using the balanced F-measure discussed in Hripcsak and Rothschild [26]. The F-values found were 0.94, 0.86, 0.94, and 0.91, indicating that terms chosen as valid were indeed valid and that valid terms were not being missed.

Results

Stage 1

The crawler visited the PatientsLikeMe.com website (marked A in Figure 1) public pages only. The parsing and n-gram extraction phases (marked B and C in Figure 1) found 88,994 n-grams. The n-gram list contained all 651 terms from the valid term list.

Stage 2

In the UMLS/CHV filter phase 1045 (1% of the total) n-grams were found in the CHV/UMLS. The total number of non-CHV terms remained large at 87,949.

The VA medical record term filter phase filtered out most of the n-grams. It eliminated all but 923 n-grams (99% reduction) and all but 215 terms from the valid term list (67% reduction).

The eliminated valid term list terms in this phase were, for the most part, long (eg, *sub mandibular injection paralyzed swallow muscles*), brand names (eg, *Nurofen*), combination terms (eg, *lipodystrophy lipoatrophy*), or biochemical terms (eg, *L-methyfolate Metafolin*). The loss of these terms was concerning, but it is possible that they would be found if a comparison was made with a larger database of medical records. Other valid term list terms excluded were consumer terms (eg, *brain fog* and *loss of time*), which may not typically be recorded in medical records.

Figure 2. Effect of termhood score thresholds on the number of candidate terms

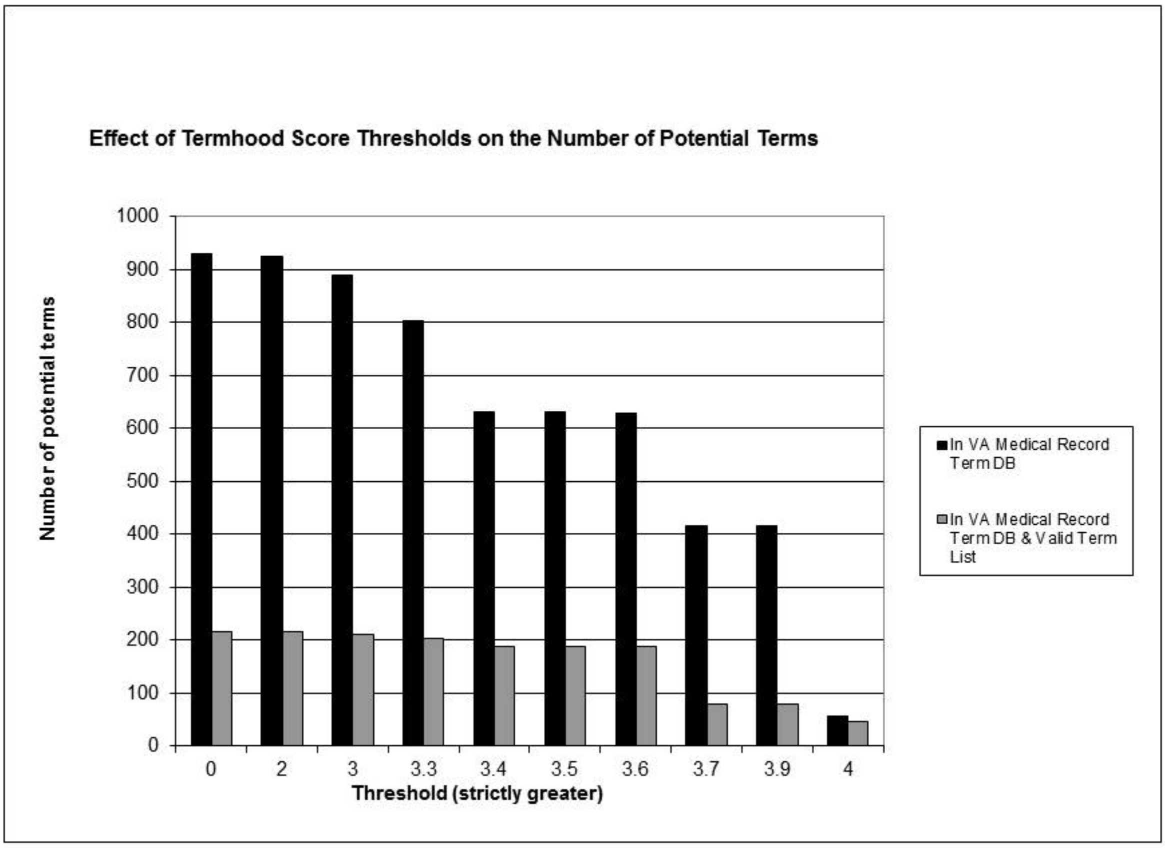
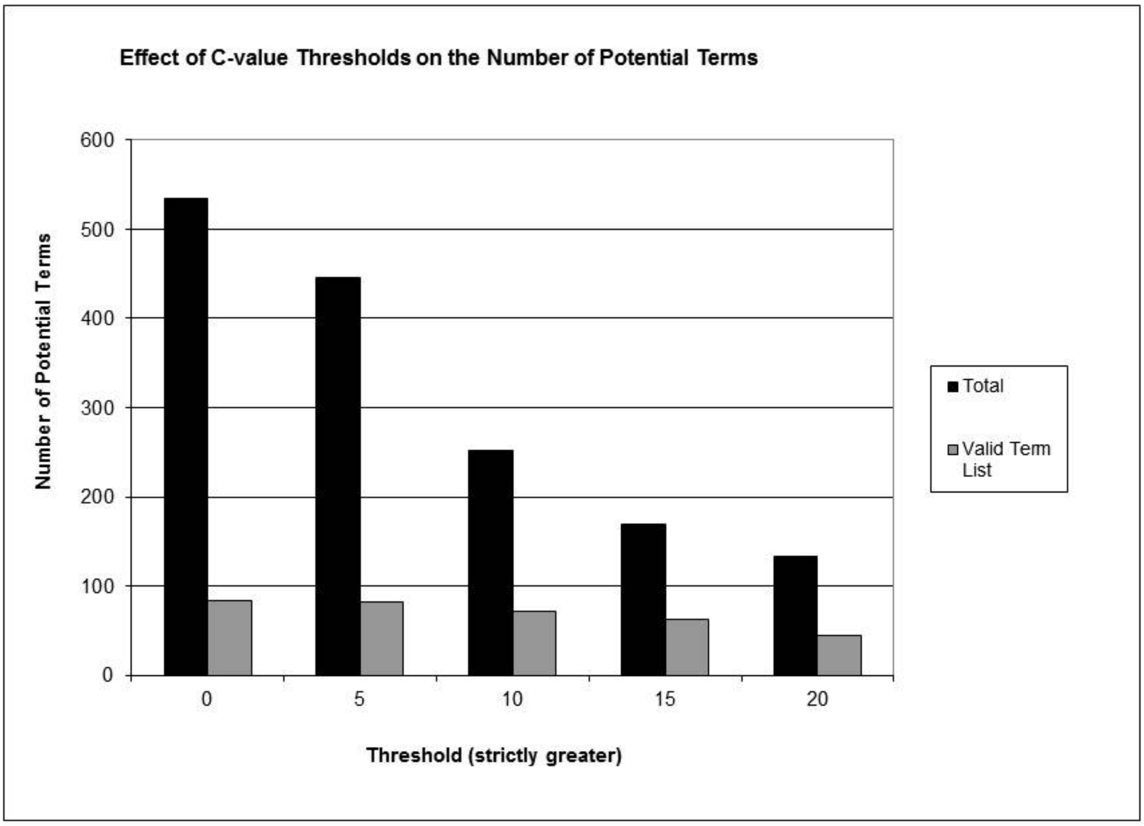


Figure 3. Effect of C-value score thresholds on the number of candidate terms



In the ATR filter phase, two filters were built by applying a threshold to the ATR scores. The first filter used was based on termhood score and was applied to each non-CH n-gram found in the VA medical record term database. From Figure 2 it can be seen that the number of terms above the threshold dropped gradually. The best yield of valid terms was achieved with thresholds between 3.4 and 3.6. Choosing the higher of these thresholds (3.6) identified 622 candidate terms (99.3% cumulative reduction) of which 189 were from the valid term list (69% cumulative reduction), a 30% valid term yield. The excluded valid terms include terms such as *asthenia symptom* and *breast-feeding*.

The second filter was based on C-value and was applied to all non-CHV n-grams. From Figure 3 it can be seen that the number of terms above this threshold also dropped gradually. The best yield of valid terms was found at a threshold of 15 (170 terms, 93.7% reduction, with 62 from the valid term list, 90% reduction), a 36% yield. This yield was higher than the termhood with VA medical record term filter, but the number of candidate terms returned was much lower. Some of the valid terms excluded here were ill-formed phrases that included the word *symptom* (eg, *asthenia symptom*, *cramps symptom*), while others were potentially more concerning (eg, *staggering walk*, *inability to raise the foot*).

When the C-value threshold was combined with the termhood threshold the number of candidate terms increased while keeping the valid term yield around 30%, the number of terms remaining was 774 with 237 valid terms. Combining the filters caught 48 more valid terms over the termhood filter alone, including *cold legs below knee, augmentative speech device, and acid reflux GER gastroesophageal reflux*.

The results are summarized in Table 1. Table 2 contains a sample list of candidate terms identified.

Table 1. Results at each stage of CAU processing

Processing Phase	System Stage (as labeled in Figure 1)	Candidate Terms	Valid Terms
1. Initial parse	1	88,994	651
2. CHV/UMLS filter	2A	87,949	651
3. VA medical record term filter	2B	923	215
4. ATR filter 1: termhood score (threshold 3.6) (VA medical record terms only)	2C	622	189
5. ATR filter 2: C-value (threshold 15) (all terms)	2C	170	62
Phase 4 and 5 filters combined	2C	774	237

Table 2. Sample list of valid and invalid candidate terms identified (displayed in no particular order and in the format in which they were found)

Sample of Candidate Terms Identified		
Valid Terms	Invalid Terms	
manage	Does not protect	acidify
Side effects	PRO survey	Permobil
Bi-Pap	Motorized recliner	Levetiracetam Treatment Report
Weakness in Hands	PRO 0	Pilates Treatment Report
Devic's Neuromyelitis Optica	MSA Multiple System	Equate Acetaminophen
	fitness	all the time
	depending	Hoveround
	is particularly	cope
	Seizures grand mal	Seage III GR
	saliva control	

Stage 3, the expert human review stage, was preempted by the use of the valid term list as gold standard in this study.

Discussion

Principal Result

We developed a vocabulary maintenance system and tested it on the PatientsLikeMe.com website. The system first identified a very high number of n-grams (n = 88,994) and then created a candidate term list of a reasonable size (n = 774) with a relatively high valid term yield (31% or 237/774). The system and the experiment are a proof-of-concept for procuring new terms using living corpora and ATR to aid vocabulary maintenance.

The system utilized NLP methods including parsing, dictionary look up, comparison with a medical record database, and ATR to filter out both the n-grams that were not related to health and ill-formed sentence fragments. Following all the filtering, the reviewers found 1 valid term among every 3 or 4 candidate terms reviewed. This is considerably better than the initial n-gram list which would have returned on average 1 valid term for every 137 candidate terms.

The system will become more efficient after each maintenance cycle. All candidate terms rejected for inclusion will be added to the stop list, which should decrease the number of candidate terms. For instance, following the experiment described in the paper we conducted a second crawl of 300 pages and obtained 240 candidate terms for human review with 71 potentially valid terms, which maintained a yield of 30% (71/240). However, since 88 of the candidate terms were place names expanding the stop list to include place names would reduce the candidate term list to 212 with 71 potentially valid terms, a valid term yield of 33% (72/212).

There have been previously reported higher yields using C-value and termhood scores [18,19]. The yield, however, is sensitive to the data and task involved in each study. Spasic et al used C-value to extract terms from full-text journal articles with a reported yield of 61%. However, they targeted all valid terms instead of new terms (ie, terms not yet included in a vocabulary), which are fewer and harder to find. In our own previous study to identify new CHV terms, both termhood score and C-value score were used. The termhood score yield was 38% and the C-value score was even lower. The data set used in that study was the query log to MEDLINEPlus. Compared with query logs, PatientsLikeMe pages contain more "noise" (ie, terms that are similar in structure to those we seek but are not health-related), which increases the number of candidate terms found. We chose not to use only either C-value or termhood scores alone on these data because the results produced were much lower than the 31% we report here.

Implications for the System

The results of this study point to the necessity of using both the termhood and C-value methods. The termhood score required first matching terms with the VA medical record term database in order to provide a concise list. This could be problematic, as consumer terms may not occur in physicians' notes. Evidence for their absence is the drop in valid terms after VA medical record term filtering from 651 to 215. C-value balances termhood by not requiring prefiltering. However, for C-value to generate a concise list, too many valid terms are excluded, only 62 out of 651.

An implication of this study specific to this type of system is the choice of threshold. We found empirically that a threshold of 3.6 for the termhood score and a threshold of 15 for the C-value score produced a list that retained enough valid terms while excluding enough invalid terms. It is possible to manipulate these thresholds. Looking at Figure 1, it can be seen that a termhood threshold of 4 produces a candidate term list which is 95% valid terms. Unfortunately, the total number of valid terms found would be only 42 out of a possible 651. We consider identifying only 6% of the available terms too inefficient. Increasing the C-value threshold produces a similar result. While the valid term yield increases to 78%, only 44 valid terms are identified. It is possible that these thresholds could be increased and the number of valid terms missed could be mitigated by processing an extremely large number of webpages.

General Implications

This system could potentially be used for vocabulary maintenance beyond CHV and even beyond the health domain. Since an increasingly large proportion of contemporary writing is published on the Internet, it is possible to crawl open-access journals, blogs, and Web news channels to identify new candidate terms for inclusion in a variety of vocabularies.

This system could also potentially be used to track the evolution of lay health language. Once the system is up-to-date, each new set of updates will be representative of the changes occurring in consumer terminology. It may be possible to use this information to recognize patients' understanding and information needs based on their vocabulary.

Limitations

A potential limitation of using the PatientsLikeMe website is the "higher level" language that occurs in the content produced by the site operators as opposed to the users. Higher level refers to language that is drawn directly from physicians' vocabulary. In this case, it is likely that the term will be contained in the UMLS Metathesaurus and thus ignored in the collection of new terms. Additionally, this broader exposure to higher level language may cause increased migration of terms. The migration of such terms would be reflected in the frequency-of-use data that are used to recommend the name preferred for use in reference to the concept (the consumer preferred CHV name). Either way, this language should not present a problem for the CAU system.

The CAU system is limited by errors in the parsing and filtering stages. Although part of speech and noun phrase parsing are relatively mature NLP technologies [27], the parsing of webpages poses extra challenges due to the prevalence of incomplete and ungrammatical sentences. The parser used in the HITEx system, OpenNLP, is trained to work with general text. HITEx was developed for the processing of clinical notes, which may be more grammatical or adhere to a different subgrammar. The continuing development of the HITEx NER system incorporated into the CAU will allow it to take advantage of any advances in the parser or mapper associated with HITEx.

Filtering using the medical record data is limited by the size and clinical characteristics of the patient population represented in the medical record database. The database could be enlarged with proper institutional review board approval. The use of C-value with the unfiltered terms also decreases the effect of this limitation.

The two ATR methods [18,19] are also imperfect. Their performance could be improved by preselecting the text to the extent it is practical in the social network setting. There may be a way to target specific locations in the website or on the webpages, perhaps by searching for key section headings or HTML tags.

Another limitation of the CAU system is the continuing need for human review. The grand goal of all automated systems is to operate completely without human intervention or possibly with only minimal expert review. Our current development is far from reaching the goal of zero human review but not from the goal of minimizing reviewer time.

The results of this method on corpora other than PatientsLikeMe require further studies. To assess the potential robustness of the technique we processed the first 300 pages encountered on a crawl of the YahooHealth.com website. We found 309 potential terms with 72 terms valid for inclusion in the CHV, that is, a 23% valid term yield. This is a lower yield than from the PatientsLikeMe.com site. However, as previously discussed, the thresholds we chose impacted the yield.

Future Research

One direction of our future efforts will be to further analyze the terms found and either map them to existing concepts or create new ones. The terms the CAU identifies are not yet included in the UMLS or CHV. It is therefore necessary to determine how to integrate them into the CHV. Since the majority of these new terms are synonyms of health concepts that already exist in the professional controlled vocabularies, it is a simple mapping to include them. However, some brand new concepts may be encountered, in which case we will utilize the characteristics of valid new consumer concepts described by Keselman et al [28] to help guide their inclusion.

In the future, we also plan to explore public participation in the collaborative review phase. In addition to discovering new terms, we plan to use live corpora to estimate the familiarity of health terms and harvest explanations.

Conclusion

Social network data can be used to provide a living corpus, which can be mined to provide new consumer health vocabulary terms. Using ATR and dictionary lookup can narrow the candidate terms discovered to produce a concise list, which allows the vocabulary to evolve with the language without requiring a large amount of human review time.

Acknowledgments

This work was funded by a supplement to the National Institutes of Health (NIH) grant RO1 LM07222. We would like to acknowledge the cooperation of PatientsLikeMe.com, especially Paul Wicks, Brant Chee, and Sally Okun. We would also like to acknowledge the JMIR reviewers whose ideas were incorporated into this paper.

Conflicts of Interest

None declared

Multimedia Appendix 1

CAU presentation PowerPoint

[PPT file (Microsoft Powerpoint File), 1,406 KB - jmir_v13i2e37_app1.ppt]

References

1. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. J Am Med Inform Assoc 2006;13(1):24-29. [doi: 10.1197/jamia.M1761] [Medline: 16221948]
2. Christiansen MH, Chater N. Language as shaped by the brain. Behav Brain Sci 2008 Oct;31(5):489-508; discussion 509. [doi: 10.1017/S0140525X08004998] [Medline: 18826669]
3. Atkinson QD, Meade A, Venditti C, Greenhill SJ, Pagel M. Languages evolve in punctuational bursts. Science 2008 Feb 1;319(5863):588 [FREE Full text] [doi: 10.1126/science.1149683] [Medline: 18239118]
4. Hurford JR, Studdert-Kennedy M, Knight C. Approaches to the Evolution of Language: Social and Cognitive Bases. Cambridge, UK: Cambridge University Press; 1998.
5. Simpson J. 2010. Oxford English Dictionary, Quarterly Updates URL: http://www.oed.com/news/updates/revisions0906.html [accessed 2010-08-16] [WebCite Cache ID 5s3Q5DqJ2]
6. The Oxford English Dictionary. The Revision Programme. Oxford, UK: Oxford University Press; 2010. URL: http://www.oed.com/about/revision.html [accessed 2010-08-16] [WebCite Cache ID 5s3QG860q]
7. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998 Nov;37(4-5):394-403. [Medline: 9865037]
8. Bakhshi-Raiez F, Cornet R, de Keizer NF. Development and application of a framework for maintenance of medical terminological systems. J Am Med Inform Assoc 2008;15(5):687-700. [doi: 10.1197/jamia.M2531] [Medline: 18579838]
9. National Cancer Institute. Term Suggestion. Bethesda, MD: US National Institutes of Health URL: http://ncitermform.nci.nih.gov/ncitermform/?dictionary=NCI%20Metathesaurus&code=C0006826 [WebCite Cache ID 5yLWpYPnF]
10. Fact Sheet: SPECIALIST Lexicon. Bethesda, MD: United States National Library of Medicine; 2006. URL: http://www.nlm.nih.gov/pubs/factsheets/umlslex.html [accessed 2010-08-16] [WebCite Cache ID 5s3QLNCby]
11. Suggestion for Medical Subject Headings Change. Bethesda, MD: United States National Library of Medicine; 2008. URL: http://www.nlm.nih.gov/mesh/meshsugg.html [accessed 2010-08-16] [WebCite Cache ID 5s3QNqGfl]
12. Kageura K, Umino B. Methods of Automatic Term Recognition a Review. Terminology 1996;3(2):259-289 [FREE Full text]
13. Krauthammer M, Nenadic G. Term identification in the biomedical literature. J Biomed Inform 2004 Dec;37(6):512-526. [doi: 10.1016/j.jbi.2004.08.004] [Medline: 15542023]
14. Harkema H, Gaizauskas R, Hepple M, Roberts A, Roberts I, Davis N, et al. A Large Scale Terminology Resource for Biomedical Text Processing. In: BioLink. 2004 Presented at: HLT/NCAAL 2004 Workshop; May 6, 2004; Boston, MA.
15. Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. J Am Med Inform Assoc 2010 May 1;17(3):229-236. [doi: 10.1136/jamia.2009.002733] [Medline: 20442139]
16. Torii M, Hu Z, Wu CH, Liu H. BioTagger-GM: a gene/protein name recognition system. J Am Med Inform Assoc 2009;16(2):247-255. [doi: 10.1197/jamia.M2844] [Medline: 19074302]
17. Collier N, Nobata C, Tsujii J. Extraction the Names of Genes and Gene Products with a Hidden Markov Model. In: Proceedings of COLING. 2000 Presented at: COLING 2000, 18th International Conference on Computational Linguistics; July 31-Aug 4, 2000; Saarbruecken, Germany. [doi: 10.3115/990820.990850]
18. Frantzi KT, Ananiadou S, Mima H. Automatic term recognition of multi-word terms: the C-value/NC-value method. International Journal on Digital Libraries 2003;3(2):115-130 [FREE Full text] [doi: 10.1007/s007999900023]
19. Zeng QT, Tse T, Divita G, Keselman A, Crowell J, Browne AC, et al. Term identification methods for consumer health vocabulary development. J Med Internet Res 2007;9(1):e4 [FREE Full text] [doi: 10.2196/jmir.9.1.e4] [Medline: 17478413]
20. Spasić I, Schober D, Sansone SA, Rebholz-Schuhmann D, Kell DB, Paton NW. Facilitating the development of controlled vocabularies for metabolomics technologies with text mining. BMC Bioinformatics 2008;9 Suppl 5:S5 [FREE Full text] [doi: 10.1186/1471-2105-9-S5-S5] [Medline: 18460187]

21. Lenhart A, Purcell K, Smith A, Zickuhr K. Social Media and Young Adults. Washington, DC: Pew Internet & American Life Project; 2010 Feb 03. URL: http://www.pewinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx [accessed 2011-04-25] [WebCite Cache ID 5yDfqtv7C]
22. Frost JH, Massagli MP. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. J Med Internet Res 2008;10(3):e15 [FREE Full text] [doi: 10.2196/jmir.1053] [Medline: 18504244]
23. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]
24. Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. BMC Med Inform Decis Mak 2006;6:30 [FREE Full text] [doi: 10.1186/1472-6947-6-30] [Medline: 16872495]
25. englishspeak.com. 2011. 1000 Most Common English Phrases URL: http://www.englishspeak.com/english-phrases.cfm [accessed 2011-04-25] [WebCite Cache ID 5yDg46KsI]
26. Hripcsak G, Rothschild AS. Agreement, the f-measure, and reliability in information retrieval. J Am Med Inform Assoc 2005;12(3):296-298. [doi: 10.1197/jamia.M1733] [Medline: 15684123]
27. Nordstrom B, Ranta A. Advances in Natural Language Processing. In: Nordstrom B, Ranta A, editors. 6th International Conference, GoTAL 2008, Gothenburg, Sweden, August 25-27, 2008, Proceedings (Lecture Notes in ... / Lecture Notes in Artificial Intelligence). Berlin, Germany: Springer; 2008.
28. Keselman A, Smith CA, Divita G, Kim H, Browne AC, Leroy G, et al. Consumer health concepts that do not map to the UMLS: where do they fit? J Am Med Inform Assoc 2008;15(4):496-505. [doi: 10.1197/jamia.M2599] [Medline: 18436906]

Abbreviations

ALS: amyotrophic lateral sclerosis
ATR: automatic term recognition
CAU: computer assisted update
CHV: consumer health vocabulary
HITEx: Health Information Text Extractor
IRB: institutional review board
LOINC: Logical Observation Identifiers Names and Codes
MeSH: Medical Subject Headings
ML: Machine Learning
MS: Multiple Sclerosis
N-gram: A group of n consecutive words
Non-CHV: Not contained in the Open Access Collaborative consumer health vocabulary
NER: named entity recognition
NIH: National Institutes of Health
NLP: natural language processing
OAC: open access collaborative
SNOMED: Systematized Nomenclature of Medicine
TS: terminology Systems
UMLS: Unified Medical Library System

Edited by G Eysenbach; submitted 25.08.10; peer-reviewed by L Slaughter, A Keselman, P Wicks, H Hochhiser; comments to author 17.09.10; revised version received 22.02.11; accepted 29.03.11; published 17.05.11.

Please cite as:
Doing-Harris KM, Zeng-Treitler Q
Computer-Assisted Update of a Consumer Health Vocabulary Through Mining of Social Network Data
J Med Internet Res 2011;13(2):e37
URL: http://www.jmir.org/2011/2/e37/
doi:10.2196/jmir.1636
PMID:21586386

Original Paper

Natural Supplements for H1N1 Influenza: Retrospective Observational Infodemiology Study of Information and Search Activity on the Internet

Shawndra Hill[1], PhD; Jun Mao[2], MD MSc; Lyle Ungar[3], PhD; Sean Hennessy[4], PharmD PhD; Charles E Leonard[4], PharmD; John Holmes[4], PhD

[1]Operations and Information Management, The Wharton School, University of Pennsylvania, Philadelphia,PA, United States
[2]Department of Family Medicine and Community Health, School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[3]Department of Computer Science, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, United States
[4]Department of Biostatistics and Epidemiology, School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Shawndra Hill, PhD
Operations and Information Management
The Wharton School
University of Pennsylvania
3730 Walnut Street, Suite 500
Philadelphia,PA, 19104
United States
Phone: 1 2155735677
Fax: 1 2158683664
Email: shawndra@wharton.upenn.edu

Abstract

Background: As the incidence of H1N1 increases, the lay public may turn to the Internet for information about natural supplements for prevention and treatment.

Objective: Our objective was to identify and characterize websites that provide information about herbal and natural supplements with information about H1N1 and to examine trends in the public's behavior in searching for information about supplement use in preventing or treating H1N1.

Methods: This was a retrospective observational infodemiology study of indexed websites and Internet search activity over the period January 1, 2009, through November 15, 2009. The setting is the Internet as indexed by Google with aggregated Internet user data. The main outcome measures were the frequency of "hits" or webpages containing terms relating to natural supplements co-occurring with H1N1/swine flu, terms relating to natural supplements co-occurring with H1N1/swine flu proportional to all terms relating to natural supplements, webpage rank, webpage entropy, and temporal trend in search activity.

Results: A large number of websites support information about supplements and H1N1. The supplement with the highest proportion of H1N1/swine flu information was a homeopathic remedy known as Oscillococcinum that has no known side effects; supplements with the next highest proportions have known side effects and interactions. Webpages with both supplement and H1N1/swine flu information were less likely to be medically curated or authoritative. Search activity for supplements was temporally related to H1N1/swine flu-related news reports and events.

Conclusions: The prevalence of nonauthoritative webpages with information about supplements in the context of H1N1/swine flu and the increasing number of searches for these pages suggest that the public is interested in alternatives to traditional prevention and treatment of H1N1. The quality of this information is often questionable and clinicians should be cognizant that patients may be at risk of adverse events associated with the use of supplements for H1N1.

(J Med Internet Res 2011;13(2):e36) doi:10.2196/jmir.1722

KEYWORDS

Internet search; pandemic; herbal supplements; H1N1 influenza

Introduction

The 2009 pandemic of influenza A (H1N1) has affected at least 199 countries, with 482,300 confirmed cases and more than 6070 deaths worldwide as of November 1, 2009 [1]. These numbers are underestimates since most countries have stopped reporting cases, and mild cases are difficult to track. Although H1N1 influenza is no more severe than typical influenzas, its widespread dispersal will result in enormous human and economic cost to society.

The Web is a source for information on almost everything, and people search the Web from all over the world. So it is no surprise that when faced with emerging diseases where treatments are few and treatment distribution is limited, individuals often search the Internet for information, for example, on H1N1 prevention or treatment. In addition, new social media technologies such as Twitter enable people to follow reputable sources to get information about H1N1 as well as communicate about epidemics with peers [2]. This is particularly true for vitamin, mineral, and herbal supplement therapies, partly because of the wide use of such modalities [2,4,5]. In this paper, we are not only interested in search queries and their link to epidemics [3,6,7], but also in which webpages people land on after they search for vitamin, mineral, and herbal supplement treatments.

The goal of this study was to examine searches of the Internet for natural products used to prevent or treat H1N1. Although the definition of natural products may be different in different contexts, for this paper we focused on natural products that are not made of synthetic compounds and can be purchased over the counter. More specifically, we are interested in those that can be consumed orally, such as herbs, vitamins, natural supplements, or homeopathic products, since these products have a greater potential to cause harm compared with topical products or nonpharmacological therapies (eg, yoga and meditation) when used in conjunction with prescribed medications. For the purposes of this paper, we call these products *supplements*.

It has been shown that the information displayed about medical supplement products on the Web is often inaccurate or even fraudulent [8]. One study found that 25% of 150 websites on supplement information contained information that could lead to direct physical harm if acted upon [8]. Another study found that medical information websites containing information on supplements were more likely to contain inaccurate information [9]. Furthermore, some supplement products purchased over the Internet contain heavy metals that could negatively affect health [10]. Physicians are often unaware of the herbal and other supplements that their patients are taking, as patients rarely disclose such therapies to their physicians [11]. Despite the potential pitfalls of using online user-contributed content for health surveillance, it is hard to ignore the potential to harness the Web for surveillance [12], and there have been recent discussions in the literature on "infodemiology" frameworks for measuring the value of different types on online digital information [13].

To our knowledge, no published study has evaluated how the rise in the H1N1 epidemic relates to Internet searches regarding supplements. However, an understanding of this issue is critical from at least two perspectives. In terms of public health, monitoring public H1N1-related search activities may provide federal agencies with estimates of the prevalence of such information-seeking behavior. Improved awareness may prompt policy development to better regulate the quality of such publically disseminated information. In terms of clinical practice, clinicians would benefit from knowing what common herbs and natural products their patients are learning about to deal with H1N1 influenza. Such information can improve patient-physician communication to promote evidence-based supplement use.

The specific objectives of our investigation were twofold: to identify and characterize websites that provide information about H1N1 alongside information on supplements and to examine trends in the public’s information-seeking behavior with respect to supplement use for preventing or treating H1N1. To meet these objectives, we sought to answer the following research questions: (1) What is the availability of information about supplements and H1N1? (2) What supplements does the public search for, and how do search patterns change over time? (3) What types of websites provide information about supplements and H1N1?

Methods

Source of Data

We used three sources to compile a list of supplements for this investigation: drug information from the National Library of Medicine [14], the HubPage on homeopathic remedies [15], and a list of supplements compiled by one of the authors (JM). This list was based on a qualitative and exhaustive Internet search using the key words *herb*, *natural*, *flu*, *cold*, *swine flu*, and *H1N1* and was conducted on October 29, 2009. A controlled vocabulary was created by merging these lists, removing duplicates, and resolving abbreviations and synonymy. We used both common and botanical names where appropriate. The final version of the controlled vocabulary contained 145 search terms and is available as Multimedia Appendix 1.

Search Method

We used a simple computer program (available on request) to perform automated Internet searches with Google, which was used because it is the most frequently utilized search engine. Our program performed a search using each item in the controlled vocabulary described above in combination with both *H1N1* and *swine flu* separately as a single search term. Examples of the resultant single search terms are *H1N1 Vitamin A* and *swine flu Vitamin A*. The results of the search for each term were written to a machine-readable file as the term with the associated number of indexed pages (“hits”). This search was performed on November 15, 2009. We used our code instead of publicly available software because our queries were straightforward to perform and we could run in batch mode, in parallel, on our university grid computer.

Metrics and Analysis

We calculated several metrics for this investigation. The first metric was $F_{supplement}$, defined as the frequency of hits counted separately for each supplement together with H1N1. Thus, $F_{supplement}$ reflects the prevalence of information about supplements co-occurring with H1N1 on the Internet, regardless of the context, including advertising, mentions on discussion boards or other social media, or health-related contexts.

The second metric is $P_{supplement}$, defined as the proportion of total hits for a given supplement that also include H1N1 on the webpage:

The third metric we used is *page rank*, a metric calculated by Google that indicates the popularity and importance of a webpage [16]. The rank of a webpage is based on page rank and the number of “incoming” links (pages from which users have navigated to arrive at a webpage) and the number of hits the page receives over time. Links to webpages are listed by order of page rank, after sponsored links, on the results of queries returned by Google. Although most queries return multiple pages of results, we focused only on the first page of search results, since most users select links from this page without exploring the others [17]. We ordered page ranks as a range from 1 (most relevant) to 10 (least relevant).

Fourth, we evaluated the variability of webpages returned by each query using an information theoretic metric, *Shannon’s entropy*:

where entropy, or H(X), is the uncertainty of variable X, $P(X_i)$ is the probability of observing category *i* in the data, where *i* represents a specific webpage in our data, for example, the webpage for the Centers for Disease Control and Prevention (CDC). In our context, entropy reflects the number of different webpages that are returned in the top slot by queries on Google. We compared two query types: queries for the supplement alone and queries with each supplement plus H1N1. We use entropy to measure how stable the webpage returned in the top slot is; for example, do we always get the CDC’s webpage in the first slot when submitting our queries, or, instead, do we get many different webpages.

As a fifth metric, we used Google Trends [16] to investigate temporal trends in search behavior. This service samples Google Web-searching activity and reports the volume of searches for a specific term relative to all Google searches over time. We investigated the trends associated with the most frequently searched-for supplements (based on $P_{Supplement}$) together with the term *flu*. We were restricted to this term because Google Trends did not detect sufficient search activity using the more descriptive terms *swine flu* or *H1N1* combined with searches for supplements. We don’t believe this restriction limited our study because we compared Google trends curves for *flu* and *swine flu* and they have the very same shape over time with the same peaks at the height of the outbreak.

The observation period was from January 1, 2009, through November 15, 2009. The search volumes were compared with specific news events that occurred during this time period to examine the effect of these events on search behavior. We measured the probability of hits returned by the queries for supplement plus H1N1 by domain or function (such as .gov, .com, news, etc.). Finally, we identified the side effect and interaction profiles of the 20 supplements with the highest $P_{Supplement}$ values.

Results

Availability of Information on the Web.

There is a significant amount of information about the H1N1 pandemic on various webpages in the form of health alerts by governments, news stories, advertisements, and user-generated content on blogs, discussion boards, and popular social networking portals like Facebook and Twitter. Table 1 displays the supplements found on the 20 pages that most frequently included both a supplement from the controlled vocabulary and H1N1, ranked by $F_{supplement}$. Table 1 also displays the 20 supplements with the highest probability of appearing on webpages with H1N1, ranked by $P_{supplement}$. We measured the public’s interest in a particular therapy by the two statistics $F_{supplement}$ and $P_{supplement.}$ These measures could be used together or alone in our context

User Search Behavior: What Do Users Search for, and How Do Search Patterns Change Over Time?

While some users contribute useful information on the Web by posting to discussion boards and blogs, others are trying to find remedies and prevention techniques in the wealth of information provided by government agencies, news organizations, and users. Figure 1 shows the temporal patterns of queries on Google for the top 10 supplements ranked by the highest relative increase in search activity during the period of the H1N1 pandemic. The baseline is the average number of searches for a supplement-plus-flu pair during the period from January 1, 2004, through December 6, 2009. The relative increase relates the peak in searches for a specific supplement during the pandemic to its baseline. Prior to April 12, 2009, very limited supplement-plus-flu search activity was reported by Google Trends; therefore, for clarity, Figure 1 displays the trends after April 12, 2009.

Figure 1 depicts search activity over time starting at the time of the first reports concerning H1N1 during the week of April 26, 2009. A precipitous drop in searches for supplement plus H1N1 occurred immediately thereafter. With the exception of searches for Orange and Vitamin D, no further significant supplement search activity (relative to all Web search activity) was detected by Google Trends until mid-September, when increasing numbers of deaths due to H1N1 were reported in the media. Supplement-plus-H1N1 search activity peaked in the last week of October, when the number of H1N1-related deaths reached 1,000. Early November 2009 is when there were reports in the news of the availability of the H1N1 vaccine. In addition to the supplements in our list, we evaluated the Federal Drug Administration’s list of banned supplements [19] for search activity associated with flu. With the exception of Airborne, the supplements on the banned list did not yield enough search activity to register in Google Trends.

Types of Websites Providing Information About Supplements and H1N1.

The choice of query terms used for Internet searches determine the quality of information returned on the Web in terms of the websites referred to. Table 2 displays the sites listed on the first page of results from queries for H1N1 alone, supplements alone, or for supplement-plus-H1N1 pairs. The average page rank for the top 10 sites found for H1N1and for supplement-plus H1N1 query pairs were 8 and 6.4, respectively. Since higher-ranked pages are generally more authoritative, users querying for H1N1 alone arrived at sources that were more likely to be authoritative and/or actively curated and vetted by experts or organizations. For example, users were presented with the CDC website in the top slot when they queried for H1N1 alone. Table 2 demonstrates a shift from government and heavily curated sites for H1N1 alone to less authoritative or curated sites when querying for H1N1 plus supplement.

Table 1. Supplements with the 20 highest frequencies of occurrence (Fsupplement, left) and those with the 20 highest probabilities of co-occurrence with H1N1 (Psupplement, right).

Rank	Supplements With the Highest Frequencies of Occurrence	Frequency of Occurrence ($F_{supplement}$)	Supplements With the Highest Probablities of Co-occurrence With H1N1	Probability of Co-occurence with H1N1 ($P_{supplement}$)
1	Orange	2,570,000	Oscillococcinum	.543
2	Juice	1,450,000	Tinospora	.308
3	Vitamin D	877,000	Guduchi	.273
4	Vitamin C	827,000	Elderberry	.263
5	Onion	823,000	North American ginseng	.195
6	Green tea	649,000	Polyphenols	.116
7	Garlic	627,000	Divya giloy sat	.113
8	Ginger	581,000	Orange Juice	.105
9	Sage	427,000	Echinacea	.105
10	Rosemary	196,000	Andrographis	.103
11	Orange Juice	181,000	Ban Lan Gen	.098
12	Vitamin A	147,000	Flavonoids	.077
13	Selenium	116,000	Vitamin D	.076
14	Vitamin E	108,000	Dulcamara	.074
15	Peppermint	93,800	Elecampane	.069
16	Oscillococcinum	84,700	Jaggery	.064
17	Echinacea	82,000	Selenium	.064
18	Dulcamara	79,900	Mullein	.063
19	Elderberry	70,300	Eupatorium perforliatum	.061
20	Sulphur	68,900	Peppermint	.060

Figure 1. Supplement plus H1N1 queries from April 5, 2009, through November 15, 2009, as reported by Google Trends (note that only Orange had sufficient searches to appear throughout the entire period of time)

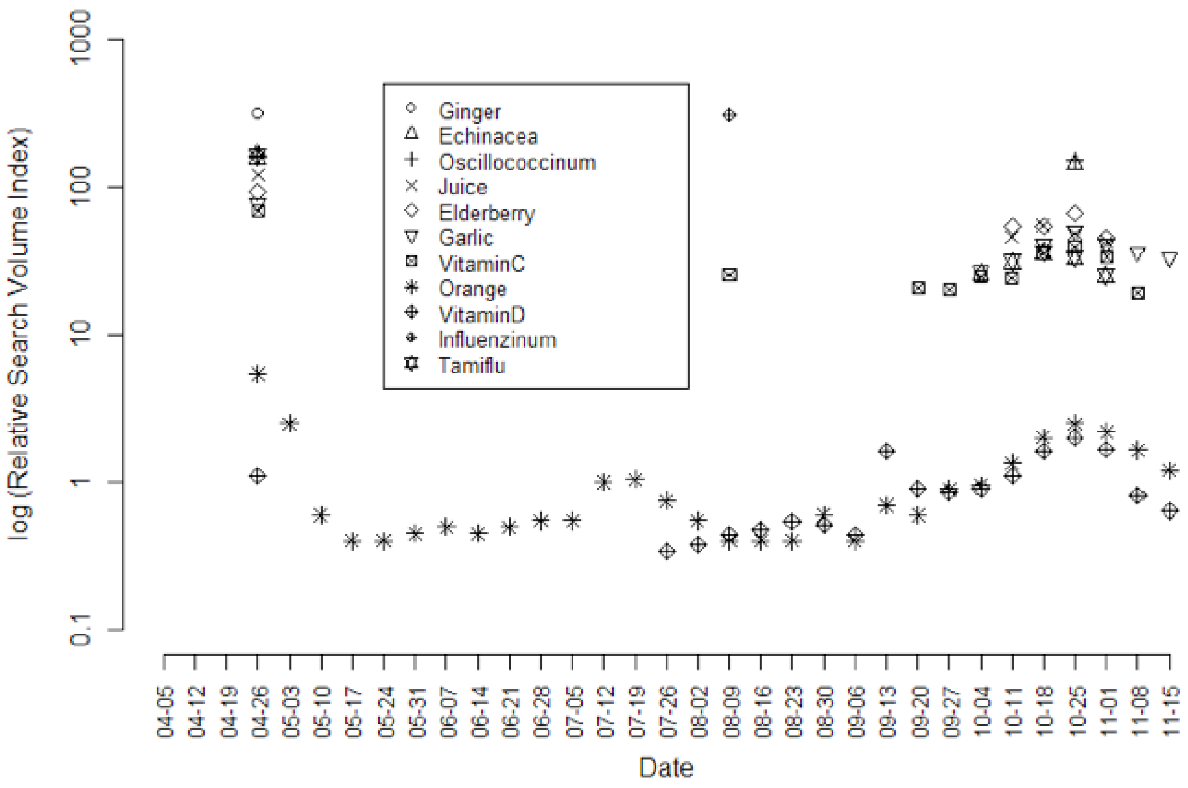
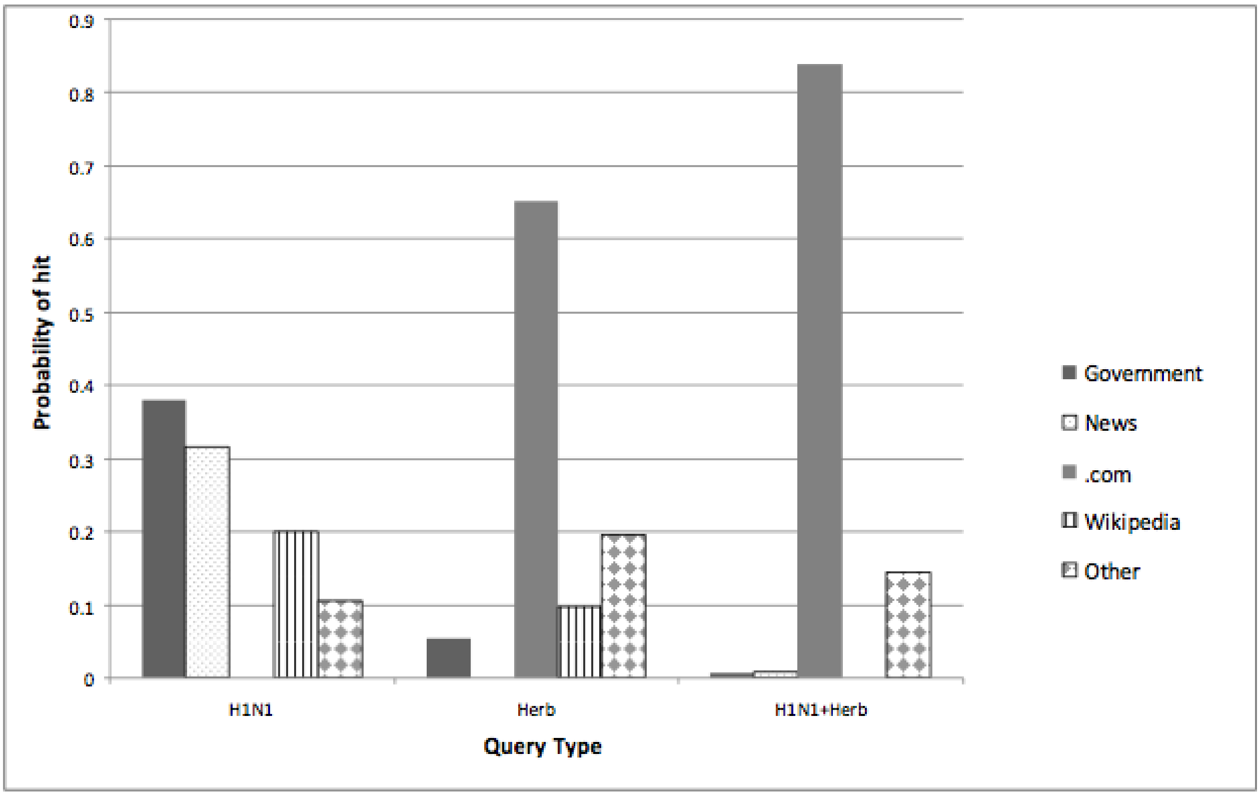


Table 2. Top 10 sources of information ranked by frequency of appearance on the first page of query results for H1N1 alone, each supplement alone, and each supplement plus H1N1

Rank	H1N1	Supplement	Supplement Plus H1N1
1	www.cdc.gov	en.wikipedia.org	ezinearticles.com
2	www.cdc.gov	images.google.com	hubpages.com
3	en.wikipedia.org	www.google.com	www.wellsphere.com
4	www.who.int	www.botanical.com	www.tcmwell.com
5	www.reuters.com	abchomeopathy.com	www.articlesbase.com
6	www.flu.gov	plants.usda.gov	www.ehow.com
7	news.yahoo.com	www.herbs2000.com	preventdisease.com
8	news.google.com	www.umm.edu	abchomeopathy.com
9	www.nlm.nih.gov	www.nlm.nih.gov	www.asiaone.com
10	www.cnn.com	www.elixirs.com	www.flutrackers.com

Figure 2. Probability distributions of hits by domain or site type for each of the three query types (distributions are based on the first page of search results)



We next calculated the proportion of hits that were from government, news, or other site types in the top 10 pages returned for all queries when querying for H1N1 alone, a supplement alone, or a supplement plus H1N1 (Figure 2). Of particular note is the change in site type proportion as the search term was changed from H1N1 alone to supplement alone to H1N1 plus supplement. The probability of arriving at a government website was .38 when querying for H1N1 alone, .05 when querying supplements alone, and only .006 when querying for supplement-plus-H1N1 pairs. When we combined supplements and H1N1 in the queries, no government entities appeared in the top 20 sites ranked by frequency of occurrence across all supplements in our sample.

Entropy of Query Results.

In addition to being of higher quality, the results of the query are also more stable as indicated by the entropy. When we queried for the 145 supplement-plus-H1N1 pairs, it was possible to return up to 145 different sites in the top slot of the first page of query results (for comparison, we performed 145 queries for H1N1 alone). If we calculated the entropy of the set of websites that are listed in the top slot when we queried for supplement plus H1N1 pairs, the result was an entropy of 0.83 as opposed to 0.51 when we queried for the supplements alone. When we queried for H1N1 alone, the top hit was always either CDC or Google News (news.google.com) resulting in an entropy of only 0.14.

Side Effect and Interaction Profiles of Commonly Searched Supplements

The known side effects and interactions of the 5 most commonly searched-for supplements on Google in the context of H1N1 are presented in Table 3.

Table 3. The 20 most frequently searched supplements in the context of H1N1, with their known uses, side effects, and interactions

Supplement	Description	Known Common Side Effects	Known Interactions
Oscillococcinum	Diluted extractions from duck liver and heart used for influenza. Contains no measurable amount of active ingredient. Not tested in pregnant women.	Unknown	Unknown
Tinospora (guduchi)	Derived from vine and used for diabetes, high cholesterol, allergic rhinitis (hayfever), upset stomach, gout, lymphoma and other cancers, rheumatoid arthritis, hepatitis, peptic ulcer disease, fever, gonorrhea, syphilis, and to boost the immune system. Not tested in pregnant women.	May reduce blood sugar in diabetics. May aggravate autoimmune diseases.	Oral hypoglycemic agents, immunosuppressants
Elderberry preparations	Used for influenza, HIV/AIDS, and boosting the immune system. It is also used for sinus pain, back and leg pain (sciatica), nerve pain (neuralgia), and chronic fatigue syndrome. Not tested in pregnant women.	May aggravate autoimmune diseases. Uncooked berries or juice can cause nausea, vomiting, and severe diarrhea.	Immunosuppressants
North American ginseng	Used for stress, to boost the immune system, to improve digestion, and as a general tonic and stimulant. Possibly unsafe in pregnant and breastfeeding women.	May cause low blood sugar, diarrhea, itching, insomnia, headache, and nervousness. Contains ginsenosides that may interfere with some estrogen-sensitive conditions.	Monoamine oxidase inhibitors, warfarin, oral hypoglycemic agents
Polyphenols	Derived from plants and includes tannins, lignins, and flavonoids. Have antioxidant properties.	Unknown	Unknown
Divya giloy sat	Ayurvedic herb with reported antiinflammatory, antipyretic properties, and immune-boosting properties. Minimal evidence exists and not tested in pregnant women.	Unknown	Unknown
Orange juice	Food product	Risk of hyperglycemia in diabetics	Unknown
Echinacea	An herb used for infections, especially the common cold and other upper respiratory infections. May decrease inflammation and boost immune system. Some limited clinical evidence and expert opinion that it may be safe in pregnancy in normal dosages.	May cause fever, nausea, vomiting, unpleasant taste, stomach pain, diarrhea, sore throat, dry mouth, headache, numbness of the tongue, dizziness, insomnia, disorientation, and joint and muscle aches. May aggravate autoimmune diseases.	Caffeine, medications metabolized by cytochrome P450 3A4 or cytochrome P450 1A2, immunosuppressants
Andrographis	Plant frequently used for preventing and treating the common cold and flu. Abortifacient	Side effects may include loss of appetite, diarrhea, vomiting, rash, headache, runny nose, and fatigue; and high doses or long-term use may cause swollen lymph glands, serious allergic reactions, and elevations of liver enzymes. May aggravate autoimmune diseases.	Antihypertensives, immunosuppressants, anticoagulants
Ban lan gen (isatis)	An herb that may have antibacterial, antiviral, antipyretic, antiinflammatory, and cancer-fighting activity. Not tested in pregnant women.	Unknown	Unknown
Flavonoids	Derived from plants, may have antiinflammatory properties. Not tested in pregnant women.	Unknown	Drugs metabolized by cytochrome P450 1A2, P-glycoprotein substrates, and anticoagulants

Supplement	Description	Known Common Side Effects	Known Interactions
Vitamin D	Vitamin used for many conditions, specifically used for boosting the immune system, preventing auto-immune diseases, and preventing cancer. Likely safe in pregnant women when used in daily amounts below 50 mcg (2000 units).	Too much vitamin D may cause weakness, fatigue, sleepiness, headache, loss of appetite, dry mouth, metallic taste, nausea, vomiting, and others.	Aluminum, calcipotriene, digoxin, diltiazem, verapamil, thiazide diuretics, cimetidine, heparin, Low molecular weight heparins.
Dulcamara	Stem from vine-like plant used for acne, itchy skin, boils, broken skin, warts, arthritis-like pain, nail bed swelling, eczema, diuretic, pain relief, and calming nervous excitement. Unsafe in pregnant women or in children.	Stem is safe, though leaves or berries are poisonous. Unsafe in children.	Unknown
Elecampane	Root from herb used for cough, asthma, bronchitis, nausea, diarrhea, worms in GI tract including hookworm, roundworm, threadworm, and whipworm. Unsafe in pregnant women.	Large amounts can cause vomiting, diarrhea, spasms, and paralysis. May cause drowsiness.	Central nervous system depressants
Jaggery	Unrefined sugar used in Ayurvedic medicine for treating lung and throat infections.	Unknown	Unknown
Selenium	Mineral used for cancer prevention, heart disease, rheumatoid arthritis, diabetes. Likely safe in pregnancy when used in low doses.	Taking high doses may cause nausea, vomiting, nail changes, loss of energy, and irritability. Poisoning from long-term use is similar to arsenic poisoning, with symptoms including hair loss, white horizontal streaking on fingernails, nail inflammation, fatigue, irritability, nausea, vomiting, garlic breath odor, and a metallic taste.	Anticoagulants including warfarin, statins, niacin, barbiturates, birth control pills, gold salts.
Mullein	Flower from plant that is used for influenza, herpes viruses, and respiratory infections. Not tested in pregnant women.	Unknown	Unknown
Eupatorium perfoliatum (boneset)	Dried leaf from plant used for cancer and bacterial infections. Is cytotoxic.	Unknown	Unknown

Discussion

In this paper, we identified a major concern regarding the types of websites on which one would most likely find information about supplements in the context of H1N1. Our ranked set of pages based on queries that contained supplements plus H1N1 or swine flu indicate that people may not be getting information from reliable sources (Table 3). Heavily curated sites that could be considered more mainstream (eg, Yahoo.com and cnn.com) or those that are more medically authoritative (eg, cdc.gov and flu.gov) did not contain supplement information in the context of H1N1.

In this study, we described the frequency with which supplements such as herbs, vitamins, and homeopathic products were displayed and how individuals searched for such information on the Web during emergence of the H1N1 pandemic. Information about the use of supplements for H1N1 was extensive, and user search activities increased and mirrored the rise in H1N1 incidence. This information was more likely to be found on alternative medicine and general information websites (eg, ehow.com), which raises the concern that those searching for supplements in the context of H1N1 will be taken to websites that are not clinically accurate, not curated for current information relative to the pandemic, or that focus on the sale of a particular supplement. One potential solution would be for government sites to provide objective information on the lack of rigorous evidence supporting the use of supplements for illnesses like H1N1 and to provide information about the potential adverse effects of supplements.

Since the Internet plays an increasing role in both public health communications and individual health–seeking behavior, this study illustrates the need for clinicians to be aware of the type and quality of health-related information available on the Internet. Our study also suggests the challenges and opportunities for Web content providers to deliver reliable and safe information to health consumers in a pandemic.

The large amount of information about supplements as they relate to H1N1 should not be surprising. Based on the 2007 report from the CDC, US consumers paid about US $14.5 billion to purchase nonvitamin, mineral, or other natural products and spent another US $2.9 billion for homeopathic medicines [20]. What is disconcerting is the lack of evidence for the efficacy for many of these products. Another major concern is the potential safety of these products and their potential interactions with conventional medications. For example, American ginseng is known to attenuate the effect of warfarin [21]. In manually inspecting the webpages returned by our searchers, we often encountered beneficial claims for these products without any mention of potential side effects or interactions with other supplements or drugs.

Our finding of an increased search activity for supplements concomitant with the rise of H1N1 incidence was probably driven by several factors. First, fear, uncertainty, and the emergence and rapid spread of an infectious disease are likely to motivate individuals to seek information. This is evident in the apparent correlation of news reports about H1N1 with search activity for supplements and Tamiflu. (Tamiflu is an antiviral drug used for the treatment of H1N1. The search activity associated with Tamiflu can be found in Figure 1.) Second, as new vaccines and treatments are developed, concerns about their safety and efficacy may lead individuals to seek alterative and natural ways of preventing or treating the disease. A recent public opinion poll found that only 40% of adults would definitely use the H1N1 vaccine if it were available [22]. Third, the distribution of effective prevention measures such as vaccines has been unequal and problematic. For example, the availability of H1N1 vaccine in primary care clinics and practices has been low, putting at risk those who are immunosuppressed or with serious chronic illnesses. Perceived unmet needs from conventional medical systems reportedly can prompt individuals to seek alternative therapies [23]. Our findings highlight that with a pandemic like the H1N1, public health agencies should consider providing on their websites objective information about supplements in the context of specific illnesses.

In an ideal clinical setting, physicians should inquire about patients' use of supplements and be knowledgeable about reported efficacy, safety, and potential side effects or drug interactions of common supplements. Prior research has suggested that patients who concomitantly use supplements and prescribed medications rarely disclose supplement use to their health care professionals. Such disclosure is even lower for Hispanic and Asian patients, who may have additional cultural and linguistic barriers [11]. Cultural beliefs about therapeutic benefit versus harm, prior experience with health communications, and a desire for greater control over their illness may be some of the factors that lead patients not to disclose use of supplements to their physicians [24]. Therefore, patients and health care professionals should be educated to discuss supplement use in an open and supportive way. This is particularly important among the chronically ill and elderly [11,25], where polypharmacy is highly prevalent in the context of concurrent supplement use. Because of the lack of education in supplements for physicians, curriculum content in supplement use and supplement-drug interactions need to be developed and taught in all undergraduate, postgraduate, and continuing medical education. Decision support tools on these issues need to be developed and available through electronic medical record or Internet to support quick clinical decision making for both physicians and patients.

Limitations

We based our first findings on results from Google queries. When people submit Web queries using a search engine such as Google, the resulting links may point to websites that no longer exist. Additionally, such sites may have a spurious association with the search terms (supplement and H1N1) due to random text matches, with both terms existing on the same page without being semantically related. Furthermore, the number of hits returned is an underestimate of the true number of webpages that contain the search terms since some websites are hidden and cannot be indexed by Google.

We based our second set of results on search query trends. Search queries are limited: a minimum number of searches is needed relative to all searches before Google Trends can provide trend data for a given term. Thus, we could not look at trends for some of the supplements with a high probability of co-occurrence with H1N1 ($P_{Supplement}$) but low frequency for the supplement alone ($F_{Supplement}$). Other issues include the inability of Google Trends to account for queries for misspelled words. For example, we would not see all of the queries for the supplement elderberry if it was misspelled as "eldaberry." Nonetheless, we do not think that any of these limitations influenced the direction of our results or our primary findings.

Conclusion

Although extensive information is available on the Internet with regard to natural products and H1N1, the source and quality of such information is questionable. During the emergence of H1N1, individuals actively searched for information on supplements for H1N1, and such Internet search patterns mirrored disease prevalence. Since the Internet will likely play an increasing role in human society, its delivery of timely, reliable, and scientifically sound information is critical to effective public health for dealing with current and future pandemic illnesses such as H1N1. Public health agencies should therefore consider providing on their websites objective information about supplements in the context of specific illnesses.

Acknowledgments

This project was supported by award number RC1LM010342 from the National Library of Medicine. Dr. Mao is supported by a career development award K23 AT004112-03 from the National Center for Complementary and Alternative Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Library of Medicine or the National Institutes of Health.

Conflicts of Interest

None declared

Multimedia Appendix 1

Search terms including tamaflu

[PDF file (Adobe PDF File), 57 KB - jmir_v13i2e36_app1.pdf]

References

1. World Health Organization. Pandemic (H1N1) 2009 - update 73. 2009. URL: http://www.who.int/csr/don/2009_11_06/en/index.html [accessed 2010-12-14] [WebCite Cache ID 5uyT0TQPp]
2. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
3. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248. [Medline: 17238340]
4. Barnes PM, Bloom B, Nahin RL. Complementary and alternative medicine use among adults and children: United States, 2007, National Health Statistics Reports, no 12. Hyattsville, MD: National Center for Health Statistics; 2008 Dec 10. URL: http://nccam.nih.gov/news/camstats/2007/ [accessed 2011-04-18] [WebCite Cache ID 5y2woK8F7]
5. Mao JJ, Farrar JT, Xie SX, Bowman MA, Armstrong K. Use of complementary and alternative medicine and prayer among a national sample of cancer survivors compared to other populations without cancer. Complement Ther Med 2007 Mar;15(1):21-29. [doi: 10.1016/j.ctim.2006.07.006] [Medline: 17352968]
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. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron A. More diseases tracked by using Google Trends. Emerg Infect Dis 2009 Aug;15(8):1327-1328 [FREE Full text] [Medline: 19751610]
8. Walji M, Sagaram S, Meric-Bernstam F, Johnson CW, Bernstam EV. Searching for cancer-related information online: unintended retrieval of complementary and alternative medicine information. Int J Med Inform 2005 Aug;74(7-8):685-693. [doi: 10.1016/j.ijmedinf.2005.01.001] [Medline: 16027034]
9. Bernstam EV, Walji MF, Sagaram S, Sagaram D, Johnson CW, Meric-Bernstam F. Commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. Cancer 2008 Mar 15;112(6):1206-1213 [FREE Full text] [doi: 10.1002/cncr.23308] [Medline: 18266210]
10. Saper RB, Phillips RS, Sehgal A, Khouri N, Davis RB, Paquin J, et al. Lead, mercury, and arsenic in US- and Indian-manufactured Ayurvedic medicines sold via the Internet. JAMA 2008 Aug 27;300(8):915-923 [FREE Full text] [doi: 10.1001/jama.300.8.915] [Medline: 18728265]
11. Mehta DH, Gardiner PM, Phillips RS, McCarthy EP. Herbal and dietary supplement disclosure to health care providers by individuals with chronic conditions. J Altern Complement Med 2008 Dec;14(10):1263-1269. [doi: 10.1089/acm.2008.0290] [Medline: 19032071]
12. Brownstein JS, Freifeld CC, Madoff LC. Influenza A (H1N1) virus, 2009--online monitoring. N Engl J Med 2009 May 21;360(21):2156 [FREE Full text] [doi: 10.1056/NEJMp0904012] [Medline: 19423868]
13. 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]
14. MedlinePlus. Drugs, Supplements and Herbal Information. Bethesda, MD: US National Library of Medicine, National Institutes of Health; 2011. URL: http://www.nlm.nih.gov/medlineplus/druginformation.html [accessed 2010-12-14] [WebCite Cache ID 5uySD3fId]
15. Licino, H. Stay Safe From The H1N1 Mexican Swine Flu: Homeopathic Remedies. HubPages URL: http://hubpages.com/hub/Stay-Safe-From-The-H1N1-Mexican-Swine-Flu-Homeopathic-Remedies [accessed 2010-12-14] [WebCite Cache ID 5uySNiFVT]
16. Langville AN, Meyer CD. Google's Page Rank and Beyond: The Science of Search Engine Rankings. Princeton, NJ: Princeton University Press; 2006.
17. Jansen BJ, Pooch U. A review of Web searching studies and a framework for future research. Journal of the American Society for Information Science and Technology 2001;52(3):235-246.
18. Google Trends. 2008. URL: http://www.google.com/trends [accessed 2010-12-26] [WebCite Cache ID 5vGcdm0gX]
19. US Food and Drug Administration. Fraudulent 2009 H1N1 Influenza Products List. Silver Spring, MD: US Department of Health and Human Services; 2010. URL: http://www.accessdata.fda.gov/scripts/h1n1flu/ [accessed 2010-12-14] [WebCite Cache ID 5uySsbG9T]
20. Nahin RL, Barnes PM, Stussman BJ, Bloom B. Costs of complementary and alternative medicine (CAM) and frequency of visits to CAM practitioners: United States, 2007. Natl Health Stat Report 2009 Jul 30(18):1-14. [Medline: 19771719]

21. Yuan CS, Wei G, Dey L, Karrison T, Nahlik L, Maleckar S, et al. Brief communication: American ginseng reduces warfarin's effect in healthy patients: a randomized, controlled Trial. Ann Intern Med 2004 Jul 6;141(1):23-27. [Medline: 15238367]
22. Blendon RJ, SteelFisher GK, Benson JM, Weldon KJ, Herrmann MJ. Influenza A(H1N1)/Swine Flu Survey III. Boston, MA: Harvard School of Public Health; 2009. URL: http://www.hsph.harvard.edu/news/press-releases/files/Swine_Flu_Topline_7.15.09.pdf [accessed 2010-12-26] [WebCite Cache ID 5vGd60j4t]
23. Mao JJ, Palmer SC, Straton JB, Cronholm PF, Keddem S, Knott K, et al. Cancer survivors with unmet needs were more likely to use complementary and alternative medicine. J Cancer Surviv 2008 Jun;2(2):116-124. [doi: 10.1007/s11764-008-0052-3] [Medline: 18648980]
24. Robinson A, McGrail MR. Disclosure of CAM use to medical practitioners: a review of qualitative and quantitative studies. Complement Ther Med 2004;12(2-3):90-98. [doi: 10.1016/j.ctim.2004.09.006] [Medline: 15561518]
25. Nahin RL, Fitzpatrick AL, Williamson JD, Burke GL, Dekosky ST, Furberg C, GEM Study Investigators. Use of herbal medicine and other dietary supplements in community-dwelling older people: Baseline data from the ginkgo evaluation of memory study. J Am Geriatr Soc 2006 Nov;54(11):1725-1735. [doi: 10.1111/j.1532-5415.2006.00942.x] [Medline: 17087700]

Abbreviations

CDC: Centers for Disease Control and Prevention

Edited by G Eysenbach; submitted 02.01.11; peer-reviewed by A Robinson, J Keelan; comments to author 26.01.11; revised version received 17.02.11; accepted 14.03.11; published 10.05.11.

Please cite as:
Hill S, Mao J, Ungar L, Hennessy S, Leonard CE, Holmes J
Natural Supplements for H1N1 Influenza: Retrospective Observational Infodemiology Study of Information and Search Activity on the Internet
J Med Internet Res 2011;13(2):e36
URL: http://www.jmir.org/2011/2/e36/
doi:10.2196/jmir.1722
PMID:21558062

Original Paper

Bias in Online Recruitment and Retention of Racial and Ethnic Minority Men Who Have Sex With Men

Patrick S Sullivan[1], DVM PhD; Christine M Khosropour[1], MPH; Nicole Luisi[1], MPH; Matthew Amsden[2], MBA; Tom Coggia[2], BSID MFA; Gina M Wingood[3], ScD MPH; Ralph J DiClemente[3], PhD MSc

[1]Rollins School of Public Health, Department of Epidemiology, Emory University, Atlanta, GA, United States

[2]Cyclogram, West Hollywood, CA, United States

[3]Rollins School of Public Health, Department of Behavioral Sciences and Health Education, Emory University, Atlanta, GA, United States

Corresponding Author:
Patrick S Sullivan, DVM PhD
Rollins School of Public Health
Department of Epidemiology
Emory University
1518 Clifton Road NE
Atlanta, GA, 30322
United States
Phone: 1 404 727 2038
Fax: 1 404 727 8737
Email: patrick.sullivan@emory.edu

Abstract

Background: The Internet has become an increasingly popular venue for men who have sex with men (MSM) to meet potential sex partners. Given this rapid increase in online sex-seeking among MSM, Internet-based interventions represent an important HIV (human immunodeficiency virus) prevention strategy. Unfortunately, black and Hispanic MSM, who are disproportionately impacted by the HIV epidemic in the United States, have been underrepresented in online research studies.

Objective: Our objective was to examine and quantify factors associated with underrecruitment and underretention of MSM of color in an online HIV behavioral risk research study of MSM recruited from an online social networking site.

Methods: Internet-using MSM were recruited through banner advertisements on MySpace.com targeted at men who reported in their MySpace profile their age as at least 18 and their sexual orientation as gay, bisexual, or unsure. Multivariable logistic regression models were used to estimate the odds stratified by race and ethnicity of the MySpace user clicking through the banner advertisement. To characterize survey retention, Kaplan-Meier survival curves and multivariable Cox proportional hazards models identified factors associated with survey dropout.

Results: Over 30,000 MySpace users clicked on the study banner advertisements (click-through rate of 0.37%, or 30,599 clicks from 8,257,271 impressions). Black (0.36% or 6474 clicks from 1,785,088 impressions) and Hispanic (0.35% or 8873 clicks from 2,510,434 impressions) MySpace users had a lower click-through rate compared with white (0.48% or 6995 clicks from 1,464,262 impressions) MySpace users. However, black men had increased odds of click-through for advertisements displaying a black model versus a white model (adjusted odds ratio [OR] = 1.83, 95% confidence interval [CI] 1.72 - 1.95), and Hispanic participants had increased odds of click-through when shown an advertisement displaying an Asian model versus a white model (adjusted OR = 1.70, 95% CI 1.62 - 1.79). Of the 9005 men who consented to participate, 6258 (69%) completed the entire survey. Among participants reporting only male sex partners, black non-Hispanic and Hispanic participants were significantly more likely to drop out of the survey relative to white non-Hispanic participants (hazard ratio [HR] = 1.6, 95% CI 1.4 - 1.8 and HR = 1.3, 95% CI 1.1 - 1.4, respectively). Men with a college-level of education were more likely to complete the survey than those with a high-school level of education (HR = 0.8, 95% CI 0.7 - 0.9), while men who self-identified as heterosexual were more likely to drop out of the survey compared with men who self-identified as gay (HR = 2.1, 95% CI 1.1 - 3.7).

Conclusions: This analysis identified several factors associated with recruitment and retention of MSM in an online survey. Differential click-through rates and increased survey dropout by MSM of color indicate that methods to recruit and retain black and Hispanic MSM in Internet-based research studies are paramount. Although targeting banner advertisements to MSM of color by changing the racial/ethnic composition of the advertisements may increase click-through, decreasing attrition of these study participants once they are engaged in the survey remains a challenge.

(J Med Internet Res 2011;13(2):e38) doi:10.2196/jmir.1797

KEYWORDS

HIV infections/prevention and control; Internet; homosexuality male; research methodology; behavioral research

Introduction

Men who have sex with men (MSM) are the most heavily impacted risk group for HIV in the United States [1]. Indeed, MSM is the only US risk group in which HIV incidence has been increasing since 2000 [2]. Similar to the United States, there is an international resurgence in HIV infections among MSM in industrialized countries in North America, Europe, and Australia [3]. Among MSM in the United States, men of color—especially younger men of color—are experiencing the most dramatic increases in HIV surveillance case reports [4,5]. At the same time, the number of HIV prevention interventions tested with MSM is low relative to the proportion of the US HIV epidemic represented by this vulnerable population [6].

The role of the Internet in the MSM epidemic—as a facilitator of transmission or as an opportunity for HIV prevention—is complex and unclear. In the results of most studies reported to date, men who meet sex partners on the Internet are also men who are more likely to engage in high risk sex [7-10] although the causality of this relationship is not clear. Regardless of causality, the Internet seems to hold promise as a way to reach men with significant sexual risk behaviors. We also know that there have been some early suggestions that Internet-based HIV prevention interventions hold promise as a strategy for accessing and increasing protective behaviors among MSM [11].

While the Internet represents a promising intervention strategy, significant gaps remain in our ability to optimally leverage the Internet to evaluate and implement HIV prevention interventions for MSM. One of the most significant barriers is that MSM of color, who bear the greatest risk for HIV infection in the United States, are systematically underrepresented in nearly all HIV prevention Internet studies to date (Table 1). Black men have been historically underenrolled by 29% to 84%, and Hispanic men underenrolled by 6% to 89%. Setting quotas for enrollment by race/ethnicity is one way to address the problem but raises additional concerns about bias among minority men who actually enroll (and how they are different from the unsurveyed men who do not enroll). Thus, underenrollment of MSM is a problem for Internet-based HIV prevention research and is ideally addressed by finding ways to obtain comparable rates of enrollment and retention for minority men.

Table 1. Selected Internet-based HIV prevention studies of men who have sex with men depicting population prevalence from recruitment location, enrolled study population prevalence, and corresponding prevalence ratio of black and Hispanic men

		Black Men			Hispanic Men		
Internet Study	Location	Population Prevalence (%)	Enrolled Prevalence (%)	Prevalence Ratio	Population Prevalence (%)	Enrolled Prevalence (%)	Prevalence Ratio
Grosskopf et al, 2010 [12]	New York City	25.1	17.9	0.71	27.4	13.5	0.49
Chiasson et al, 2009 [11]	United States	12.4	6.3	0.51	15.1	14.2	0.94
Rosser et al, 2009[a] [13]	United States	12.4	16.4	1.3	15.1	25.1	1.7
Berg et al, 2007 [14]	United States	12.4	2.5	0.20	15.1	1.7	0.11
Mackellar et al, 2007[b] [15]	6 US cities	25.3	8.6	0.34	30.2	18.8	0.62
Chiasson et al, 2007 [16]	United States and Canada	11.3	4.6	0.41	15.1	7.1	0.57
Bull et al, 2004[c] [17]	United States	12.4	6.6	0.53	15.1	10.9	0.72
Hirshfield et al, 2004 [18]	United States	12.4	2.0	0.16	15.1	5.5	0.36

[a] Recruitment was capped at 750 participants in each racial/ethnic group to ensure a diverse sample.

[b] The Web-Based HIV Behavioral Surveillance (WHBS) Study Group

[c] Recruitment strategies included print ads and flyers.

Our study addressed two aspects of this problem: (1) challenges in recruiting MSM of color (ie, the extent to which MSM of color click through banner advertisements to enroll in research studies) and (2) challenges in retaining MSM of color (ie, the extent to which MSM of color complete an online survey once they begin taking it). Our goals were to quantify the extent of underrecruitment of MSM of color, to identify factors associated with recruitment, and to determine the extent of underretention of MSM of color. To address these questions, we conducted an Internet survey of MSM recruited through banner advertisements shown on MySpace.com, a popular online social networking site.

Methods

Recruitment and Participation

Internet-using MSM were recruited from March 19, 2009, through April 16, 2009, through selective placement of banner advertisements on MySpace.com. During the recruitment period, advertisements were displayed to MySpace members based on self-reported demographic MySpace profile information. Exposures were made to males 18 years of age and over logging into MySpace whose profile indicated a residence in the United States and who reported their sexual orientation as gay, bisexual, or unsure. Participants who clicked through the banner advertisements were taken to an eligibility screener for an Internet-based survey.

A total of six banner advertisements were used, all with similar text and graphical design (Figure 1). Of the six, two of the advertisements presented a white male model, two presented a black male model, and two presented an Asian male model. Asian male models were used as controls since our hypotheses centered primarily on black/white differences. Data on the number of advertisement exposures and “click-throughs” were collected for each combination of model race, respondent race/ethnicity, respondent sexual orientation (gay, bisexual, unsure), and level of education (less than high school, high school, some college, college graduate and higher). Participant data on race/ethnicity, sexual orientation, and level of education for the “click-through” analysis were obtained from participants’ MySpace.com profiles. Therefore, these categories, which were preset options on MySpace.com, are not identical to those that are presented in the analysis of survey completion, where we used the data from participants’ survey responses (and which contained different category options for sexual identity and education level).

Figure 1. Shown are six banner advertisements displaying white (left), black (middle), and Asian (right) models used to recruit potential participants from MySpace.com for an online HIV behavioral risk study in the United States in 2009

Participants referred to the survey site after clicking through the banner advertisement were first screened for eligibility. Participants were eligible for the survey if they were male, 18 years of age and over, and reported at least one male sex partner in the last 12 months. Eligible participants were provided an online informed consent document, and consenting participants were passed into an online survey. In the survey, participants were asked for relevant demographic information as well as questions about the use of the Internet to meet sex partners, recent sexual risk behaviors, use of technology, HIV testing history, and interest in specific, new HIV prevention interventions. Participants did not receive an incentive to participate in the baseline survey. The study was reviewed and approved by the institutional review board (IRB) of Emory University.

Analysis

Characteristics Associated With Click-Through of Advertisements

Internet advertisement exposures and click-throughs were totaled and stratified by race/ethnicity of the MySpace user. Within each stratum of race/ethnicity, data were summarized by race of the model displayed in the advertisement, education, and sexual orientation of the respondent as reported in the MySpace profiles. We utilized multivariable logistic regression to model factors associated with clicking through the banner advertisement, stratified by race/ethnicity of the MySpace user. With click-through as the outcome and race of the model, education, and sexual identity as the independent factors of interest, we calculated adjusted odds ratios (ORs) with 95% confidence intervals (CIs) for each racial/ethnic group.

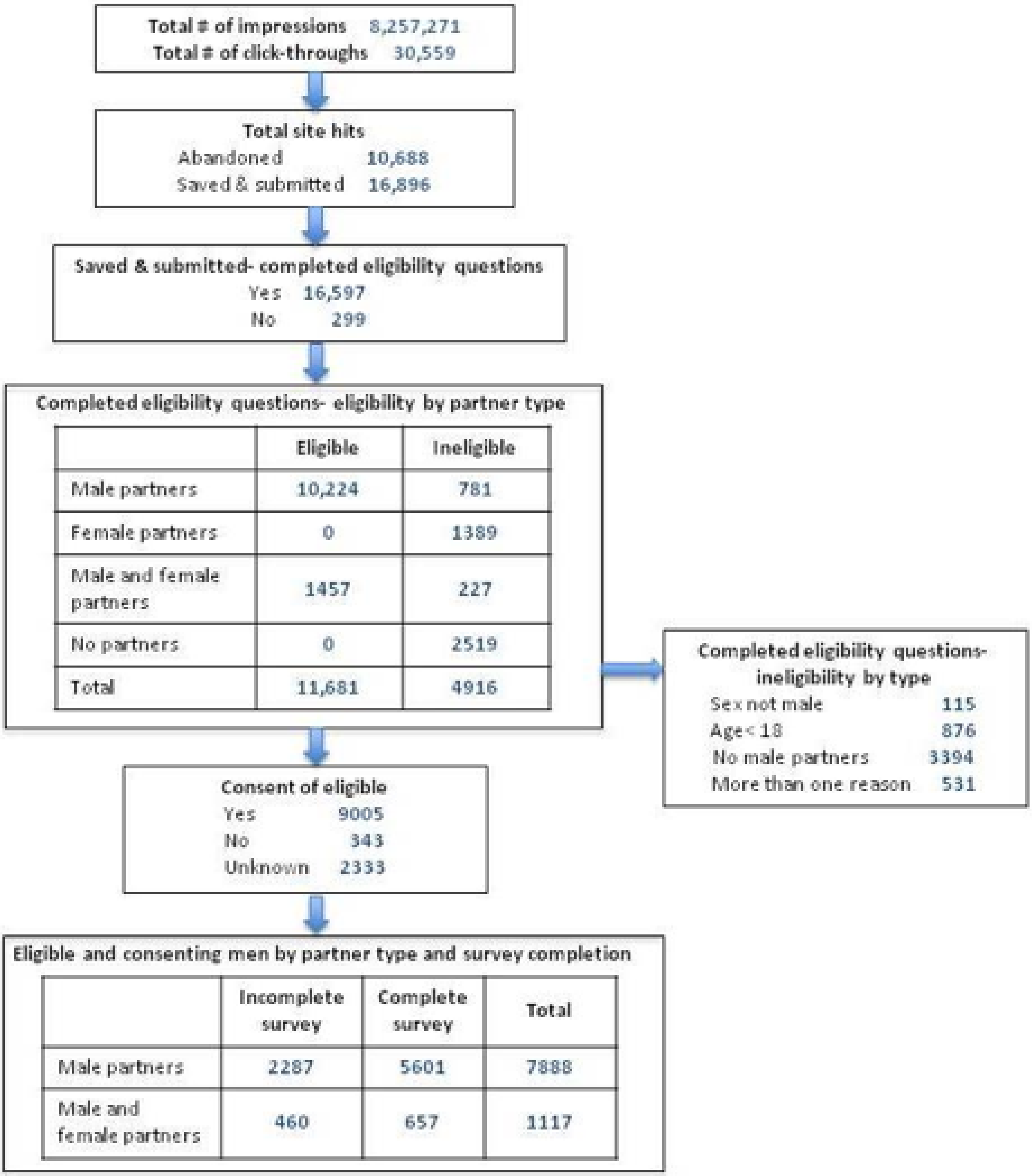
Characteristics Associated With Survival in the Survey

To determine whether there was any bias in completion of the survey, we used survival analysis to model factors associated with survey dropout. We evaluated survey dropout by race, age group, sexual orientation, education, geographic region, and gender of sexual partners in the last 12 months (male only or male and female) as reported in the online survey. Participants were assigned a *complete* or *incomplete* status for each page of the survey. A *complete* status was assigned if the participant answered at least 50% of the questions on the survey page. Failure (dropout from the survey) was defined as at least two consecutive incomplete pages. Kaplan–Meier survival methods and log-rank tests were used to evaluate survey dropout by individual demographic factors. For multivariable analysis, Cox proportional hazards regression was used to identify variables associated with survey dropout. Models were constructed separately for men who had only male partners and those who had both male and female partners; the reason for this was that men with male and female partners were asked additional questions about their female partners, which resulted in a longer survey. Thus, it was not possible to equitably assess dropout in one model that included both groups. The multivariable model findings are reported as hazard ratios (HR) with 95% CIs. All data analysis was completed using SAS version 9.2 (SAS Institute, Cary, North Carolina, USA).

Results

The recruitment and enrollment of study participants is described in Figure 2. A total of 8,257,271 MySpace advertising impressions over a 29-day period resulted in 30,559 (0.37%) click-throughs to our eligibility screener. Of those potential participants, 35% (10,688 of 30,559) abandoned the page without attempting the survey. Of the 16,896 respondents who saved or submitted their survey responses, 98% (16,597 of 16,896) completed the three sex, age, and partner questions used to determine eligibility. Approximately 30% (4916 of 16,597) of respondents who completed the screening questions were ineligible. Ineligibility was most often due to reporting no male sex partners in the past year (3394 of 4916 potential participants or 69%), followed by reporting an age of less than18 years (876 of 4916 potential participants or 18%). Thus, a total of 11,681 men were eligible for participation and were referred to informed consent. Another 20% (2333 of 11,681) of participants discontinued the survey at this point, not completing the informed consent process. Of those who completed the consent process, the vast majority, 96% (9005 of 9348), consented to participate.

Figure 2. Flow chart of participant recruitment, eligibility, and enrollment in an online HIV behavioral risk study conducted in the United States in 2009



	Eligible	Ineligible
Male partners	10,224	781
Female partners	0	1389
Male and female partners	1457	227
No partners	0	2519
Total	11,681	4916

	Incomplete survey	Complete survey	Total
Male partners	2287	5601	7888
Male and female partners	460	657	1117

Of the 9005 consenting participants, 39% (3473 of 9005) were white non-Hispanic, 14% (1293 or 9005) were black non-Hispanic, and 31% (2809 of 9005) were Hispanic. The median age of participants was 21, with 68% (6157 of 9005) of participants aged 18 to 24. Most reported having only male sex partners, and 12% (1117 of 9005) reported both male and female partners. In all, 69% (6176 of 8409 reporting) of the men self-identified as homosexual or gay, while 23% (2077 of 8409 reporting) self-identified as bisexual, and less than 1% (60 of 8409 reporting), as heterosexual or straight. Over half of participants (4815 of 8357 reporting) attended at least some college, and about a third (2914 of 8357) completed high school or a general equivalency diploma (GED). Over 60% of participants (4843 of 8063 reporting) had unprotected anal intercourse (UAI) with a male sex partner in the past 12 months, while over a quarter of participants (1775 of 6623 reporting) indicated they had never been tested for HIV.

Characteristics Associated With Click-Through of Advertisements

The overall click-through rate for the banner advertisements was 0.37% (30,599 clicks from 8,257,271 impressions); this varied by race, education, sexual identity, and race of model. In general, black, Hispanic, and men of other races were less likely to click through on the banner advertisements than were white men, with click-through rates of 0.36% (6474 clicks from 1,785,088 impressions) for black men, 0.35% (8873 clicks from 2,510,434 impressions) for Hispanic men, and 0.33% (8271 clicks from 2,497,487 impressions) for men of other races versus 0.48% (6995 clicks from 1,464,262 impressions) for white men. Men with more than a high school level of education were more likely to click through, as were men who self-identified as gay or bisexual. Click-through rates were higher when the models displayed were black (0.33% or 9006 clicks from 2,751,415 impressions) or Asian (0.48% or 13,241 clicks from 2,753,684 impressions) versus white (0.30% or 8312 clicks from 2,752,172 impressions) (data not shown).

Because our primary research question was focused on how to optimize banner advertisement to achieve comparable click-through rates for racial/ethnic minority men, we conducted multivariable analyses stratified by race/ethnicity (Table 2). Among all nonwhite racial/ethnic groups analyzed, men with greater than a high school education had higher odds of click-through compared with those with less education. Black, white, and Hispanic men that self-identified as gay or bisexual also had higher odds of click-through. Hispanic men had higher odds of click-through when exposed to Asian models; black men had higher odds of click-through when exposed to black models. White men had higher odds of click-through when shown an Asian model, but lower odds of click-though when shown a black model.

Table 2. Odds of clicking on study banner advertisements by MySpace.com users controlling for self-reported education, sexual identity, and race of model in advertisements and stratified by race of the MySpace.com user in the United States in 2009

	White Men	Black Men	Hispanic Men	Other Men
Characteristic	Adjusted OR (95% CI)	Adjusted OR (95% CI)	Adjusted OR (95% CI)	Adjusted OR (95% CI)
Education				
< High School (referent)				
> High School	0.99 (0.95 - 1.04)	*1.20 (1.14 - 1.26)* [a]	*1.05 (1.01 - 1.10)*	*1.10 (1.04 - 1.16)*
Identity				
Unsure (referent)				
Gay	*2.10 (1.98 - 2.24)*	*1.62 (1.53 - 1.71)*	*1.45 (1.38 - 1.52)*	*3.07 (2.88 - 3.28)*
Bisexual	*1.63 (1.53 - 1.74)*	*1.78 (1.67 - 1.89)*	*1.58 (1.49 - 1.67)*	*2.83 (2.63 - 3.04)*
Race of model				
White (referent)				
Black	*0.74 (0.70 - 0.79)*	*1.83 (1.72 - 1.95)*	1.05 (0.99 - 1.11)	0.95 (0.89 - 1.00)
Asian	*1.56 (1.47 - 1.64)*	*1.46 (1.37 - 1.56)*	*1.70 (1.62 - 1.79)*	*1.61 (1.52 - 1.69)*

[a] Results presented in italics denote significance at $P < .05$.

Characteristics Associated With Survival in the Survey

Of the 9005 participants, 69% (6258 of 9005) completed the entire survey, and 31% (2747 of 9005) dropped out of the survey after starting. Demographic characteristic of participants by survey completion status are provided in Table 3, and hazards of dropping out of the survey from the Cox proportional hazards model are described in Table 4. Figures 3 and 4 display the Kaplan-Meier curves for survival in the survey by participant race/ethnicity and sexual orientation, respectively, among participants reporting having only male partners in the past 12 months. Among white study participants, 77% (2670 of 3473) completed the entire survey compared with 66% (849 of 1293) of black participants and 71% (1987 of 2809) of Hispanic participants. Among those with only male sex partners in the last year, black men had about a 60% higher hazard of dropping out than white men, and Hispanic men had an increased hazard of about 30%. Among those with both male and female partners in the last year, both black men and Hispanic men had an increased hazard of dropout (30% and 50% increased hazard, respectively) compared with white men. Only 58% (35 of 60) of men who self-identified as heterosexual completed the survey compared with 75% (4653 of 6176) of men who self -identified as gay, resulting in heterosexually identified men having over twice the hazard of dropping out of the survey compared with men identifying as homosexual. Participants who reported less than a high school level of education had an increased hazard of dropout compared with those with greater than a high school level of education. There was no significant difference in hazard of dropout related to age or geographical region.

Table 3. Demographic characteristics of survey participants enrolled in an online HIV behavioral risk study by survey completion status (n = 9005) in the United States in 2009

Characteristics of Participants[a]	Completed Each Page in Survey (n = 6258) n	%	Did Not Complete Each Page in Survey (n = 2747) n	%
Race				
White[b]	2670	77	803	23
Black[b]	849	66	444	34
Hispanic	1987	71	822	29
Asian/Pacific Islander[b]	138	66	70	34
Native American/Alaska Native[b]	123	67	61	33
Multiracial[b]	332	73	120	27
Other[b]	100	75	33	25
Age (years)				
18-24	4181	68	1976	32
25-29	1014	70	432	30
30-34	431	75	142	25
35-45	477	77	146	23
> 45	155	75	51	25
Sexual identity				
Bisexual	1445	70	632	30
Homosexual or gay	4653	75	1523	25
Heterosexual or straight	35	58	25	42
Other[c]	70	73	26	27
Education				
College/postgraduate	969	75	315	25
Some college/associate degree	2694	76	837	24
High school or GED	2082	71	832	29
Less than high school	444	71	184	29
Sexual Partners, past 12 months				
One or more men	5601	71	2287	29
Both men and women	657	59	460	41
UAI with a male sex partner, past 12 months				
Yes	4064	84	779	16
No	2189	68	1031	32
Ever been tested for HIV				
Yes	4501	95	258	5
No	1674	94	101	6
Urban versus rural [d]				
Rural	2376	75	805	25
Urban	3608	72	1381	28

[a] Totals for most variables do not equal the total number of participants due to missing data.

[b] non-Hispanic

[c] Participants could write in a text response for “Orientation”; the most frequent responses were “queer,” “curious,” and “questioning.”

[d] The categorization of rural versus urban was based on population density (per square mile) of the respondents’ zip codes; respondents who lived in a zip code with a population density of < 1000 persons per square mile were considered to live in rural areas.

Table 4. Hazards of failure to complete all pages of an online HIV behavioral risk survey, by gender of sexual partners of participants in the past 12 months in the United States in 2009

Characteristics of Participants	Male Partners Only Hazard Ratio (95% CI)	Male and Female Partners Hazard Ratio (95% CI)
Race		
White[a] (referent)		
Black[a]	*1.6 (1.4-1.8)*[b]	1.3 (1.0-1.8)
Hispanic	*1.3 (1.1-1.4)*	*1.5 (1.2-1.9)*
Other[a]	*1.3 (1.1-1.5)*	1.0 (0.8-1.5)
Age (years)		
18-24 (referent)		
25-29	1.0 (0.9-1.1)	1.0 (0.8-1.3)
30-34	0.9 (0.8-1.1)	0.7 (0.4-1.1)
35-45	0.8 (0.7-1.0)	0.8 (0.5-1.2)
> 45	0.8 (0.6-1.1)	1.0 (0.5-2.2)
Sexual identity		
Homosexual or gay (referent)		
Bisexual	1.0 (0.8-1.1)	0.8 (0.6-1.1)
Heterosexual or straight	*2.1 (1.1-3.7)*	0.8 (0.4-1.4)
Other[c]	0.9 (0.5-1.5)	0.6 (0.3-1.1)
Education		
College/postgraduate	*0.8 (0.7-0.9)*	0.9 (0.7-1.3)
Some college/associate degree	*0.8 (0.7-0.9)*	1.0 (0.8-1.3)
High school or GED (referent)		
Less than high school	1.0 (0.8-1.2)	1.3 (0.9-1.9)
Urban versus rural [d]		
Rural	0.9 (0.8-1.0)	1.2 (1.0-1.5)
Urban (referent)		

[a] non-Hispanic

[b] Results presented in italics denote significance at $P < .05$.

[c] Participants could write in a text response for “Orientation”; most frequent responses were “queer,”“curious,” and “questioning.”

[d] The categorization of rural versus urban was based on population density (per square mile) of the participants’ zip codes; participants who lived in a zip code with a population density of < 1000 persons per square mile were considered to live in rural areas.

Figure 3. Retention in an online behavioral risk survey among participants reporting only male partners in the past 12 months, by race/ethnicity of the participants in the United States in 2009

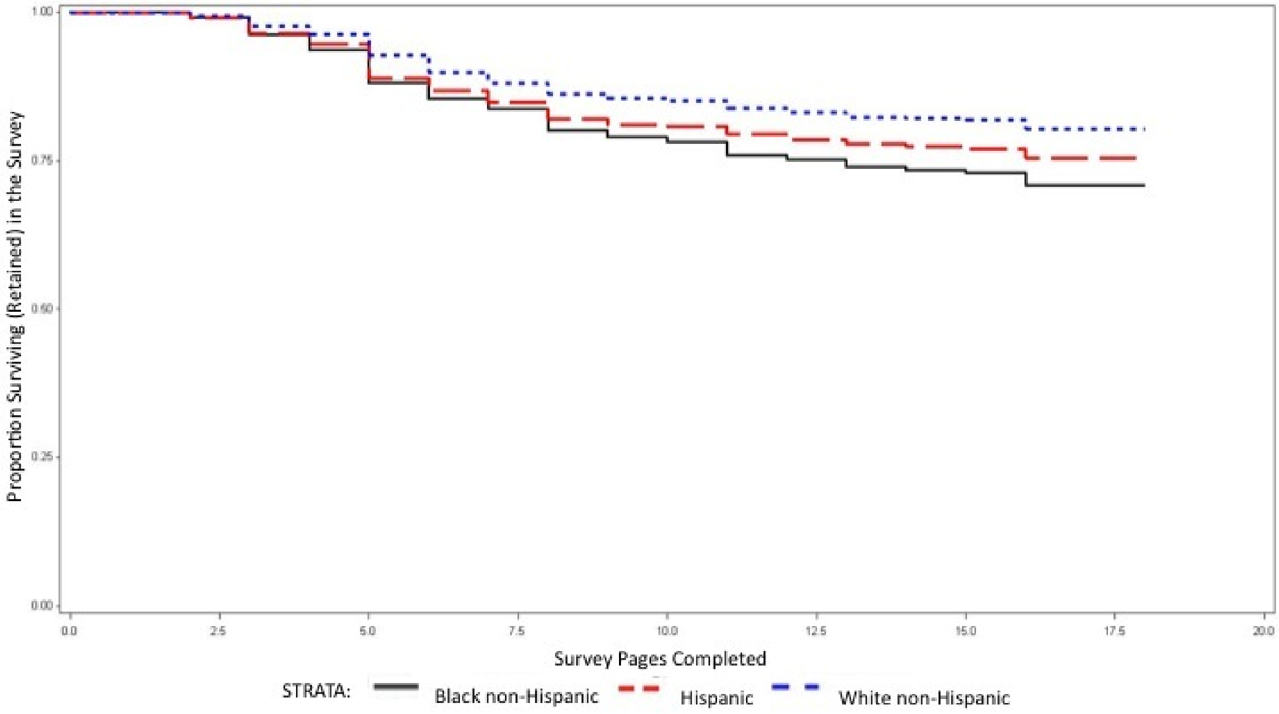
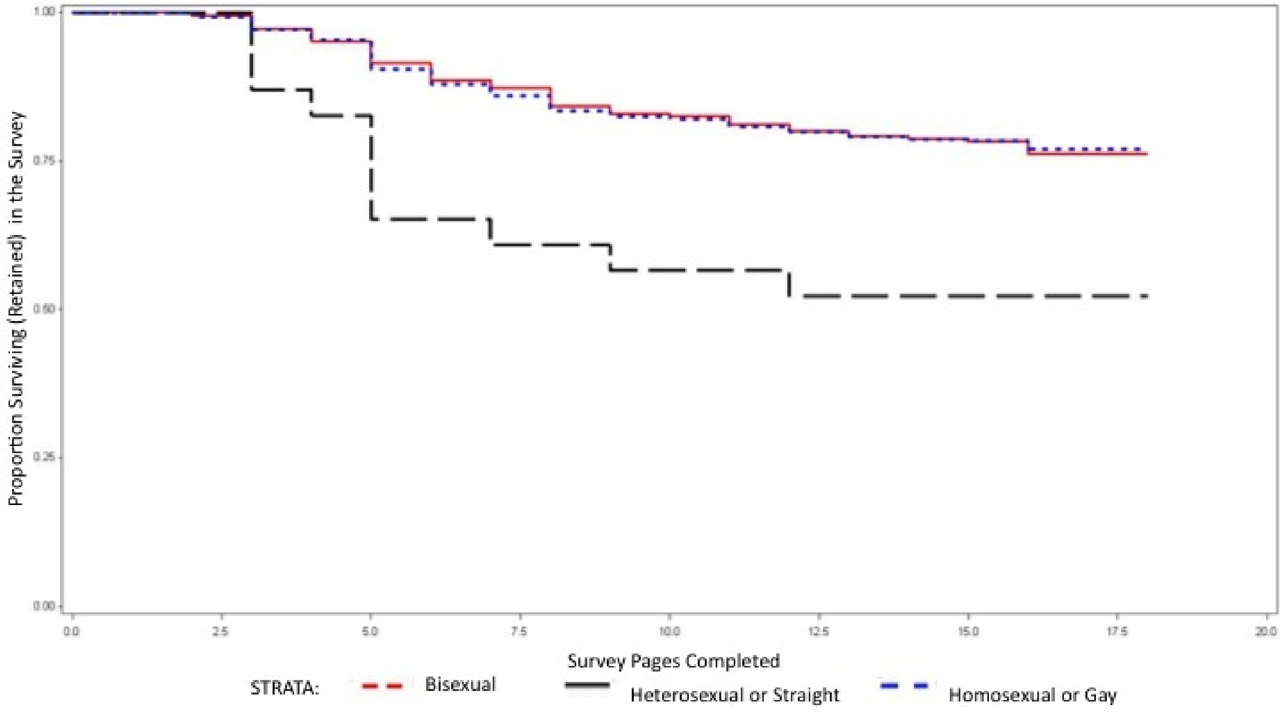


Figure 4. Retention in an online behavioral risk survey among participants reporting only male partners in the past 12 months by self-identified sexual orientation of the participant (United States, 2009)



Discussion

Our results illustrate two levels at which bias may occur in online HIV prevention research with MSM. Based on differential click-through rates, our data suggest that online surveys recruited through banner advertisements similar to ours may underrepresent black and Hispanic MSM, MSM with less education, and MSM who do not identify as gay. Once engaged in the survey, further bias is introduced by differential dropout during the survey; in this case, our data suggest that we were more likely to lose as participants men who were black, Hispanic, or of other nonwhite races, as well as men who were less educated, and men who identify as heterosexual.

Our approach furthers knowledge in this field because we used an Internet venue for advertisement in which we could collect data on each exposure to the banner advertisement (each impression) along with the race, education, and sexual identity of the person who saw the advertisement (as reported in their profile). This allowed us, for the first time to our knowledge, to calculate race-, education-, and sexual identity-specific click-through rates and model how the characteristics of Internet users who saw the banner advertisements related to their likelihood of clicking through that advertisement. The results indicate that the black men who saw our advertisements were much more likely to click through an advertisement displaying a black model than one displaying a white model, whereas the white men who saw our advertisements were much less likely to click through an advertisement with a black model than a white model. All of the racial/ethnic groups were more likely to click through on the advertisements displaying the Asian models compared with the advertisements displaying the white models; however, it is unclear why the Asian model advertisements yielded the highest click-through.

These data suggest that it may be possible to mitigate the differentially low click-through of black and Hispanic MSM by changing the racial/ethnic composition of the banner advertisements that they are shown. For example, black Internet users may be shown, as a group, banner advertisements that predominantly depict black models. It is important to emphasize that, to our understanding, an important goal is to achieve comparable click-through rates by racial or ethnic group. It will always be possible to set quotas to ensure that equal numbers of men in each racial/ethnic group are included, but the central question must be what biases attend the inclusion of those specific men. In other words, if 25% fewer black men are clicking through than white men, how are those men who are not clicking through different from the ones who are? Striving for comparable click-through rates by race/ethnicity is one way to address this concern.

Although MSM of color were also more likely to drop out of our survey prematurely, our data are less definitive in identifying solutions to this source of bias. However, we propose several possible reasons for our findings and options to reduce differential dropout. First, it is well documented that black Americans are less likely to have access to private, high speed Internet than are white Americans [19]. We speculate that, if black non-Hispanic participants were taking the current survey either without high-speed Internet access or in a public place (for example, a library), they may have discontinued participation early because of frustration with page load times or because of privacy concerns as questions became more personal. Possible solutions include optimizing our Internet interfaces for nonbroadband use or allowing an option to complete surveys via text messaging, as black and Hispanic Americans exceed white Americans on mobile phone ownership and text message use [20]. Another possible explanation for our results is that participants with less education drop out sooner because their literacy skills are lower. Possible solutions are attempting to lower the reading level of the survey and adding an option for text-to-speech service to the survey website. An additional explanation may be that the sensitive nature of the questions (ie, sexual history and HIV testing information) may have led some participants to discontinue their participation. An informal analysis of the pages from which men dropped out of the survey did not indicate that dropouts were clustered on a certain page; therefore, if the content of questions was a cause for dropout it was more likely due to the general nature of the topics and not to a specific question. However, it is important to note that bisexual- and straight-identified participants, although they had already reported having male sex partners, tended to drop out when the series of questions about male partners began. In future versions, we may choose to ask questions about male sex partners later in the survey, after questions about female sex partners.

Although not the primary focus of our study, we also note that social networking sites may be fruitful places for HIV prevention research among MSM. Most previous studies have recruited from sites that were either more explicitly gay-identified [13,16] or that were sex-seeking sites [11]. Using MySpace, we found that we were able to recruit a large number of behaviorally eligible MSM quickly. Also, the demographic and behavioral characteristics of these men suggest that they were at least as high risk and in need of prevention services as men recruited in "'real-world" venues. The greatest expansion in the US HIV epidemic among MSM is among MSM aged 13 to 24. The median age of our consenting participants recruited through a social networking site was 21; for comparison, the average age of participants in the 2008 National HIV Behavioral Surveillance System (NHBS) was 32 [21]. Also, our participants were relatively high risk: 60% reported unprotected anal intercourse with a male sex partner in the past 12 months compared with 47% of 2003-2005 NHBS respondents [22]. Use of prevention services among our participants was low, with 28% never having been tested for HIV versus 10% of 2008 NHBS respondents who had never been tested [21]. Thus, recruiting through a social networking site geared toward younger users yielded a young, high-risk sample with low levels of prior HIV testing—an ideal population for HIV prevention research and, eventually, Internet-delivered interventions.

There are a number of limitations to our study. First, our participants are not a representative sample, and our conclusions cannot be generalized to all MSM users of MySpace, to users of other social networking sites, or to US MSM more generally. Because we only showed our banner advertisements to men who identified themselves in their MySpace profiles as gay, bisexual, or unsure, we did not display banner advertisements to men who may have had male sex partners but who identified themselves as straight. Further, we could not verify that certain self-reported characteristics of our respondents—for example, male sex, or race/ethnicity—were correct. However, we recruited men based on their MySpace profile, and therefore we would only refer men to the eligibility screener if they had identified those characteristics in their profile, independent of the research study. Our data were subject to recall bias because of our 12-month recall period and to potential social desirability bias. There is also concern about participants taking the survey multiple times. In our case, we think that this was not common for two reasons. First, participants could only enter the survey site by being referred from a link in the banner advertisement. Therefore, a participant could only participate a second time if he was shown the banner advertisement a second time—a low probability event. Also, we did not allow multiple surveys to be completed from the same Internet protocol (IP) address, which would further require that a participant change his IP address, or take the survey a second time from a different computer.

Our study has several implications for those conducting online HIV prevention research with MSM. First, we have demonstrated that it is possible in some Internet settings to collect denominator data, characterize nonrespondents (ie, men who clicked on the advertisement but did not answer any survey questions), and model factors associated with participation. Second, our data suggest that social networking sites may offer an appropriate alternate recruitment venue to sex-seeking sites, especially for studies that seek to enroll younger, higher risk MSM. Finally, our results suggest the need for further work to attempt to reduce disparities in click-through rates for MSM of color and improve retention in surveys of MSM of color, of those with less education, and of those who are not gay-identified. These are difficult problems to address, but the epidemiology of the MSM epidemic in the United States demands that we address them, and increasing technological capacities of Web services where MSM congregate offer new opportunities to apply the highest standards of prevention research to online HIV prevention studies.

Acknowledgments

The authors disclose receipt of the following financial support for the research and/or authorship of this article: Emory Center for AIDS Research (P30 AI050409) and the National Center for Minority Health and Health Disparities (1RC1MD004370).

Conflicts of Interest

None declared

References

1. Diagnoses of HIV infection and AIDS in the United States and Dependent Areas, 2008.: Centers for Disease Control and Prevention; 2010. URL: http://www.cdc.gov/hiv/surveillance/resources/reports/2008report/index.htm [accessed 2010-12-16] [WebCite Cache ID 5v3FSBy7N]
2. Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, HIV Incidence Surveillance Group. Estimation of HIV incidence in the United States. JAMA 2008 Aug 6;300(5):520-529 [FREE Full text] [doi: 10.1001/jama.300.5.520] [Medline: 18677024]
3. Sullivan PS, Hamouda O, Delpech V, Geduld JE, Prejean J, Semaille C, Annecy MSM Epidemiology Study Group. Reemergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996-2005. Ann Epidemiol 2009 Jun;19(6):423-431. [doi: 10.1016/j.annepidem.2009.03.004] [Medline: 19460672]
4. Centers for Disease Control and Prevention (CDC). Subpopulation estimates from the HIV incidence surveillance system--United States, 2006. MMWR Morb Mortal Wkly Rep 2008 Sep 12;57(36):985-989 [FREE Full text] [Medline: 18784639]
5. Centers for Disease Control and Prevention (CDC). Increase in newly diagnosed HIV infections among young black men who have sex with men--Milwaukee County, Wisconsin, 1999-2008. MMWR Morb Mortal Wkly Rep 2011 Feb 4;60(4):99-102 [FREE Full text] [Medline: 21293324]
6. Sullivan PS, Zapata A, Benbow N. New U.S. HIV incidence numbers: heeding their message. Focus 2008 Nov;23(4):5-7. [Medline: 19062780]
7. Bolding G, Davis M, Hart G, Sherr L, Elford J. Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners? AIDS 2005 Jun 10;19(9):961-968. [Medline: 15905678]
8. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. J Acquir Immune Defic Syndr 2004 Apr 1;35(4):421-426. [Medline: 15097159]
9. Benotsch EG, Kalichman S, Cage M. Men who have met sex partners via the Internet: prevalence, predictors, and implications for HIV prevention. Arch Sex Behav 2002 Apr;31(2):177-183. [Medline: 11974643]
10. Liau A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. Sex Transm Dis 2006 Sep;33(9):576-584. [doi: 10.1097/01.olq.0000204710.35332.c5] [Medline: 16540884]
11. Chiasson MA, Shaw FS, Humberstone M, Hirshfield S, Hartel D. Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). AIDS Care 2009 Sep;21(9):1081-1089. [doi: 10.1080/09540120902730013] [Medline: 20024766]
12. Grosskopf NA, Harris JK, Wallace BC, Nanin JE. Online Sex-Seeking Behaviors of Men Who Have Sex With Men in New York City. Am J Mens Health 2010 Aug 26. [doi: 10.1177/1557988310372801] [Medline: 20798150]
13. Rosser BR, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL. HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's INTernet Sex Study-II (MINTS-II). AIDS Behav 2009 Jun;13(3):488-498. [doi: 10.1007/s10461-009-9524-3] [Medline: 19205866]
14. Berg RC. Barebacking among MSM Internet users. AIDS Behav 2008 Sep;12(5):822-833. [doi: 10.1007/s10461-007-9281-0] [Medline: 17676278]

15. Mackellar DA, Hou SI, Whalen CC, Samuelsen K, Sanchez T, Smith A, for the WHBS Study Group. Reasons for Not HIV Testing, Testing Intentions, and Potential Use of an Over-the-Counter Rapid HIV Test in an Internet Sample of Men Who Have Sex With Men Who Have Never Tested for HIV. Sex Transm Dis 2010 Dec 22. [doi: 10.1097/OLQ.0b013e31820369dd] [Medline: 21183863]
16. Chiasson MA, Hirshfield S, Remien RH, Humberstone M, Wong T, Wolitski RJ. A comparison of on-line and off-line sexual risk in men who have sex with men: an event-based on-line survey. J Acquir Immune Defic Syndr 2007 Feb 1;44(2):235-243. [doi: 10.1097/QAI.0b013e31802e298c] [Medline: 17179769]
17. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. AIDS Care 2004 Nov;16(8):931-943. [doi: 10.1080/09540120412331292507] [Medline: 15511725]
18. Hirshfield S, Remien RH, Humberstone M, Walavalkar I, Chiasson MA. Substance use and high-risk sex among men who have sex with men: a national online study in the USA. AIDS Care 2004 Nov;16(8):1036-1047. [doi: 10.1080/09540120412331292525] [Medline: 15511735]
19. US Census Bureau. 2009. Internet Use in the United States: October 2009 - Detailed Tables URL: http://www.census.gov/population/www/socdemo/computer/2009.html [accessed 2011-01-25] [WebCite Cache ID 5w0XsE6ge]
20. Smith A. Mobile Access 2010. Washington, DC: Pew Internet & American Life Project; 2010. URL: http://www.pewinternet.org/Reports/2010/Mobile-Access-2010.aspx [accessed 2011-01-23] [WebCite Cache ID 5w0Y095d0]
21. Centers for Disease Control and Prevention (CDC). Prevalence and awareness of HIV infection among men who have sex with men --- 21 cities, United States, 2008. MMWR Morb Mortal Wkly Rep 2010 Sep 24;59(37):1201-1207 [FREE Full text] [Medline: 20864920]
22. Sanchez T, Finlayson T, Drake A, Behel S, Cribbin M, Dinenno E, Centers for Disease Control and Prevention (CDC). Human immunodeficiency virus (HIV) risk, prevention, and testing behaviors--United States, National HIV Behavioral Surveillance System: men who have sex with men, November 2003-April 2005. MMWR Surveill Summ 2006 Jul 7;55(6):1-16 [FREE Full text] [Medline: 16826162]

Abbreviations

CI: confidence interval
GED: general equivalency diploma
HIV: human immunodeficiency virus
HR: hazard ratio
IP: Internet protocol
IRB: institutional review board
MSM: men who have sex with men
NHSB: National HIV Behavioral Surveillance System
OR: odds ratio
UAI: unprotected anal intercourse

Edited by G Eysenbach; submitted 08.03.11; peer-reviewed by J Rosenberger; comments to author 01.04.11; accepted 12.04.11; published 13.05.11.

Please cite as:
Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, DiClemente RJ
Bias in Online Recruitment and Retention of Racial and Ethnic Minority Men Who Have Sex With Men
J Med Internet Res 2011;13(2):e38
URL: http://www.jmir.org/2011/2/e38/
doi:10.2196/jmir.1797
PMID:21571632

Original Paper

Relationship Between Health Service Use and Health Information Technology Use Among Older Adults: Analysis of the US National Health Interview Survey

Namkee Choi[1], PhD

University of Texas at Austin, School of Social Work, Austin, TX, United States

Corresponding Author:
Namkee Choi, PhD
University of Texas at Austin
School of Social Work
1 University Station, D3500
Austin, TX, 78712-0358
United States
Phone: 1 512 232 9590
Fax: 1 512 471 9600
Email: nchoi@mail.utexas.edu

Abstract

Background: Older adults are the most frequent and heaviest users of health services in the United States; however, previous research on older adults' use of health information technology (HIT) has not examined the possible association of HIT use among older adults with their use of health services.

Objective: This study examined the relationship between US older adults' use of health services and their use of the Internet for health-related activities, controlling for socioeconomic characteristics and aging-related limitations in sensory and cognitive function. It also examined gender differences in the pattern of association between the types of health services used and HIT use.

Methods: The data for this study were drawn from the 2009 US National Health Interview Survey (NHIS), which was the first nationally representative household survey to collect data on HIT (Internet) use. First, the rates of lifetime and 12-month HIT use among sample adults (n = 27,731) by age group (18-29 to 85 and over) were analyzed. Second, bivariate analysis of sociodemographic characteristics, health status, and health service use by HIT use status among those aged 65 or older (n = 5294) was conducted. Finally, multivariate binary logistic regression analysis was used to test the study hypotheses with 12-month HIT use as the dependent variable and 12-month health service uses among the age group 65 or older as possible correlates.

Results: The rates of HIT use were significantly lower among the age groups 65 or older compared with the younger age groups, although the age group 55 to 64 was not different from those younger. The rates of HIT use decreased from 32.2% in the age group 65 to 74 to 14.5% in the age group 75 to 84 and 4.9% in the 85 and older age group. For both genders, having seen or talked to a general practitioner increased the odds of HIT use. However, having seen or talked to a medical specialist, eye doctor, or physical therapist/occupational therapist (PT/OT) were significantly associated with HIT use only for older women, while having seen or talked to a mental health professional only marginally increased the odds of HIT use only for older men. Having visited or talked to a chiropractor and having had overnight hospitalization, surgery, and/or homecare services were not associated with the odds of HIT use for either gender.

Conclusions: Older-adult users of general health services were more likely to use HIT than nonusers of general health services, while older-adult users of specialized health services were not different from nonusers of specialized health services in their odds of HIT use. The findings have implications for narrowing the age-related and socioeconomic status-related gaps in HIT use. The access gaps among racial/ethnic minority older adults and poorly educated and/or low-income older adults are especially striking and call for concerted efforts to facilitate Internet access and HIT use among these disadvantaged older adults.

(J Med Internet Res 2011;13(2):e33) doi:10.2196/jmir.1753

KEYWORDS

Health information technology; older adults; health service use; Internet use

Introduction

An increasing number of Americans use the Internet to search for health information and engage in other health-related activities such as participating in Internet chat groups, filling and refilling prescriptions, and using email to communicate with their healthcare providers [1,2]. However, studies have consistently found significant age group differences in the rates and frequencies of Internet use for health-related activities, with older adults (aged 65 and over) lagging behind younger ones [3-5]. Older adults, in general, are less likely than younger adults to use Internet technology and engage in Internet activities such as email, social networking, and accessing information such as news and weather. Age-related disparity in computer ownership, digital subscriber line, cable, or satellite Internet connections from home or work, levels of education and income, and levels of literacy and health literacy are likely to be contributing factors to the age disparity in the use of Internet technology for accessing health information and engaging in other health-related activities [6-8].

Among older adults, socioeconomic factors also have been found to be associated with Internet use versus nonuse. A study of older adults (aged 55-74) in Spain found that although Internet users appeared to have better self-rated health than nonusers, this relationship disappeared once social class (derived from the cross-classification of occupation and educational attainment of the family's primary income earner) was entered in the regression model [9]. In a US national public opinion survey, no African American or Hispanic American older adult in the sample reported going online for health information [4]. Other possible barriers to older adults' using the Internet for health-related activities include factors related to the aging process itself. One study found that the oldest adults (ages 85 and over) had negative reactions to using health information webpages that lacked the design accommodations for older adults recommended by the US National Institute on Aging and the National Library of Medicine [10]. This finding implies that for some older adults, visual and other sensory impairment and slowing information-processing capacity may be barriers to seeking health information online. Another study of the role of Internet knowledge and cognitive abilities in Web-based information seeking found that older adults (aged 60 and over) performed at a lower level than younger ones (aged 18-39) only when search problems were complex, a finding that also implies that older adults with age-related cognitive deficits may face barriers to using health information technology (HIT) [11].

Older adults are the most frequent and heaviest users of US health services, including visits to general practitioners, medical specialists, emergency departments, ambulatory surgeries, inpatient hospitalizations, and home health care [12]. However, previous research on older adults' use of HIT has not examined the possible association of their use of HIT with their use of health services. A 2001 survey of a representative adult sample of the US population found that more than 90% of the Internet health information seekers reported no impact of their Internet use on their numbers of visits to and telephone contacts with their physicians [1]. However, another national survey, conducted in 2003, found that 55% of Internet health information seekers contacted a health care professional because of information they had found online [13]. And those who had searched health information for a specific personal or loved one's health or medical condition were significantly more likely to contact a health care provider following their search than were seekers of information unrelated to a specific personal or loved one's health or medical condition. On the other hand, the same study also found that those who reported that they used Internet health information because it was free or because seeing a health professional was expensive were 90% less likely to contact a health care professional because of information found online than were those who did not mention cost factors [13]. Another study of Internet health information seeking among the chronically ill found that about 8% sought care from different doctors or providers than the ones they had been seeing because of the Internet information they had found and that about 30% used the Internet information to improve self-management of their conditions [14].

Although the samples that these studies used included older adults, the studies did not examine any age-specific pattern of association between health service use and HIT use. The purpose of the present study was to examine the relationship between US older adults' use of health services and their use of the Internet for health-related activities. When socioeconomic characteristics and aging-related limitations in sensory and cognitive function are controlled for, older adults who use health services may be more likely to have engaged in online health-related activities than their peers who do not use such services for the following reasons: (1) they are likely to have greater needs for health care information in order to manage their acute or chronic medical conditions; (2) they may want sources of health care information to supplement and enhance information and knowledge they obtain from their health care providers; and (3) they are more likely to be put in situations where they have to engage in certain online health-related activities (eg, filling or refilling prescriptions, scheduling medical appointments, and emailing their health care providers).

Specifically, this study examined the question of whether specific types of health services are more likely than others to be associated with older adults' HIT use and whether the relationship patterns differ by gender. Older adults rely heavily on their primary care physicians, who tend to be general practitioners, to deal with a variety of physical and mental health care needs ranging from preventive checkups and treatment to specialist referrals. Those who talked with or visited a general practitioner may be more likely than those who did not to have used HIT, because their talking with or visiting the doctor may indicate that they were having health problems and/or that they had a high level of health consciousness [15]. The findings of previous studies also suggest that those with chronic medical conditions and other serious illnesses (eg, cancer) may be more likely than others to search online health information [2,14,16]. Thus, it is possible that older adults who saw or visited a medical specialist or eye doctor or used such health services as inpatient hospitalization, surgery, physical therapy/occupational therapy (PT/OT), and home care may be more likely to have used HIT than their peers who did not use these health services.

Previous studies also showed that adults with mental health problems and other stigmatizing health conditions (eg, urinary incontinence or sexually transmitted diseases) were more likely to turn to the Internet for health information and communicate with a health care provider online [17,18]. One study also found that using the Internet increased health care use among those with psychiatric conditions [18]. Thus, older adults with mental health conditions may also be more likely to use online health information than discuss these conditions with those in their support network, or they may have visited mental health professionals after their online search for information about their mental health problems. Gender differences also needed to be examined, given the findings that women, including older women, use the Internet and HIT more than men [2,5,9,14].

This study tested the following hypotheses: Controlling for demographic and socioeconomic factors, self-reported sensory and memory limitations, and self-rated health, (1) older adults who had visited or talked to a general practitioner in the preceding 12 months compared with their peers who had not done so were more likely to have used HIT during the same period; (2) older adults who had visited or talked to a medical specialist or had used other health services (ie, eye doctor, PT/OT, chiropractor, inpatient hospitalization, surgery, and homecare) in the preceding 12 months were more likely to have used HIT during the same period than their peers who had not used these health services; (3) older adults who had visited or talked to a mental health professional in the preceding 12 months were more likely than their peers who had not done so to have used HIT during the same period; and (4) the pattern of association between the types of health services used and HIT use were likely to vary by gender. Although higher HIT use among women may be associated with their higher health service use, a directional hypothesis regarding the relationship between their HIT use and specific types of health services was not posited for lack of previous empirical data.

Methods

Data Source and Sample

The data for this study were drawn from the 2009 US National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics (NCHS). The NHIS, conducted annually since 1957, is designed to collect data on the health of nationally representative samples. The survey employs a multistage sample designed to represent the civilian, noninstitutionalized population of the United States. The interviewed sample for 2009 consisted of 33,856 households, which yielded 88,446 persons in 34,460 families. Of the 88,446 persons, 27,731 persons aged 18 or older were designated "sample adults" and were asked some additional questions. In 2009, the NHIS was to our knowledge the first nationally representative household survey to collect data on Internet use of health information and medical communication. The 10 questions asked of the sample adults were fielded in the HIT supplement [19].

The present study used the public-use data file for all 27,731 sample adults to describe the rates of Internet use of health information and medical communication among different age groups. Then the focus was on the sample adults who were aged 65 or older to examine the relationship between health service use and HIT use (ie, Internet use of health information and medical communication). Of the 5493 sample adults aged 65 or older, responses from 33 individuals who were not non-Hispanic white, non-Hispanic black, Hispanic, or Asian were excluded from the analysis along with an additional 166 whose responses were answered by proxy, resulting in the final analysis sample size of 5294 adults aged 65 or older.

Measures

Health Information Technology Use

Of 10 questions in the HIT supplement, 5 focused on Internet use for the following activities without a specific time frame. These were: (1) Have you ever looked up health information on the Internet? (2) Have you ever used chat groups to learn about health topics? (3) Have you ever refilled a prescription on the Internet? (4) Have you ever scheduled a medical appointment on the Internet? (5) Have you ever communicated with a healthcare provider by email? The other 5 questions focused on the respondent's Internet use for the same activities during the preceding 12 months (eg, Did you look up health information on the Internet in the past 12 months?). Summary measures of lifetime and 12-month HIT use represented engagement (coded 1) or no engagement (coded 0) in any of the 5 activities.

Health Service Use

The following nine types of health services used in the preceding 12 months were selected to be included in the analysis as they represent a wide range of health services that a significant proportion of older adults use: (1) saw or talked to a general practitioner, (2) saw or talked to a medical specialist, (3) saw or talked to an eye doctor, (4) saw or talked to a PT/OT, (5) saw or talked to a chiropractor, (6) was hospitalized overnight, (7) had any surgery, (8) used homecare services, and (9) saw or talked to a mental health professional.

Demographic, Socioeconomic, and Health Status Covariates

Demographic, socioeconomic, and health status covariates were gender, age, race/ethnicity (non-Hispanic black, Hispanic, Asian, and non-Hispanic white, the reference group); marital status (widowed, divorced/separated, never married, and married, the reference group); level of education (less than high school, general equivalency diploma (GED) or high school diploma, some college or associate's degree, bachelor's degree, and master's or doctoral degree, the reference group); family income-to-needs ratio (less than 1, 1-1.99, 2-3.99; missing, and 4 or higher, the reference group); paid work status (worked in the preceding 12 months vs did not work); any self-reported activity limitations due to a vision or hearing problem (yes vs no); any self-reported limitations due to difficulty remembering (yes vs no); and self-ratings of health (rated on a 5-point scale from 1, poor to 5, excellent). The latter was treated as a ratio-level variable.

Analysis Strategy

First, the rates of lifetime and 12-month HIT use by age group (18-29, 30-44, 45-54, 55-64, 65-74, 75-84, and 85 and over) in each of 5 activity areas as well as in summary measures were presented. Second, bivariate analysis of sociodemographic characteristics, health status, and health service use by HIT use status among those age 65 or older were presented. Finally, to test the study hypotheses, multivariate binary logistic regression analysis—for both genders and then separately for men and women—was done with the summary measure of 12-month HIT use as the dependent variable and the 12-month health service uses among the 65 and over age group as possible correlates. The respondents with missing information as to their level of education (n = 37) were excluded from the multivariate analysis. Because of the cross-sectional nature of the data, the relationships examined were correlational, not causal. Analyses were conducted with svy commands in Stata 11 (StataCorp LP, College Station, Texas, USA) to account for the NHIS's complex multistage sampling design.

Results

Health Information Technology Use in Different Age Groups

Table 1 shows that almost 51% of all adults reported that they had ever looked up health information on the Internet, but only 3.4% to 7% of them reported that they had ever used HIT for other health-related activities. Thus, the summary measure of lifetime HIT use—52.3% among all adults—appears to reflect that adults mostly had searched health information on the Internet but had not used it for other health-related activities. In the preceding 12 months, about 45% of all adults reported that they had looked up health information on the Internet, but only 3.2% to 5.9% of them reported that they had used HIT for other health-related activities. However, these average rates for all adults mask significant differences by age group, especially the differences in use between individuals in the age groups under 65 and the age groups 65 and over. For both lifetime and 12-month HIT use, the rates were significantly lower among individuals in the age groups over 65 than among individuals in the younger age groups. For example, more than half of those in the age groups under 55 and nearly half of those in the age group 55 to 64 had used HIT in the preceding 12 months compared with less than one third of those in the age group 65 to 74, less than one sixth of those in the age group 75 to 84, and less than one twentieth of the age group 85 and over. Gender differences or lack thereof by age group are also informative. In groups younger than age 65, the unadjusted rates of HIT use were significantly higher among women than among men, while in the age group 65 to 74, the rates were virtually the same. In the age groups 75 to 84 and 85 and over, the rates were significantly lower among women than men.

Table 1. Weighted percentage of persons who used health information technology by age group

Use of Health Information Technology	All	Age Group						
		18-29	30-44	45-54	55-64	65-74	75-84	85+
N=	27,731	5457	7087	5084	4360	2989	2002	752
Have ever (%)								
Looked up health information on the Internet	50.8	59.9	60.4	55.6	52.5	35.7	16.3	5.7
Used chat groups to learn about health topics	4.2	5.0	5.7	4.2	3.6	2.8	1.1	0.3
Refilled prescription on Internet	7.0	4.0	7.1	8.7	10.6	8.2	3.7	0.4
Scheduled medical appointment on Internet	3.4	3.9	4.5	4.0	3.2	1.6	0.8	0.3
Communicated with health care provider by email	5.8	4.4	7.4	7.5	6.8	4.2	1.5	0.4
Done any of the above (%)	52.3	61.3	61.8	57.1	54.2	37.2	17.4	6.1
Male[a]	47.6	53.8	52.9	49.7	49.6	37.0	20.5	11.3
Female[a]	56.1	68.3	69.5	63.6	58.1	37.4	15.6	3.6
In the preceding 12 months (%)								
Looked up health information on the Internet	44.8	53.3	54.1	48.4	46.4	30.3	13.3	4.5
Used chat groups to learn about health topics	3.2	4.0	4.5	3.3	2.7	1.9	0.9	0.1
Refilled prescription on Internet	5.9	3.3	5.5	7.4	9.5	7.2	3.3	0.4
Scheduled medical appointment on the Internet	2.6	2.9	3.4	3.2	2.5	1.3	0.7	0.1
Communicated with health care provider by email	5.8	4.4	7.4	7.5	6.8	4.2	1.5	0.4
Done any of the above (%)	46.5	54.7	55.6	50.2	48.8	32.2	14.5	4.9
Male[a]	41.4	46.4	46.0	42.7	44.3	32.2	17.5	9.3
Female[a]	50.7	62.5	64.0	56.8	52.5	32.3	12.8	2.8

[a] Gender difference in each age group, except the 65-74 group, is significant at $P < .01$.

Sample Characteristics Among Persons Aged 65 and Over by 12-Month HIT Use Status

Table 2 shows that older adults who used HIT were significantly different from their age peers who did not do so in terms of demographic, socioeconomic, and health characteristics, and in terms of the rates of utilization of all nine types of health services. As compared with nonusers, the users included higher proportions of men, persons in the age group 65 to 74, non-Hispanic whites, married persons, and those who had worked in the preceding 12 months but included lower proportions of persons without college education and with lower income (ie, income-to-needs ratio < 2). A significantly lower proportion of users than nonusers reported limitations due to sensory or memory problems, and self-rated health was higher among users than nonusers. Despite their higher self-ratings of health, a higher proportion of users had seen or talked to a general practitioner, medical specialist, eye doctor, PT/OT, chiropractor, and/or mental health professional, or had had surgery, but a lower proportion of users had had an overnight hospitalization or had used homecare services.

Table 2. Sample characteristics among those 65 and over by HIT use status in the preceding 12 months: Weighted statistics

Sociodemographic and health characteristics	All N=5294 (100%)	Did Not Use HIT N=4078 (77%)	Used HIT N=1215 (23%)	*P* Value[a]
Gender				< .001
Male	38.9	37.6	43.3	
Female	61.1	62.4	56.7	
Age group				< .001
65-74	52.9	46.3	74.7	
75-84	34.4	37.9	22.7	
85+	12.7	15.7	2.6	
Race/ethnicity				< .001
Non-Hispanic white	81.1	78.5	89.8	
Non-Hispanic black	9.6	10.9	5.2	
Hispanic	6.3	7.5	2.3	
Non-Hispanic Asian	3.0	3.1	2.7	
Marital status				< .001
Married/cohabiting	43.6	38.5	60.5	
Widowed	36.7	41.8	19.5	
Divorced/separated	14.8	14.5	16.0	
Never married	4.9	5.2	4.0	
Education				< .001
< High school	22.4	27.9	4.0	
High school diploma or GED	32.2	35.4	21.3	
Some college or associate's degree	23.8	21.7	31.0	
Bachelor's degree	12.5	9.5	22.6	
Some graduate school, MA/MS/PhD degree	8.5	4.8	20.9	
Missing	0.6	0.7	0.2	
Family income-to-need ratio				< .001
<1	10.1	12.4	2.4	
1-1.99	19.4	22.1	10.3	
2-3.99	26.4	25.8	28.2	
4+	22.4	16.2	43.1	
Missing	21.8	23.5	16.0	
% worked in the preceding 12 months	19.3	15.6	31.7	< .001
% reporting limitation due to hearing/vision problem	4.2	4.7	2.3	< .001
% reporting limitation due to memory impairment	6.9	7.8	3.9	< .001
Self-ratings of health (1=poor, 5=excellent), mean (SD)	3.31 (1.09)	3.21 (1.10)	3.63 (1.09)	< .001
Health care service use in 12 months				
% saw or talked to a general practitioner	85.8	84.7	89.5	< .001
% saw or talked to a medical specialist	44.9	41.8	55.0	< .001
% saw or talked to an eye doctor	57.3	55.0	65.0	< .001

Sociodemographic and health characteristics	All	Did Not Use HIT	Used HIT	P Value[a]
	N=5294 (100%)	N=4078 (77%)	N=1215 (23%)	
% saw or talked to a PT/OT	14.0	12.4	19.3	< .001
% saw or talked to a chiropractor	8.7	7.9	11.4	< .001
% had overnight hospitalization	17.1	17.8	14.9	0.019
% had any surgery	18.6	17.2	23.3	< .001
% used homecare services	7.5	8.4	4.6	< .001
% saw or talked to a mental health professional	3.7	3.0	5.9	< .001

[a] *P* denotes difference between nonusers and users shown from chi-square tests or independent samples *t* tests.

As expected, a majority of older adults had seen/talked to a general practitioner. Further analysis (not shown in Table 2) found that those who had not seen/talked to a general practitioner were younger than those who had done so (mean 73.74 [SD 6.55] vs mean 74.78 [SD 6.67], $P < .001$) and that they had significantly better self-ratings of health (mean 3.61 [SD 1.09] vs mean 3.26 [SD 1.08], $P < .001$) and fewer chronic illnesses (mean 0.21 [SD 0.61] vs mean 0.34 [SD 0.78], $P < .001$ when hypertension, arthritis, diabetes, heart disease, stroke, lung disease, and cancer were included). The nonusers of general practitioners' service were also less likely to have seen/talked to other health care providers, even though they did not differ from the users in educational level, income, and Medicare coverage. Thus, it appears that those who had not seen/talked to a general practitioner had less need for health services than those who had done so. On the other hand, those who had seen/talked to a medical specialist had significantly lower self-ratings of health (mean 3.12 [SD 1.09] vs mean 3.46 [SD 1.07], $P < .001$) and more chronic illness (mean 0.43 [SD 0.89] vs mean 0.22 [SD 0.63], $P < .001$ when hypertension, arthritis, diabetes, heart disease, stroke, lung disease, and cancer were included) than those who had not seen/talked to a medical specialist. However, these two groups did not differ in age.

Further analysis (not shown in Table 2) also found that women were older than men (mean 75.07 [SD 6.71] vs mean 73.95 [SD 6.53], $P < .001$) but did not differ from them in self-ratings of health. A higher proportion of women than men had seen or talked to a general practitioner and a medical specialist, had had overnight hospitalization and surgery, or had received homecare services, while no gender differences were found in the rates of seeking/talking to an eye doctor, a PT/OT, and/or a chiropractor.

Relationship Between Health Service Use and Health Information Technology Use

For older adults of both genders, binary logistic regression results (model likelihood ratio $\chi^2_{29} = 1433.64$, $P < .001$) in Table 3 show that having visited or talked to a general practitioner, medical specialist, eye doctor, PT/OT, and/or mental health specialist in the preceding 12 months increased a person's odds of having used HIT in the same period. On the other hand, having seen or talked to a chiropractor and having had overnight hospitalization, surgery, and/or homecare services were not significantly associated with HIT use. Although use of mental health services, compared with physical health services, appears to be highly correlated with HIT use, the odds ratios indicate that visiting or talking to other health care providers (ie, general practitioner, medical specialist, eye doctor, or PT/OT) had similar odds of increased HIT use. Significant covariates were gender, age, race/ethnicity, education, income-to-needs ratio, and self-ratings of health. Female gender was associated with higher odds of HIT use, while older age, being non-Hispanic black, Hispanic, or Asian, and being unmarried were associated with decreased odds of HIT use. As opposed to holding a master's or doctoral degree, all the other educational levels were associated with decreased odds of HIT use. With respect to family income-to-needs ratio, those with levels between 2 and 3.99 were not different from those with levels greater than 4, but lower ratios or missing categories were associated with decreased odds of HIT use. Higher self-ratings of health were associated with increasing odds of HIT use, but self-reported limitations due to sensory or memory problems were not associated with HIT use.

Gender-separate analysis found gender-neutral as well as gender-specific correlational patterns (model likelihood ratio $\chi^2_{28} = 588.49$, $P < .001$ for men and model likelihood ratio $\chi^2_{28} = 887.96$, $P < .001$ for women). For both men and women, having seen or talked to a general practitioner increased the odds of HIT use. However, having seen or talked to a medical specialist, eye doctor, or PT/OT were significantly associated with HIT use only for older women. Having visited or talked to a chiropractor or having had overnight hospitalization, surgery, and/or homecare services were not associated with the odds of HIT use for either gender. Interestingly, when gender-separate analysis was done, having seen or talked to a mental health professional only marginally ($P = .06$) increased the odds of HIT use only for older men, and it was not significantly associated with older women's HIT use.

As compared with non-Hispanic white men, non-Hispanic black and older Hispanic men, but not Asian men, had lower odds of having used HIT, while Hispanic and older Asian women, but not non-Hispanic black women, had lower odds of having used HIT. Men's marital status was not a factor significantly associated with the odds of their HIT use, while all single women had lower odds of having used HIT than all married women. Women with bachelor's degrees did not differ from women with master's or doctoral degrees. With respect to family income-to-needs ratio, women with levels between 1 and 3.99

did not differ from women with ratios greater than 4 in their odds of HIT use, but those with income-to-needs ratios less than 1 or with missing income data were less likely than women with ratios greater than 4 to use HIT. For men, however, level of education and income-to-needs ratio appeared to have a linear relationship with the odds of HIT use. Work status and limitations due to sensory or memory problems were not significantly associated with either gender, and self-rating of health was a significant factor only for women.

Table 3. Relationship between health information technology use and health service utilization in the preceding 12 months: Logistic regression analysis results

Predictor	All (N=5256)		Men (N=2059)		Women (N=3197)	
	Odds Ratio (SE)	95% CI	Odds Ratio (SE)	95% CI	Odds Ratio (SE)	95% CI
Gender						
Male	1.00	1.00				
Female	1.24 (0.13)[c]	1.01-1.52				
Age	0.91 (0.01)[a]	0.89-0.92	0.93 (0.01)[a]	0.91-0.96	0.89 (0.01)[a]	0.87-0.91
Race/ethnicity						
Non-Hispanic white	1.00	1.00	1.00	1.00	1.00	1.00
Non-Hispanic black	0.61 (0.12)[b]	0.42-0.89	0.49 (0.14)[c]	0.27-0.86	0.72 (0.17)	0.45-1.14
Hispanic	0.38 (0.08)[a]	0.25-0.56	0.52 (0.15)[c]	0.30-0.92	0.29 (0.08)[a]	0.16-0.50
Asian	0.55 (0.12)[b]	0.36-0.85	0.68 (0.23)	0.34-1.34	0.47 (0.14)[b]	0.27-0.83
Marital status						
Married/cohabiting	1.00	1.00	1.00	1.00	1.00	1.00
Widowed	0.55 (0.06)[a]	0.44-0.70	0.78 (0.17)	0.51-1.20	0.47 (0.07)[a]	0.35-0.62
Divorced/separated	0.69 (0.09)[b]	0.53-0.89	0.73 (0.16)	0.48-1.11	0.60 (0.11)[b]	0.41-0.86
Never married	0.45 (0.12)[b]	0.27-0.75	0.54 (0.23)	0.24-1.23	0.39 (0.11)[a]	0.23-0.68
Education						
Some graduate school/MA/MS/PhD	1.00	1.00	1.00	1.00	1.00	1.00
< High school	0.08 (0.02)[a]	0.05-0.13	0.06 (0.02)[a]	0.03-0.11	0.12 (0.04)[a]	0.06-0.22
High school diploma or GED	0.22 (0.03)[a]	0.16-0.30	0.16 (0.04)[a]	0.10-0.25	0.31 (0.06)[a]	0.21-0.47
Some college/associate’s degree	0.47 (0.07)[a]	0.35-0.63	0.37 (0.08)[a]	0.24-0.56	0.63 (0.13)[c]	0.42-0.94
Bachelor’s degree	0.70 (0.12)[c]	0.49-0.99	0.50 (0.11)[b]	0.32-0.78	1.03 (0.23)	0.67-1.60
Family income-to-need ratio						
4+	1.00	1.00	1.00	1.00	1.00	1.00
<1	0.32 (0.08)[a]	0.19-0.54	0.30 (0.13)[b]	0.13-0.75	0.39 (0.12)[b]	0.21-0.72
1-1.99	0.58 (0.09)[a]	0.44-0.78	0.45 (0.12)[b]	0.27-0.85	0.72 (0.14)	0.48-1.07
2-3.99	0.84 (0.10)	0.67-1.07	0.57 (0.11)[b]	0.39-0.78	1.15 (0.18)	0.85-1.58
Missing	0.58 (0.07)[a]	0.44-0.75	0.51 (0.11)[b]	0.33-0.78	0.70 (0.12)[c]	0.51-0.97
Did not work in 12 months	1.00	1.00	1.00	1.00	1.00	1.00
Worked	1.10 (0.12)	0.89-1.37	1.05 (0.18)	0.75-1.47	1.23 (0.19)	0.91-1.66
No hearing/vision problem	1.00	1.00	1.00	1.00	1.00	1.00
Hearing/vision problem	0.93 (0.27)	0.52-1.66	0.69 (0.25)	0.33-1.42	1.24 (0.47)	0.59-2.61
No memory problem	1.00	1.00	1.00	1.00	1.00	1.00
Memory problem	1.04 (0.22)	0.69-1.59	0.93 (0.33)	0.46-1.88	1.21 (0.32)	0.73-2.03
Self-ratings of health	1.16 (0.06)[b]	1.05-1.28	1.04 (0.08)	0.90-1.21	1.29 (0.09)[a]	1.12-1.49
Did not see/talk to a GP	1.00	1.00	1.00	1.00	1.00	1.00
Saw/talked to a GP	1.50 (0.20)[b]	1.15-1.97	1.53 (0.29)[c]	1.05-2.21	1.52 (0.30)[c]	1.03-2.24
Did not see/talk to a specialist	1.00	1.00	1.00	1.00	1.00	1.00

Predictor	All (N=5256)		Men (N=2059)		Women (N=3197)	
	Odds Ratio (SE)	95% CI	Odds Ratio (SE)	95% CI	Odds Ratio (SE)	95% CI
Saw/talked to a medical specialist	1.35 (0.15)[b]	1.09-1.67	1.21 (0.20)	0.87-1.68	1.48 (0.21)[b]	1.12-1.95
Did not see/talk to an eye doctor	1.00	1.00	1.00	1.00	1.00	1.00
Saw/talked to an eye doctor	1.27 (0.12)[b]	1.06-1.53	1.11 (0.15)	0.85-1.44	1.41 (0.19)[b]	1.08-1.82
Did not see/talk to a PT/OT	1.00	1.00	1.00	1.00	1.00	1.00
Saw/talked to a PT/OT	1.46 (0.18)[b]	1.14-1.87	1.20 (0.24)	0.81-1.78	1.76 (0.29)[a]	1.27-2.44
Did not see/talk to a chiropractor	1.00	1.00	1.00	1.00	1.00	1.00
Saw/talked to a chiropractor	1.02 (0.15)	0.77-1.36	0.93 (0.25)	0.55-1.57	1.10 (0.19)	0.78-1.55
Did not have hospitalization	1.00	1.00	1.00	1.00	1.00	1.00
Had overnight hospitalization	0.92 (0.16)	0.66-1.29	1.44 (0.35)	0.90-2.31	0.66 (0.15)	0.42-1.04
Did not have surgery	1.00	1.00	1.00	1.00	1.00	1.00
Had any surgery	1.21 (0.16)	0.94-1.57	1.07 (0.20)	0.74-1.55	1.29 (0.26)	0.87-1.91
Did not use home care services	1.00	1.00	1.00	1.00	1.00	1.00
Used home care services	0.76 (0.16)	0.51-1.14	0.66 (0.22)	0.35-1.27	0.85 (0.23)	0.50-1.44
Did not see/talk to a mental health professional	1.00	1.00	1.00	1.00	1.00	1.00
Saw/talked to a mental health professional	1.78 (0.43)[c]	1.10-2.88	1.93 (0.68)	0.97-3.87	1.70 (0.55)	0.90-3.22

[a] $P < .001$

[b] $P < .01$

[c] $P <.05$

Discussion

Despite many previous studies that found a significant age-related digital divide in HIT use, previous research did not extensively identify contributors and barriers to older adults' HIT use. Given that the older adult groups are the most frequent, heaviest users of health services of all age groups, the goal of the study was to examine the relationship between their HIT use and their health service use. The analysis of the US NHIS data confirmed the findings of previous studies that the rates of HIT use were significantly lower in persons in the age group 65 and over compared with persons in younger age groups, although persons in the age group 55 to 64 were not different from those in the younger age groups. The age group difference was also conspicuous among those aged 65 and over, with the rates of HIT use decreasing from 32.2% in the age group 65 to 74 to 14.5% in the age group 75 to 84 and 4.9% in the age group 85 and over. In addition to age, other demographic and socioeconomic variables (race/ethnicity and levels of education and family income) were significant determinants of HIT use among older adults.

Multivariate logistic regression results fully support hypothesis 1 (ie, a positive association between visits/talk to a general practitioner and HIT use) but only partially support hypothesis 2 (ie, a positive association between visits/talk to other health services and HIT use) and hypothesis 3 (ie, a positive association between visits/talk to a mental health professional and HIT use). Partial support for hypotheses 2 and 3 was attributable to the findings that having visited or talked to a chiropractor and having had overnight hospitalization, surgery, and/or homecare services were not associated with the odds of HIT use, while having visited or talked to a medical specialist, eye doctor, or PT/OT was significantly associated with women's HIT use only. Given these gender-specific patterns of association between health service use and HIT use, hypothesis 4 (ie, gender difference) was supported.

The findings imply that—controlling for demographic, socioeconomic, and health status—older adults with more general health care needs were more likely to use HIT than those with fewer general health care needs, as seeing/talking to a general practitioner was a significant correlate for both genders. On the other hand, the lack of association for both genders between HIT use and the use of overnight hospitalization, surgery, chiropractic care, and homecare appears to suggest that these more severe or specialized health care needs are not significantly associated with the odds of HIT use. The older adults who had undergone an overnight hospitalization and/or outpatient or inpatient surgery and/or had received homecare (usually following a hospitalization) were likely to have received health care information specific to their medical conditions from their health care providers, decreasing the need for online information seeking. Also, serious and/or multiple medical problems may have made it difficult for these older adults to use HIT. Some older adults have difficulty searching for complex problems online or understanding complex medical information [11,20].

Gender differences in HIT use are very interesting. The unadjusted rates of HIT use were lower among older women than older men; however, in line with the findings of previous studies that included all adult age groups [3,5,8,21], multivariate analysis results show that, when other things were equal, older women were more likely than their male counterparts to have used HIT. The finding that older women who had seen or talked to a medical specialist or eye doctor were also more likely to have used HIT than their peers who had not seen or talked to those health care providers suggests that older women may be more likely than their male counterparts to look for a wide range of online health information or engage in other health-related activities. An intriguing finding is the gender-specific association between mental health service use and HIT use. The finding that only older men who had seen or talked to a mental health professional were more likely than their peers who had not done so to have used HIT suggests two possibilities: (1) as found in previous studies [17,18], the stigma of having mental health problems may have influenced older men to a greater extent than older women to utilize HIT as a source of information or vehicle for other related activities; and (2) HIT use may have influenced older men to a greater extent than older women to use mental health services.

The study has a few limitations. First, since NHIS is a cross-sectional data set, the time order between HIT use and health service use could not be determined. As a result, only correlations, not causations, were deduced. Second, the NHIS HIT questions did not ask the respondents for whom they had searched health information and engaged in health-related activities on the Internet. Although most of those who used HIT are likely to have used it for both themselves and their loved ones, some older adults may have used it exclusively for others, such as their spouse and their relatives. Longitudinal and qualitative data will help examine the timing and the context of older adults’ use of HIT. Third, the NHIS data are self-reported, thus the reliability of some data on health service use and HIT use may be questionable, especially for those with some memory issues. Future research needs to find more objective measures to examine the relationship between health service use and HIT use.

Despite these limitations, this was one of the first studies to have examined the association between older adults’ use of HIT and their health service use. The findings have implications for narrowing the age-related and socioeconomic status–related gaps in HIT use. The access gaps among racial/ethnic minority older adults and poorly educated and/or low-income older adults are especially striking and call for concerted efforts to facilitate Internet access and HIT use among these disadvantaged older adults. Previous studies show that training classes and technical support may help (1) low-income persons use the Internet use, and (2) older adults in general who are willing to use the Internet as a general source of health information and email to communicate with their physicians [22-24]. However, the content of health information and the webpage design issues are also important considerations when attempting to accommodate low level of literacy/health literacy as well as the aging-related sensory and cognitive limitations among these older adults [8,10,11]. To increase HIT use among those with complex medical conditions, ease of comprehension also needs to be considered for all age groups [20]. Given the comparable rates of HIT use between the age group 55 to 64 and the younger age groups, the HIT use of future older adults is likely to increase. However, the digital divide between racial/ethnic minority, poorly educated, and/or low-income older adults and their non-Hispanic white, better-educated, and high-income counterparts will likely continue unless there are targeted efforts to reduce the access gaps.

Conflicts of Interest

None declared

References

1. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406 [FREE Full text] [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
2. Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, et al. Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). J Med Internet Res 2007;9(3):e20 [FREE Full text] [doi: 10.2196/jmir.9.3.e20] [Medline: 17627929]
3. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
4. Miller EA, West DM, Wasserman M. Health information Websites: characteristics of US users by race and ethnicity. J Telemed Telecare 2007;13(6):298-302. [doi: 10.1258/135763307781644915] [Medline: 17785031]
5. 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]
6. Birru M, Steinman RA. Online health information and low-literacy African Americans. J Med Internet Res 2004 Sep 3;6(3):e26 [FREE Full text] [doi: 10.2196/jmir.6.3.e26] [Medline: 15471752]
7. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]

8. Jensen JD, King AJ, Davis LA, Guntzviller LM. Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. J Aging Health 2010 Sep;22(6):804-826. [doi: 10.1177/0898264310366161] [Medline: 20495159]
9. Gracia E, Herrero J. Internet use and self-rated health among older people: a national survey. J Med Internet Res 2009;11(4):e49 [FREE Full text] [doi: 10.2196/jmir.1311] [Medline: 19955041]
10. Chaffin AJ, Maddux CD. Accessibility accommodations for older adults seeking e-health information. J Gerontol Nurs 2007 Mar;33(3):6-12. [Medline: 17378186]
11. Sharit J, Hernández MA, Czaja SJ, Pirolli P. Investigating the Roles of Knowledge and Cognitive Abilities in Older Adult Information Seeking on the Web. ACM Transactions on Computer-Human Interaction 2008;15(1):3. [doi: 10.1145/1352782.1352785] [Medline: 20011130]
12. National Center for Health Statistics. Health, United States, 2009: With Special Feature on Medical Technology URL: http://www.cdc.gov/nchs/data/hus/hus09.pdf [accessed 2011-04-06] [WebCite Cache ID 5xkQjD9j7]
13. 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]
14. Wagner TH, Baker LC, Bundorf K, Singer S. Preventing Chronic Disease Public Health Research, Practice, and Policy; 1(4). 2004. Use of the Internet for health information by the chronically ill URL: http://www.cdc.gov/pcd/issues/2004/oct/2004/04_0004.htm [accessed 2011-04-06] [WebCite Cache ID 5xkQoE8SE]
15. Dutta-Bergman MJ. Health attitudes, health cognitions, and health behaviors among Internet health information seekers: population-based survey. J Med Internet Res 2004 May 28;6(2):e15 [FREE Full text] [doi: 10.2196/jmir.6.2.e15] [Medline: 15249264]
16. LaCoursiere SP, Knobf MT, McCorkle R. Cancer patients' self-reported attitudes about the Internet. J Med Internet Res 2005 Jul 1;7(3):e22 [FREE Full text] [doi: 10.2196/jmir.7.3.e22] [Medline: 15998613]
17. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
18. Powell J, Clarke A. Internet information-seeking in mental health: population survey. Br J Psychiatry 2006 Sep;189:273-277 [FREE Full text] [doi: 10.1192/bjp.bp.105.017319] [Medline: 16946364]
19. National Center for Health Statistics. Atlanta, GA: Centers for Disease Control and Prevention; 2010. 2009 National Health Interview Survey (NHIS) Public Use Data Release: NHIS Survey Description URL: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2009/srvydesc [accessed 2011-04-06] [WebCite Cache ID 5xkgRlLCH]
20. Murero M, D'Ancona G, Karamanoukian H. Use of the Internet by patients before and after cardiac surgery: telephone survey. J Med Internet Res 2001;3(3):E27 [FREE Full text] [doi: 10.2196/jmir.3.3.e27] [Medline: 11720969]
21. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
22. Campbell RJ, Nolfi DA. Teaching elderly adults to use the Internet to access health care information: before-after study. J Med Internet Res 2005;7(2):e19 [FREE Full text] [doi: 10.2196/jmir.7.2.e19] [Medline: 15998610]
23. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. J Med Internet Res 2007;9(4):e31 [FREE Full text] [doi: 10.2196/jmir.9.4.e31] [Medline: 17951215]
24. Singh H, Fox SA, Petersen NJ, Shethia A, Street RL. Older patients' enthusiasm to use electronic mail to communicate with their physicians: cross-sectional survey. J Med Internet Res 2009;11(2):e18 [FREE Full text] [doi: 10.2196/jmir.1143] [Medline: 19632972]

Abbreviations

GED: general equivalency diploma
HIT: health information technology
NCHS: National Center for Health Statistics
NHIS: National Health Interview Survey
PT/OT: physical therapy/occupational therapy

Edited by G Eysenbach; submitted 09.02.11; peer-reviewed by C Maddux; comments to author 05.03.11; accepted 10.03.11; published 20.04.11.

Please cite as:
Choi N
Relationship Between Health Service Use and Health Information Technology Use Among Older Adults: Analysis of the US National Health Interview Survey
J Med Internet Res 2011;13(2):e33
URL: http://www.jmir.org/2011/2/e33/
doi:10.2196/jmir.1753
PMID:21752784

Original Paper

Internet Skills Performance Tests: Are People Ready for eHealth?

Alexander J A M van Deursen[1], PhD; Jan A G M van Dijk[1], PhD

University of Twente, Department of Media, Communication and Organization, Enschede, Netherlands

Corresponding Author:
Alexander J A M van Deursen, PhD
University of Twente
Department of Media, Communication and Organization
Cubicus C1.47
PO Box 217
Enschede, 7500 AE
Netherlands
Phone: 31 534891021
Fax: 31 534894259
Email: a.j.a.m.vandeursen@utwente.nl

Abstract

Background: Despite the amount of online health information, there are several barriers that limit the Internet's adoption as a source of health information. One of these barriers is highlighted in conceptualizations of the digital divide which include the differential possession of Internet skills, or "eHealth literacy". Most measures of Internet skills among populations at large use self-assessments. The research discussed here applies a multifaceted definition of Internet skills and uses actual performance tests.

Objective: The purpose of this study was to assess how ready a sample of the general population is for eHealth. More specifically, four types of Internet skills were measured in a performance test in which subjects had to complete health-related assignments on the Internet.

Methods: From November 1, 2009, through February 28, 2010, 88 subjects participated in the study. Subjects were randomly selected from a telephone directory. A selective quota sample was used divided over equal subsamples of gender, age, and education. Each subject had to accomplish assignments on the Internet. The Internet skills accounted for were categorized as operational (basic skills to use the Internet), formal (navigation and orientation), information (finding information), and strategic (using the information for personal benefits). The tests took approximately 1.5 hours and were conducted in a University office, making the setting equally new for all. Successful completion and time spent on the assignments—the two main outcomes—were directly measured by the test leader.

Results: The subjects successfully completed an average of 73% (5.8/8) of the operational Internet skill tasks and an average of 73% (2.9/4) of the formal Internet skill tasks. Of the information Internet skills tasks, an average of 50% (1.5/3) was completed successfully and, of the strategic Internet skills tasks, 35% (0.7/2). Only 28% (25/88) of the subjects were able to successfully complete all operational skills tasks, 39% (34/88) all formal skills tasks, 13% (11/88) all information skills tasks, and 20% (18/88) both the strategic skill tasks. The time spent on the assignments varied substantially. Age and education were the most important contributors to the operational and formal Internet skills. Regarding the formal Internet skills, years of Internet experience also had some influence. Educational level of attainment was the most important contributor to the information and strategic Internet skills.

Conclusions: Although the amount of online health-related information and services is consistently growing, it appears that the general population lacks the skills to keep up. Most problematic appear to be the lack of information and strategic Internet skills, which, in the context of health, are very important. The lack of these skills is also problematic for members of younger generations, who are often considered skilled Internet users. This primarily seems to account for the operational and formal Internet skills. The results of the study strongly call for policies to increase the level of Internet skills.

(J Med Internet Res 2011;13(2):e35) doi:10.2196/jmir.1581

KEYWORDS

Internet; eHealth; online; skills; literacy; information; digital divide

Introduction

The Internet nowadays functions as an important source of health information for the general population. The use of websites in the domain of health care varies from searching for information, getting support from Internet-based peer groups, getting online consultations, and obtaining health interventions [1]. The most commonly reported function is searching for health information [2]. People often search for health information on the Internet when prescribed a new medication or course of treatment, when dealing with a medical condition, when having unanswered questions after a doctor's visit, or when deciding to change diets or exercise habits [3]. Popular topics include fitness, drugs, hospitals, treatments, alternative medicines, and doctors [4-6]. Among the potential benefits of providing health information online are saving time and effort, easier access, getting help when feeling embarrassed or stigmatized, healthier lifestyles, early detection of potential medical problems, collaborative treatment of illnesses, and access to treatments that a local provider may not have access to [1,7].

To realize the Internet's potential for improving the health of the public, there are some important drawbacks that should be accounted for: (1) there are only a few tools available to help people find relevant information among the excessive amount of information that is available [8]; (2) the available information is often incomplete [8]; (3) the scientific quality of online health information is often questionable, and there are several examples of the supply of potentially harmful information on topics such as cancer rates, smoking cessation methods, and fever management in children [9-11]; and (4) users must be able to understand the information found and put it into practice [12]. Difficulties with technical terms and required reading skill levels appear to be problematic [8]. All these barriers that limit the Internet's widespread adoption as a source of health information relate to one of the most important factors in the current digital divide debate. While early research on the digital divide primarily focused on a binary classification of having or not having physical access to the Internet, a more refined understanding has appeared with the differential possession of Internet skills as key aspect [13-16].

Levels of Internet Skills

Deficient levels of Internet skills may prevent people from recognizing that information is missing, from understanding the difference between biased and unbiased information, from distinguishing evidence-based claims, and from interpreting the information intended for health professionals [17]. An indication of the low levels of search-related Internet skills is that one of the most common complaints about online health information searches is the amount of time required to process the documents that are found online [18]. While Internet search engines help to identify a large number of health-related documents, use of search engines calls for advanced skills that not all information consumers may possess [19]. In general, search engines and simple search terms do not seem to provide efficient access to health information [20]. Furthermore, most users seem to focus on finding information quickly rather than on evaluating the information found [21]. Most people only explore the first few links obtained from a search using a general search engine [22].

Besides search-related problems, more basic problems also limit the general public's use of online health information and services. Actual performance tests in the United States revealed that the general user population in the United States lacks an understanding of the basics of surfing the Internet [23]. In the Netherlands, the results of performance tests that have been conducted support the search-related problems described above [24]. Furthermore, these tests also revealed that a large part of the population struggles with deficiencies of other Internet skills [24]. Not everyone seems to be able to save or even open PDF documents. In addition, the basic operations needed to access the Internet such as typing a URL in the address bar or typing search queries in the search bar are not self-evident for all Internet users. Furthermore, many Internet users experience problems with website designs and when new browser windows are opened, something not unusual on the Internet.

Measuring Internet Skills

The literature concerning Internet skills is not consistent in the terms used or in the underlying concepts applied. Though there are many converging views, there is no agreement on the exact definition. The types of Internet skills that have received the most of the attention are basic technological knowledge and search behavior. In health-related literature, a popular concept is health information–seeking behavior [25]. Unfortunately, most studies provide little insight into this concept's specific meaning [25]. Contrary to the many related definitions of Internet skills, only a few measurable dimensions seem to have been applied within specific settings. One exception can be found in the field of health care. Here, the 8-item eHealth Literacy Scale (eHEALS) was developed to measure consumers' combined knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems [26]. The scale was based on a model that distinguishes between six types of literacy that are organized into analytic (traditional, media, information) and context-specific (computer, scientific, health) types. Although several skills required for the general Internet user population are considered, the main problems with this instrument are that the scale lacks validation and is based on an individual's perception of one's own skills or knowledge. This method is used in almost all strategies for measuring Internet skills for the obvious reason that observational studies are time-consuming and expensive. Unfortunately, measurements that use surveys and self-reports have significant problems of validity [27,28].

There are few measurements and scientific investigations of the actual levels of Internet skills possessed by populations at large [23,24,28]. We draw upon the definitions of Internet skills developed by Van Deursen and Van Dijk to encourage researchers to focus on in-depth skills measurements [24,28]. Van Deursen and Van Dijk elaborated and validated four types of Internet skills from an extensive literature overview. These are termed operational, formal, information, and strategic Internet skills.

Operational Internet skills indicate a set of basic skills in using an Internet browser, search engine, or Web-based form.

Formal Internet skills relate to the hypermedia structure, which requires the following skills: (1) navigating through different Web and menu layouts and (2) keeping a sense of orientation (disorientation is the most frequently cited problem in hypermedia use [29]).

Information Internet skills are derived from staged approaches in explaining the actions via which users try to fulfill their information needs [30]. Some of these actions are present in studies concerning health information–seeking behavior. However, in these studies, it is often difficult to locate explicit definitions. The actions considered here are: problem definition, choosing a way of searching, defining search queries, selecting relevant information, and evaluating the information found.

Strategic Internet skills enhance the capacity to use the Internet as a means of reaching particular goals including the general goal of improving one’s position in society. Strategic Internet skills are derived from the classical approach to decision making, where emphasis lies on procedures through which decision makers can reach optimal solutions as efficiently as possible [31].

By classifying the four Internet skills into medium-related Internet skills (operational and formal) and content-related Internet skills (information and strategic), Van Deursen and Van Dijk avoid a technological deterministic viewpoint since the command of hardware and software is not the only focus of attention. Both technical aspects related to the use of the Internet and substantial aspects related to the content provided by the Internet are accounted for. Furthermore, the four Internet skills have a sequential and conditional nature [28]. They illustrate that the possession of operational and formal skills alone is a necessary but not sufficient condition when using the Internet. Table 1 shows the specific skills included in each of the four Internet skill categories.

Table 1. Internet skills definitions [24,28]

Skill Category	Specific Skills
Medium-related Internet skills	
Operational skills in using Internet browsers	• Opening websites by entering the URL in the location bar • Navigating forward and backward between pages using the browser buttons • Saving files on the hard disk • Opening various common file formats (eg, PDFs) • Bookmarking websites • Changing the browser’s preferences • Using text or images with hyperlinks
Operational skills using Internet-based search engines	• Entering keywords in the proper field • Executing the search operation • Opening search results in the search result lists
Operational skills operating Internet-based forms	• Using the different types of fields and buttons Submitting a form
Formal skill used to navigate on the Internet	• Using hyperlinks embedded in different formats such as texts, images, or menus
Formal skill in maintaining a sense of location while navigating	• Not becoming disoriented when navigating within a website • Not becoming disoriented when navigating between websites • Not becoming disoriented when opening and browsing through search results
Content-related Internet skills	
Information skills used to locate required information	• Choosing a website or a search system to seek information • Defining search options or queries • Selecting information on websites or in search results • Evaluating information sources
Strategic skills used to take advantage of the Internet	• Developing an orientation toward a particular goal • Taking the right action to reach this goal • Making the right decision to reach this goal • Gaining the benefits resulting from this goal

The purpose of this study was to assess how ready the general population is for the transformation of information and services to the Internet in the domain of health care. The Internet skills of the Dutch population were measured in a performance test in which subjects had to complete health-related assignments on the Internet. The first research question was: What are the levels of Internet skills of Dutch citizens when using the Internet in the domain of health care?

When measuring Internet skills, the following variables should be accounted for [32]: Gender, age, educational attainment, time spent online, years of Internet experience, social resources, socioeconomic position, the location of Internet use, and participation in an Internet course. The second research question was: What are the characteristics of individuals who are most likely to suffer from inadequate levels of Internet skills in the domain of health care?

Methods

Recruitment

Subjects were recruited using random digit dialing in cities and villages in the region of Twente in the eastern region of the Netherlands. This region is fairly representative of the country as a whole because the demographics of the population as well as the proportion of the population living in rural versus urban settings are similar to those in the country as a whole. In line with procedures applied in prior research [24,28,32], a condition of participation was use of the Internet at least once every month for purposes other than email alone. Although this condition excluded approximately 20% of the Dutch population, it ensured that low frequency users who are nevertheless familiar with the Internet were also included. The invitation policy put people who feared a test at ease. Only Dutch-speaking adults 18 years of age and older were included and were promised €25 for their participation in a one-and-a-half hour research session.

To increase the representativeness of the findings, the subjects were recruited by applying a stratified random sampling method. First, a sample was randomly selected from a telephone directory. Subsequently, a selective quota sample was drawn from the strata to reach equal subsamples of gender, age (equal number of subjects in the categories of age 18-29, 30-39, 40-54, and 55-80), and educational level of attainment (equal number of subjects in the categories low, middle, and high). The result of this sampling procedure is that the results are not representative of the whole Dutch population. The focus is on the relative differences between the subsamples in terms of relative skill levels by age, gender, and education controlling for variables such as Internet experience. When respondents indicated they were willing to participate, their contact and email address were recorded and a time for the research session was scheduled. Respondents received a follow-up letter in the mail for confirmation and with directions to the research site. The day before the study, respondents were reminded of the session by phone.

Procedure

The performance tests were conducted from November 1, 2009, through February 28, 2010, in a university office. Prior to the test, a 10-minute questionnaire was administered to gather personal data. Subjects were asked for their year of birth, gender, educational level of attainment, amount of Internet use (hours per week), Internet experience (in years), location of respondents' regular Internet use, social support networks, and socioeconomic status.

In the performance tests, subjects used a keyboard, a mouse, and a 17-inch monitor connected to a laptop with a high-speed Internet connection. The laptop was programmed with the three most popular Internet browsers (Internet Explorer, Mozilla Firefox, and Google Chrome), which allowed the subjects to replicate their regular Internet use. No default page was set on the browsers, and all the assignments started with a blank page. To ensure that subjects were not influenced by a previous user's actions, the browser was reset after each session by removing temporary files, cookies, and favorites. In addition, downloaded files, history, forms, and passwords were removed and the laptop was rebooted.

During the assignment completion, subjects themselves decided when they were finished or wanted to give up on an assignment. No encouragements were given because the pressure to succeed was already higher in the laboratory setting than at home. After a specified ample amount of time had passed (determined for every task based on the results of 12 pilot tests), the test leader gently asked the subjects to move on to the next assignment (see Multimedia Appendix 1 for a complete overview of the assignments, the corresponding Internet skills, and the maximum time allowed). If the correct answer was not found, the assignment was rated as not completed. Both successful completion and time required—the main outcomes of the performance test—were directly noted during the sessions.

Assignments

The assignments the subjects had to complete were all health-related and accessible to the general user population. All assignments were fact-based and had a specific correct action or answer. Open-ended tasks were avoided because of the ambiguity of interpretation of the many potential answers. Included were two assignments (consisting of 8 tasks) to measure operational Internet skills, two (consisting of 4 tasks) to measure formal Internet skills, three to measure information Internet skills, and two to measure strategic Internet skills. In the operational assignments, subjects were, for example, asked to open a health website, save a file, or add a website to the “favorites.” Examples of tasks in the formal skills assignments were navigating different health-related menu and website designs and surfing between different websites. The information skill assignments charged subjects with finding health-related information on the Internet (requiring the skills described in Table 1 to locate information). Subjects were, for example, asked to find the name of a specific medical condition or to find out whether it would be a good idea to start with a treatment after being infected. Finally, the strategic skill assignments forced subjects to extract information from different sources, make decisions based on the information found, and gain personal benefits by making the right decisions. For example, subjects were asked to find out whether it would be a good idea to give a 3-year-old boy Vitamin A and D supplements, or to find a homecare organization with a specific caring program. All assignments were pilot tested with 12 subjects to ensure comprehensibility and applicability.

Results

Characteristics of the Sample Population

The characteristics of the subjects that participated in the performance test are shown in Table 2. The average number of years of Internet experience reported was 9.3 (SD 4.3) and the average amount of Internet use reported was 12.2 (SD 13.7) hours per week. Overall, the people who participated represented a diverse group of Internet users.

Table 2. Distribution of subjects by demographic variables (gender, education, and age) and the control variables (location of Internet use, needing assistance, socioeconomic status, and participation in an Internet course)

Characteristic	n (%)
Gender	
Male	45 (51%)
Female	43 (49%)
Age	
18-29	24 (27%)
30-39	18 (21%)
0-54	23 (26%)
55-80	23 (26%)
Level of education (highest completed)	
Low (primary school)	25 (28%)
Middle (high school)	32 (36%)
High (college or university)	31 (35%)
Primary location of Internet use	
At home	75 (85%)
At work	1 (1%)
At school	8 (9%)
At friends or family	3 (3%)
At a library	1 (1%)
Assistance when using the Internet	
No	49 (56%)
Yes, from family	18 (21%)
Yes, from friends	17 (20%)
Yes, from colleagues	4 (5%)
Yes, from a helpdesk	0 (0%)
Socioeconomic status	
Employee	30 (34%)
Retired	14 (16%)
Student	21 (24%)
Housemen/housewife	4 (4%)
Employer	6 (7%)
Disabled	4 (5%)
Unemployed	9 (10%)
Participation in an Internet course	
No	63 (72%)
Yes	25 (28%)

Task Completion and Time Spent

Table 3 shows that the subjects successfully completed an average of 5.8 (73%) of the 8 operational Internet skill tasks and an average of 2.9 (73%) of the 4 formal Internet skill tasks. Of the 3 information Internet skills tasks, an average of 1.5 (50%) were completed successfully, and of the 2 strategic Internet skills tasks, an average of 0.7 (35%) was completed successfully. The time spent on all assignments varies substantially.

Table 3. Overview of successful task completion and time spent

Internet Skills (Number of Tasks)	Average Task Completion		Seconds Spent	
	Mean (SD)	%	Mean (SD)	Minimum/Maximum time
Operational tasks (8)	5.8 (2.1)	73	427 (198)	118/980
Formal tasks (4)	2.9 (1.2)	73	450 (218)	180/1143
Information tasks (3)	1.5 (0.9)	50	960 (336)	343/1717
Strategic tasks (2)	0.7 (0.8)	35	1613 (545)	441/2500

Table 4. Number of tasks subjects failed to complete successfully

	Number of Failed Tasks	Number of Subjects (%)
Operational Internet skills	0	25 (28%)
	1	18 (21%)
	2	14 (16%)
	3	9 (10%)
	4	9 (10%)
	5	8 (9%)
	6	4 (4%)
	7	3 (3%)
	8	0 (0%)
Formal Internet skills	0	34 (39%)
	1	25 (28%)
	2	15 (17%)
	3	10 (11%)
	4	4 (5%)
Information Internet skills	0	11 (13%)
	1	35 (40%)
	2	30 (34%)
	3	12 (14%)
Strategic Internet skills	0	18 (20%)
	1	31 (35%)
	2	39 (44%)

Table 4 reveals that 28% of the subjects (25/88) were able to complete all operational Internet skills tasks, 39% (34/88) all formal Internet skills tasks, 13% (11/88), all information Internet skills tasks, and 20% (18/88), both of the strategic skills tasks. Furthermore, 44% of the subjects (39/88) could not complete either of the 2 strategic Internet skills tasks successfully. This was 13% (11/88 subjects) regarding the 3 information Internet skills tasks. The second strategic Internet skills tasks was the hardest and could only be completed successfully by 25% of the subjects (22/88). In this task, subjects were asked to find a homecare organization in the city of Enschede with a special caring program for individuals suffering from dementia and impaired hearing.

Contributors to the Level of Internet Skills

To identify factors that contribute to the level of Internet skills, two linear regressions for all 4 skills were conducted: one with the number of assignments completed and one with the time spent on these assignments as dependent variable. The independent variables in the regression model were gender, educational level attained (coded from 1, low to 3, high), age (years since birth), Internet experience (years online), amount of time spent on the Internet (hours per week), using social support (yes vs no), the primary location of Internet use (at home vs elsewhere), and socioeconomic status (active vs inactive).

Table 5 contains the linear regression results of the number of operational tasks completed successfully ($R^2 = .55$, $F_{9,87} = 10.11$, $P < .001$) and the time spent ($R^2 = .54$, $F_{9,87} = 10.11$ ($P < .001$). Age and education are the two significant contributors to the number of operational Internet skills tasks successfully completed and to the time spent on these tasks. Age is the strongest contributor. For seniors and people with lower levels of education, the most problematic task was saving PDF files.

Table 5. Linear regression results of the number of operational tasks completed successfully and the time spent

Independent Variables	Number of Tasks Completed		Time Spent	
	Beta	*P*	Beta	*P*
Gender (male/female)	−.10	.21	−.02	.65
Age (in years)	−71	< .001	.51	< .001
Education (low to high)	.18	.04	−.27	.01
Internet experience (in years)	.12	.17	−.15	.09
Time online (hours per week)	.02	.85	−.04	.64
Followed an Internet course (no/yes)	.02	.83	.10	.33
Using peers for help (no/yes)	.03	.72	.08	.32
Primary location of use (at home/elsewhere)	.01	.87	−.09	.28
Working situation (inactive/active)	−.11	.27	−.06	.59

Table 6 contains the linear regression results of the number of formal tasks completed successfully (R^2 = .61, $F_{9,87}$ = 13.60, $P < .001$) and the time spent (R^2 = .68, $F_{9,87}$ = 18.20, $P < .001$). Again age and education were significant contributors to both the number of successfully completed tasks and the time spent. In addition, number of years of Internet experience contributed significantly to both equations. Again seniors and people with lower levels of education experienced the most problems, especially with navigating different Web layouts.

Table 6. Linear regression results of the number of formal tasks completed successfully and the time spent

Independent Variables	Number of Tasks completed		Time Spent	
	Beta	*P*	Beta	*P*
Gender (male/female)	−.09	.27	.00	.36
Age (in years)	−.62	< .001	.61	< .001
Education (low to high)	.40	< .001	−.34	< .001
Internet experience (in years)	.17	.04	−.15	.03
Time online (hours per week)	−.11	.16	.00	.33
Followed an Internet course (no/yes)	−.05	.63	.07	.37
Using peers for help (no/yes)	−.06	.38	.03	.28
Primary location of use (at home/elsewhere)	.07	.32	−.04	.22
Working situation (inactive/active)	−.13	.18	−.02	.19

Table 7 contains the linear regression results of the number of information tasks completed successfully (R^2 = .34, $F_{9,87}$ = 4.48, $P < .001$) and the time spent (R^2 = .09, $F_{9,87}$ = 1.71, $P < .001$). Educational level of attainment is the strongest significant contributor to the number of information tasks completed successfully. Furthermore, participation in an Internet course positively contributes to the number of tasks completed successfully, and using peers for help negatively contributes significantly to the time spent on the tasks. The most difficult information Internet skills task was completed by only 28% (25/88) of the subjects. This task asked the subjects whether it is a good idea to start an antiviral (remedy against viral infections) for Lyme borreliosis. Subjects with lower levels of education, in particular, used broad search queries (eg, searching for the word *tick*) and did not venture past the first 3 search results. Also remarkably was that almost none of the subjects evaluated the information found.

Table 7. Linear regression results of the number of information tasks completed successfully and the time spent

Independent variables	Number of asks Completed		Time Spent	
	Beta	*P*	Beta	*P*
Gender (male/female)	.09	.26	−.13	.19
Age (in years)	−.06	.72	.08	.37
Education (low to high)	.56	< .001	−.08	.63
Internet experience (in years)	.01	.50	−.08	.50
Time online (hours per week)	−.01	.91	.00	.98
Followed an Internet course (no/yes)	−.24	.04	−.00	.77
Using peers for help (no/yes)	−.04	.46	.28	.01
Primary location of use (at home/elsewhere)	−.09	.50	−.05	.60
Working situation (inactive/active)	−.00	.71	−.04	.76

Table 8 contains the linear regression results of the number of strategic tasks completed successfully (R^2 = .45, $F_{9,87}$ = 7.10, *P* < .001) and the time spent (R^2 = .07, $F_{9,87}$ = 0.62, *P*= .21). Educational level of attainment is the only significant contributor to the number of successfully completed strategic Internet skills tasks. Hours spent online weekly and Using peers for help appeared to be significant contributors to the time spent on the tasks. The second strategic Internet skills tasks was the most difficult and could only be completed successfully by 25% of the most highly educated subjects (22/88). In this task, subjects were asked to find a homecare organization in the city of Enschede with a special caring program for people with dementia and impaired hearing.

Table 8. Linear regression results of the number of strategic tasks completed successfully and the time spent

Independent Variables	Number of Tasks Completed		Time Spent	
	Beta	*P*	Beta	*P*
Gender (male/female)	.11	.11	.11	.11
Age (in years)	.01	.55	.17	.14
Education (low to high)	.58	< .001	.11	.12
Internet experience (in years)	.07	.52	−.14	.15
Time online (hours per week)	.03	.63	.02	.54
Followed an Internet course (no/yes)	−.05	.89	−.06	.44
Using peers for help (no/yes)	−.00	.74	−.03	.46
Primary location of use (at home/elsewhere)	.02	.70	−.13	.22
Working situation (inactive/active)	.12	.22	.10	.20

Discussion

Principal Results

This study examined the level of Internet skills of a sample of the Dutch population when using the Internet for health-related information and services. Furthermore, it was examined whether skill levels can be predicted by demographic and socioeconomic factors. The study applied an in-depth definition of Internet skills by distinguishing between operational, formal, information, and strategic Internet skills. All four types of skills were measured in an actual performance test. While the test is not statistically representative for the general Dutch population, the results suggest that the sample on average possesses a sufficient level of operational and formal Internet skills when using the Internet for health-related topics. However, the levels of information skills and especially strategic Internet skills attained are probably much lower. Age and educational attainment are the most important contributing factors. Age appeared significant for the levels of operational and formal Internet skills, but not for the levels of information and strategic Internet skills.

Especially in the domain of health care, having sufficient levels of information and strategic skills is very important since the quality of the information offered is often questionable and unfortunately too often seems to be taken for granted. This might even mean that the lack of information and strategic Internet skills can become vital in the most literal sense. This happens when people with lower levels of Internet skills cannot find the hospital with the shortest waiting list for surgery or the best qualifications. This also occurs when they lack any other crucial information that helps them in preventing or relieving a particular urgent disease or when they are not able to ask for a second opinion about a proposed treatment. This is alarming, especially when considering that outside the artificial test situation created here, performance might be even lower (although we did not explicitly encourage the subjects).

It appears that the younger generations also are in need of improved information and strategic Internet skills. Younger generations are often considered to be skilled users of the Internet. In contrast to this, older people are often regarded as lagging behind in the adoption of new innovations. Together with the difficulties in learning new skills, resistance to change has also been suggested as a barrier to Internet use by the elderly [33]. However, in the performance tests reported here, it appears that younger generations do not score better on the information and strategic Internet skills. Similar results appeared in performance tests conducted in settings outside the domain of health [28,32]. This is an important finding, since there are also innumerable health sources on the Internet relevant for the younger generations. The results of the performance tests raise doubts as to whether younger people have sufficient information and strategic Internet skills to benefit from online health information provision. The same accounts for people with lower levels of education. Educational attainment appears very important in predicting who is likely to command low levels of Internet skills. Educational level of attainment proved significant for operational, formal, information, and strategic skills. Often, education is considered to be the most consistent global predictor for the use of information and communication technologies. More highly educated people are more likely to own computers, have Internet access at home, connect through broadband, and spend more time online [34]. Furthermore, more highly educated people are able to keep up with technological advancements and therefore increase their lead over people who are not able to keep up [35].

People who spend more time online—whether at work or any other location—are expected to acquire more knowledge about the Internet and are thus expected to have better online skills [27]. Moreover, people who have been Internet users for a longer period of time are presumed to be better at finding information online because they have more experience to draw on [27]. However, we found that operational, information, and strategic Internet skills do not grow with years of Internet experience and amount of time spent online weekly. Internet experience only had an effect on the formal Internet skills. Of all other variables, participation in an Internet course had some minor positive influence on the level of information Internet skills, and getting help from peers a negative influence on the same skills.

Limitations

This study provides an overview of the levels of four types of Internet skills among different segments of the Dutch population. However, only the absolute levels of the four types of Internet skills are considered. The specific aspects of the subjects' skills—for example the different steps concerning information or strategic Internet skills—are not considered here. A future qualitative analysis is required to provide more details about the specific skills indices.

A second limitation is that in this study, Internet use was limited to information retrieval. Communication skills were not measured because this would have made the performance tests that already required 1.5 hours from the subjects an unrealistic effort. Furthermore, content creation and sharing have also been ignored. These activities refer to so-called Web 2.0 applications. We consider information and strategic Internet skills as crucial for these activities, even more so than for information retrieval. Active participation and user-generated content require a high level of Internet skill, particularly for "serious" as compared to entertainment applications. Both limitations are a job for future researchers developing operational definitions and measurements of skills of communication, interaction, and peer-to-peer networking on the Internet.

Because of the major labor intensity of performance tests and the very high travel costs of drawing subjects to the university lab nationwide, it was not possible to test a random sample of 1200 people from the whole Dutch population. Ultimately, 88 subjects participated in the performance tests. Although this is not enough to generalize to the whole population, the applied quota sample for the categories of gender, age, and education greatly improved representativeness. Furthermore, to rate the overall representativeness of the sampling approach, the approach should be compared with the standards of an experiment rather than a survey. For an experiment, the number of subjects in our study was quite high. But larger than average experimental groups were required because large social and cultural differences in computer use and experience had to be taken into account.

Conclusions

In conclusion, this study found that operational and formal Internet skills are not sufficient when using the Internet for health purposes. The information and strategic Internet skills that are very important for seeking health information and for making decisions based on the retrieved information appear to be quite problematic. People with lower levels of education in all age groups seem to be the most likely candidates for lacking these Internet skills. The results of this study strongly call for policies that try to improve the level of Internet skills. The gap between the content provided and the content that people are able to manage must be acknowledged and remedied [26]. This is possible by accounting for Internet skills from a demand (user) and from a supply (Internet developers) perspective [36].

From a demand perspective, it appears that systematic training of operational and formal Internet skills in all types of adult education and computer classes would benefit older generations. If people in older groups improve these skills, they are likely to perform better on the information and strategic Internet skills than the younger generations. Training in information and strategic Internet skills should be accounted for in educational programs and other training environments. Unfortunately, learning to use the Internet is not a standard component of the current curriculum in education. On the contrary, it is generally believed that technologies such as the Internet by themselves empower learners and are regarded as an easy fix to learning itself [37]. This, however, is dubious and research on the matter is often misrepresented [37].

From a supply perspective, further research should be conducted into strategies to improve health information and service provision on the Internet. Results of this research indicate that seniors will probably benefit the most when websites require low levels of operational and formal Internet skills. After all, seniors do not seem to be inferior to younger citizens on either information or strategic skills. The lack of information and strategic Internet skills point to a major need for the improvement of the provision of online health-related information and services. Unfortunately the quality of online information varies substantially and the excessive amount of information offered online only makes the relevant sources harder to find for many users.

Conflicts of Interest

None declared

Multimedia Appendix 1

Assignments

[PDF file (Adobe PDF File), 86 KB - jmir_v13i2e35_app1.pdf]

References

1. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10. [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
2. Fox S, Rainie L. The online health care revolution: how the Web helps Americans take better care of themselves. Washington, DC: Pew Internet & American Life Project; 2000 Nov 26. URL: http://www.scribd.com/doc/219514/The-Online-Health-Care-Revolution-How-the-Web-helps-Americans-take-better-care-of-themselves [accessed 2010-05-20] [WebCite Cache ID 5ps8ntwXj]
3. Fox S, Rainie L. Vital decisions, a Pew Internet health report. Washington, DC: Pew Internet & American Life Project; 2002 May 22. URL: http://www.pewinternet.org/Reports/2002/Vital-Decisions-A-Pew-Internet-Health-Report.aspx [accessed 2010-05-20] [WebCite Cache ID 5ps92xhPm]
4. Fox, S. Health information online. Washington, DC: Pew Internet & American Life Project; 2005 May 17. URL: http://www.pewinternet.org/Reports/2005/Health-Information-Online.aspx [accessed 2010-05-20] [WebCite Cache ID 5ps9BGXCg]
5. Satterlund MJ, McCaul KD, Sandgren AK. Information gathering over time by breast cancer patients. J Med Internet Res 2003;5(3):e15 [FREE Full text] [doi: 10.2196/jmir.5.3.e15] [Medline: 14517106]
6. Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, et al. Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). J Med Internet Res 2007;9(3):e20 [FREE Full text] [doi: 10.2196/jmir.9.3.e20] [Medline: 17627929]
7. Mittman R, Cain M. The Future of the Internet in Health Care: Five Year Forecast.: California Healthcare Foundation; 1999 Jan. URL: http://www.chcf.org/publications/1999/01/the-future-of-the-internet-in-health-care-fiveyear-forecast [accessed 2010-05-20] [WebCite Cache ID 5ps9Pzbof]
8. Benigeri M, Pluye P. Shortcomings of health information on the Internet. Health Promot Int 2003 Dec;18(4):381-386 [FREE Full text] [Medline: 14695369]
9. Biermann JS, Golladay GJ, Greenfield ML, Baker LH. Evaluation of cancer information on the Internet. Cancer 1999 Aug 1;86(3):381-390 [FREE Full text] [Medline: 10430244]
10. Cheh JA, Ribisl KM, Wildemuth BM. An assessment of the quality and usability of smoking cessation information on the Internet. Health Promot Pract 2003 Jul;4(3):278-287. [Medline: 14610998]
11. Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. BMJ 1997 Jun;314(7098):1875-1879. [Medline: 9224132]
12. D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med 2001;155:807-812. [Medline: 11434848]
13. Warschauer M. Technology and social inclusion: rethinking the digital divide. Cambridge, MA: MIT Press; 2003.
14. van Dijk J. The Deepening Divide: Inequality in the Information Society. London: Sage Publications, Inc; 2005.
15. Mossberger K, Tolbert CJ, Stansbury M. Virtual Inequality: Beyond the Digital Divide. Washington, DC: Georgetown University Press; 2003.
16. DiMaggio P, Hargittai E. From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use As Penetration Increases. Working Paper Series 15: Princeton University Center for Arts and Cultural Policy Studies; 2001.
17. 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]
18. Hogan TP, Palmer CL. Information preferences and practices among people living with HIV/AIDS: results from a nationwide survey. J Med Libr Assoc 2005 Oct;93(4):431-439. [Medline: 16239938]

19. Ivanitskaya L, O'Boyle I, Casey AM. Health information literacy and competencies of information age students: results from the interactive online Research Readiness Self-Assessment (RRSA). J Med Internet Res 2006;8(2):e6 [FREE Full text] [doi: 10.2196/jmir.8.2.e6] [Medline: 16867969]
20. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 12020305]
21. Crespo J. Training the health information seeker: quality issues in health information web sites. Libr Trends 2004;53(2):360-374.
22. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
23. Hargittai E. Second-Level Digital Divide: Differences in People's Online Skills. First Monday 2002;7(4).
24. Van Deursen AJAM, Van Dijk JAGM. Using the Internet: Skill related problems in users' online behavior. International Journal of Human Computer Interaction 2009;21(5-6):393-402.
25. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]
26. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
27. Hargittai E. Survey measures of web-oriented digital literacy. Social Science Computer Review 2005;23(3):371-379.
28. Van Deursen AJAM, Van Dijk JAGM. Measuring Internet skills. International Journal of Human-Computer Interaction 2010;26(10):891-916.
29. Lee MJ. Expanding hypertext: Does it address disorientation? Journal of Computer-Mediated Communication 2005;10(3).
30. Marchionini G. Information Seeking in Electronic Environments. New York, NY: Cambridge University Press; 1995.
31. Miller K. Organizational Communication: Approaches and Processes, 4th edition. Belmont, CA: Wadsworth Cengage Learning; 2006.
32. Van Deursen AJAM. Internet Skills, Vital Assets in an Information Society. Enschede, the Netherlands: University of Twente; 2010.
33. Castells M. The Internet Galaxy: Reflections on the Internet, Business, and Society. Oxford, UK: Oxford University Press; 2002.
34. Buente W, Robbin A. Trends in Internet information behavior, 2000-2004. Journal of the American Society for Information Science and Technology 2008;59(11):1743-1760.
35. Goldin C, Katz LF. The Race between Education and Technology. Cambridge, MA: Harvard University Press; 2008.
36. Van Deursen AJAM, Van Dijk JAGM. Improving digital skills for the use of online public information and services. Government Information Quarterly 2009;26(2):333-340.
37. Buckingham D. Beyond Technology: Children's Learning in the Age of Digital Culture. Cambridge, UK: Polity; 2007.

Abbreviations

eHEALS: eHealth Literacy Scale

Edited by G Eysenbach; submitted 04.06.10; peer-reviewed by L Bowler; comments to author 13.11.10; revised version received 01.12.10; accepted 10.03.11; published 29.04.11.

Please cite as:
van Deursen AJAM, van Dijk JAGM
Internet Skills Performance Tests: Are People Ready for eHealth?
J Med Internet Res 2011;13(2):e35
URL: http://www.jmir.org/2011/2/e35/
doi:10.2196/jmir.1581
PMID:21531690

Original Paper

Impact of Health Portal Enrollment With Email Reminders on Adherence to Clinic Appointments: A Pilot Study

Monica Horvath[1], PhD; Janet Levy[2], PhD; Pete L'Engle[1], MPH; Boyd Carlson[1]; Asif Ahmad[1], MBA; Jeffrey Ferranti[1,3], MD MS

[1]Duke Health Technology Solutions, Duke University Health System, Durham, NC, United States
[2]Duke University School of Nursing, Durham, NC, United States
[3]Department of Pediatrics, Duke University School of Medicine, Durham, NC, United States

Corresponding Author:
Monica Horvath, PhD
Duke Health Technology Solutions
Duke University Health System
2424 Erwin Road
Durham, NC, 27705
United States
Phone: 1 919 668 6327
Fax: 1 919 668 3241
Email: monica.horvath@duke.edu

Abstract

Background: Internet portal technologies that provide access to portions of electronic health records have the potential to revolutionize patients' involvement in their care. However, relatively few descriptions of the demographic characteristics of portal enrollees or of the effects of portal technology on quality outcomes exist. This study examined data from patients who attended one of seven Duke Medicine clinics and who were offered the option of enrolling in and using the Duke Medicine HealthView portal (HVP). The HVP allows patients to manage details of their appointment scheduling and provides automated email appointment reminders in addition to the telephone and mail reminders that all patients receive.

Objective: Our objective was to test whether portal enrollment with an email reminder functionality is significantly related to decreases in rates of appointment "no-shows," which are known to impair clinic operational efficiency.

Methods: Appointment activity during a 1-year period was examined for all patients attending one of seven Duke Medicine clinics. Patients were categorized as portal enrollees or as nonusers either by their status at time of appointment or at the end of the 1-year period. Demographic characteristics and no-show rates among these groups were compared. A binomial logistic regression model was constructed to measure the adjusted impact of HVP enrollment on no-show rates, given confounding factors. To demonstrate the effect of HVP use over time, monthly no-show rates were calculated for patient appointment keeping and contrasted between preportal and postportal deployment periods.

Results: Across seven clinics, 58,942 patients, 15.7% (9239/58,942) of whom were portal enrollees, scheduled 198,199 appointments with an overall no-show rate of 9.9% (19,668/198,199). We found that HVP enrollees were significantly more likely to be female, white, and privately insured compared with nonusers. Bivariate no-show rate differences between portal enrollment groups varied widely according to patient- and appointment-level attributes. Large reductions in no-show rates were seen among historically disadvantaged groups: Medicaid holders (OR = 2.04 for nonuser/enrollee, 5.6% difference, $P < .001$), uninsured patients (OR = 2.60, 12.8% difference, $P < .001$), and black patients (OR = 2.13, 8.0% difference, $P < .001$). After fitting a binomial logistic regression model for the outcome of appointment arrival, the adjusted odds of arrival increased 39.0% for portal enrollees relative to nonusers (OR = 1.39, 95% CI 1.22 - 1.57, $P < .001$). Analysis of monthly no-show rates over 2 years demonstrated that patients who registered for portal access and received three reminders of upcoming appointments (email, phone, and mail) had a 2.0% no-show rate reduction ($P < .001$), whereas patients who did not enroll and only received traditional phone and mail reminders saw no such reduction ($P < .09$).

Conclusions: Monthly no-show rates across all seven Duke Medicine clinics were significantly reduced among patients who registered for portal use, suggesting that in combination with an email reminder feature, this technology may have an important and beneficial effect on clinic operations.

(J Med Internet Res 2011;13(2):e41) doi:10.2196/jmir.1702

KEYWORDS

Reminder systems; Health Information Technology for Economic and Clinical Health Act; medical informatics applications

Introduction

As the availability of the Internet continues to expand and health care consumers grow increasingly comfortable obtaining information online, surveys show that patients desire access to their personal health care information [1]. Patients have also reported altering their health care behavior based on information they find on the Internet [2]. Given these developments, health portals—novel Web-based applications that allow patients to securely and privately review portions of their electronic health record (EHR), schedule appointments, find educational information, review medications, and even send messages to providers—have the potential to revolutionize patients' involvement in their own care.

Recently, the Health Information Technology for Economic and Clinical Health **(**HITECH) act (July 2010) defined a detailed program through which Medicaid and Medicare providers could receive incentives for deploying and demonstrating "meaningful use" of certified EHRs [3]. This program may result in increased patient portal use, as one measure of meaningful use requires that more than 50% of patients requesting an electronic copy of their medical record receive access to that information within 3 business days. Portal applications afford an attractive means for meeting this objective, especially given the fact that patients have been shown to frequently use these tools to access aftercare summaries [4].

Reports of portal usage and enrollee demographics have been sporadically published over the past few years; most indicate patients primarily use portals for viewing lab and radiology results [4-6]. The majority of portals described in the literature are health system-dependent, meaning patients can only view their EHR as stored by a single health network. Examination of these studies' demographic characteristics shows that portal enrollees are more likely to be middle-aged, female, privately insured patients with a higher degree of morbidities [4,7]. These studies typically describe portal deployment and report characteristics of early adopters of the technology, including Kaiser Permanente, the Veterans Affairs health system, Group Health, and Beth Israel Deaconess Medical Center. Meanwhile, there is a notable paucity of follow-up reports describing whether enrollee populations broaden and how portal activities change over time. This absence calls for additional research to assess portal impact on patient outcomes, understand how patients use them, and define a business case for wider adoption [8-11].

Because portals provide patients with direct access to appointment scheduling details, their use may reduce missed appointments—a known barrier to clinic efficiency [12,13] that occurs in 10% to 30% of all appointments [14]. This study describes the demographic characteristics of patients enrolled in the Duke Medicine HealthView portal (HVP) and investigates how portal enrollment may influence appointment attendance at seven Duke Medicine clinics over a 1-year period. Specifically, we sought to test whether portal enrollment coupled with an email appointment reminder function is significantly related to decreases in rates of appointment "no-shows."

Methods

The Duke Medicine HealthView Portal (HVP)

The HVP (deployed in February 2007) is internally developed and supported by a full-time team from Duke Health Technology Solutions (DHTS), an entity within the Duke University Health System (DUHS) that is responsible for information technology initiatives and supports nearly 20,000 full-time health system employees. The HVP is a secure website constructed using IBM WebSphere Portal Server Architecture. The development team includes a manager, two business analysts, a technical support representative, two application system administrators, and three programmers. Additional staff from the DHTS Infrastructure and Operations teams provide ongoing hardware, networking, storage, and database support. The HVP is tested to ensure compatibility with Windows, Macintosh, and Linux operating systems running Firefox, Internet Explorer, and Safari browsers. HVP registration requires patients to provide an active email address and select a password. The HVP website is available outside of the DUHS firewall and, at the time of this study, affords access to appointment and billing information and clinical data pertaining to service at any outpatient Duke Medicine facility (Figure 1). Parents or legal guardians may link children to their own HVP accounts and act as surrogate users. By selecting Medical Records, patients are able to see their laboratory and radiology test results, which may be annotated by physicians. Other functions include viewing and scheduling appointments, reviewing accounts, editing personal and insurance profiles, selecting a preferred method for communication, and managing account settings such as password, email address, and required security questions. All HVP enrollees receive an email reminder 1 week prior to a scheduled appointment as well as an automated telephone call and a physical letter by mail. Nonusers receive only the telephone and mail reminders of upcoming appointments.

Figure 1. Screenshot of the DUHS HealthView portal

Welcome back to HealthView JANE

Managing account of: JANE DOE - 01/01/1970

Medical Records

for JANE DOE

From: 6/11/07

To: 6/11/08

Date	Description
04/09/2008	THYROID STIMULATING HORMONE
04/09/2008	BASIC METABOLIC PANEL
04/09/2008	AUTOMATED BLOOD COUNT
04/09/2008	LIPID PANEL (CALCULATED LDL)

Annotated Report

Your provider has written the following message for you about this report.

I have reviewed your results and everything is normal.

--Annotated By:PETER J. WOODS On: 2008-06-11

Patient: DOE, JANE AR6018

```
GENLAB Chemistry: Final   04/09/2008 09:50  Acc#  000810003746   Acct# 000HN6

THYROID STIMULATING HORMONE
                                                        Reference
  THYROID STIMULATING HORMONE        1.16        uIU/mL      [0.34-5.66]

  ORDERING MD:
   ROOKWOOD,JACQUELINE LOUI

  PERFORMED BY
   DUHS FRANKLIN LAB 4425 BEN FRANKLIN BLVD DURHAM, NC 27704
```

Setting and Patient Population

This study included all DUHS patients with a scheduled appointment in the period from January 1, 2008, through December 31, 2008, at one of seven study clinics located in Durham County, North Carolina: three primary care clinics, three children’s primary care clinics, and one specialty clinic. Clinics were selected based on overall appointment volume and HVP enrollee penetrance. The three primary care clinics were affiliated with Duke Primary Care (DPC), which offers full-service family medicine, general internal medicine, and pediatric medicine practices. Approximately 2600 patients attend 7800 appointments each month under the service of 29 physicians, 26 residents, 9 physician assistants, 2 nurse practitioners, and 2 pharmacists. The three Duke Children's Primary Care (DCPC) clinics provide comprehensive care for newborns, children, and adolescents. Approximately 1300 patients attend 4500 appointments each month under the service of 77 pediatric-focused physicians, 6 physician assistants, and 5 nurse practitioners. We also included one specialty clinic focused on hematology/oncology, as anecdotal evidence in the literature suggests that this group of patients is particularly interested in monitoring their clinical information online [15]. This clinic diagnoses, treats, and helps patients manage solid tumors, lymphomas, and a variety of complex bleeding and clotting disorders. The clinic practice comprises 9 physicians: 6 focused on hematology and 3 on oncology. Approximately 650 patients attend a total of 2400 appointments each month.

Study Design and Data Collection

We used a retrospective cross-sectional study to assess (1) demographic profiles of HVP enrollees and nonusers and (2) the relationship between HVP enrollment and appointment no-show rates where no-show rates are calculated as the percentage of all arrived and no-show appointments in total that were no-show appointments.

For direct demographic comparisons between groups, patients were classified as *enrollees* or *nonusers* according to HVP registration status as of December 31, 2008. Patient-level data extracted from the organizational data warehouse included age, patient-reported race, patient-reported ethnicity, sex, and HVP registration date. Patient age was classified into one of six groups and considered as categorical (not continuous) data, as this modeling better fit the observed variability of overall appointment scheduling by age.

For no-show rate comparisons, all 2008 appointment-level data (appointment date, location, visit type, and payor) for the seven study clinics were extracted and classified as *arrived* or *no-show*, indicating whether an appointment was kept, and as *enrollee appointment* or *nonuser appointment*, indicating whether the patient had a portal registration prior to the appointment date. “Payor” was defined as the payor identified by the patient at the time of the appointment scheduling and subsequently recorded in our data warehouse (the ultimate party who paid the appointment expenses could potentially differ from the payor identified by the patient). For a secondary analysis, we also collected all 2006 appointments from these seven clinics for the same cohort of patients in order to measure no-show rates among the same cohort prior to HVP deployment.

Statistical Methods

We tested for significant differences in demographic profiles between portal enrollees and nonusers using chi-square tests for categorical data and nonparametric Wilcoxon rank-sum tests for continuous data. *P* values less than or equal to .05 were considered statistically significant. For clinic appointments scheduled in 2008, no-show rates were compared between enrollees and nonusers and further stratified by patient- and appointment-level characteristics. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated as the odds of no-show among nonusers divided by the odds of no-show among HVP enrollees. All analyses were performed using JMP 8.0 (SAS Corporation, Cary, NC).

Multiple binomial logistic regression was used to describe the effect of HVP registration with an email appointment reminder feature on 2008 appointment arrivals (coded as 1 = arrived, 0 = no-show) while adjusting for other recognized confounders that could affect this outcome. All modeling was performed in SAS 9.1.3 (SAS Corporation, Cary, NC). An initial model was created using categorical stratification variables deemed relevant to appointment arrival based on ORs from the previous bivariate analysis: HVP registration, appointment clinic, race, age bracket, appointment type, payor, sex, and ethnicity. All variables, as well as their first-order interactions with HVP registration, were entered into the model with reference coding using the PROC LOGISTIC procedure. Each possible predictor variable had multiple levels. Levels within the same variable that had significant overlap in the CIs of the ORs for the no-show appointment outcome were collapsed. The backwards elimination method was used to remove nonsignificant predictor variables and interactions (Wald χ^2 test, $P > .05$). This defined the most parsimonious model describing appointment arrival. The -2 log likelihood ratio test was used to test the overall model significance, and the Wald χ^2 tests were used to assess the significance of the predictor variables.

Monthly no-show rates were calculated and plotted over time for clinic appointments from January to December 2006 (prior to HVP deployment) and then again from January to December 2008. Monthly no-show rates were subdivided by appointments for HVP enrollees versus nonusers. Only those patients who had at least one appointment in both 2006 and 2008 were included. Error bars were calculated as the 95% CI of each month’s no-show rate. Monthly no-show rates between compared groups were matched by month and the paired *t* test was used to assess statistical significance.

Results

Characteristics of HealthView Portal Enrollees

Across our seven study clinics, 58,942 patients scheduled a total of 198,119 appointments in 2008. A total of 13,265 patients failed to arrive for 19,668 appointments, resulting in an overall no-show rate of 9.9%. Of all 58,942 patients, 2838 (4.8%) were registered for HVP use at the beginning of the study, and an additional 9239 patients (15.7%) had registered by December 31, 2008. Consistent with previous reports [4-7,11], the 12,077 portal enrollees were significantly more likely to be female (63.8%, 7702/12,077), white (70.3%, 8495/12,077), to hold private health insurance (83.5%, 10,080/12,077), and to be between 40 and 65 years old (52.6%, 6,354/12,077) (Table 1, all $P < .001$). Portal enrollees in the 17 years of age and under age group comprised only 4.7% (563/12,077) of the enrollee population but 39.1% (18,345/46,865) of the nonuser population, likely because children usually have portal access through a parental surrogate. Both enrollees and nonusers had a median of two scheduled appointments per year, although nonusers scheduled more appointments. Nearly all comparisons of demographic factors between enrollees and nonusers were statistically significant with the exception of the proportion of individuals in the age group 18 to 29 years ($P = .75$) and the age group 65 years of age and older ($P = .88$).

Table 1. Demographic characteristics of patients seen at seven Duke Medicine clinics during 2008

	All Patients (n = 58,942) n (%)	HVP Enrollees[a] (n = 12,077) n (%)	Nonusers[a] (n = 46,865) n (%)	*P* [b]
Female	34,094 (57.8)	7702 (63.8)	26,392 (56.3)	< .001
Age group				
≥ 65	8569 (14.5)	1750 (14.5)	6819 (14.6)	.88
55-64	7441 (12.6)	2536 (21.0)	4905 (10.5)	< .001
40-54	11,619 (19.7)	3818 (31.6)	7801 (16.6)	< .001
30-39	6601 (11.2)	2211 (18.3)	4390 (9.4)	< .001
18-29	5804 (9.8)	1199 (9.9)	4605 (9.8)	.75
≤ 17	18,908 (32.1)	563 (4.7)	18,345 (39.1)	< .001
Ethnicity				
Hispanic/Latino	2263 (3.8)	147 (1.2)	2116 (4.5)	< .001
Other	56,679 (96.2)	11,930 (98.8)	44,749 (95.5)	< .001
Race				
White	29,621 (50.2)	8495 (70.3)	21,126 (45.1)	< .001
Black	22,182 (37.6)	2415 (20.0)	19,767 (42.2)	< .001
Asian	2135 (3.6)	557 (4.6)	1578 (3.4)	< .001
Other	5004 (8.5)	610 (5.1)	4394 (9.4)	< .001
Payor class [c]				
Medicaid	9888 (16.8)	199 (1.6)	9689 (20.7)	< .001
Medicare	8301 (14.1)	1539 (12.7)	6762 (14.4)	< .001
Private	39,243 (66.6)	10,080 (83.5)	29,163 (62.2)	< .001
Uninsured/self-pay	848 (1.4)	63 (0.5)	785 (1.7)	< .001
Unknown	662 (1.1)	196 (1.6)	466 (1.0)	< .001

[a] Enrollee status as of 12/31/2008

[b] *P* value by 2-tailed χ^2 test

[c] Payor as of last 2008 appointment

Analysis of Appointment No-Show Rates Across User Groups

Table 2 shows appointment no-show rates in both portal enrollees and nonusers during the study period. For an appointment to be classed as an enrollee appointment, the patient must have registered for an HVP account before the scheduled appointment date. In this way, patients who became enrollees over the course of the study could still have their appointment arrival activity appropriately categorized and analyzed. The unadjusted ORs describe the odds of a nonuser failing to keep a scheduled appointment relative to the odds of an enrollee failing to keep a scheduled appointment. Overall, nonusers had 2.26 times the odds of missing a scheduled appointment relative to enrollees receiving email reminders ($P < .001$). Although male nonusers had higher no-show rates than female nonusers, male enrollees had lower no-show rates than female enrollees, suggesting that HVP use may specifically improve appointment keeping in this subgroup. All age brackets had statistically significant no-show rates between enrollees and nonusers, but the greatest reduction in no-show rates was observed in the age group 18 to 29 years (OR = 2.35, rate difference of 8.11), whereas a more modest difference in no-show rates was observed among those in the age group 65 years and over (OR = 1.74).

Although all comparisons were statistically significant, uninsured/self-pay patients and Medicaid holders displayed the largest between-group difference in payor-stratified no-show rates (ORs 2.60 and 2.04, respectively), which represents a no-show rate difference between enrollee groups of 9.0% to 12.7%. This is a much greater difference than was seen between other payor types such as private insurance or Medicare (ORs 1.62 and 1.79, respectively), which represents a statistically different and yet much smaller no-show rate difference of 2.8% to 3.1% between enrollee groups. This pattern remained when race was examined, as black patients (a historically disadvantaged group) had an OR of 2.14 (no-show rate difference between groups of 8.0%, $P < .001$), whereas the OR calculation in white patients was 1.42 (no-show rate difference between groups of 1.6%, *P* < .001). Asians were the only group that exhibited a statistically similar no-show rate between enrollees and nonusers (4.1% and 5.5%, *P* = .09).

In terms of appointment-level characteristics, the hematology/oncology clinic had the lowest OR (1.58) for nonusers versus enrollees, a finding consistent with the fact that these patients are chronically ill and are thus more likely than the primary care groups to keep their appointments. When scheduled appointments were examined by type, consults (OR = 2.77) and new visits (OR = 2.53) had the largest no-show rate differences among visit classes. The day of week was also examined. Interestingly, there was an 8.8% difference in no-show rates between nonusers and enrollees on Saturdays (OR = 3.83), but only a 5.3% to 5.8% difference on other days (OR range 1.95 - 2.37). Taken together, these data suggest appointments outside of a patient’s normal personal schedule have a higher chance of arrival relative to other appointment types if registration for the HVP exists and an email reminder of upcoming appointments is sent.

Table 2. No-show rates by portal enrollee status at time of clinic appointment

Parameter	Nonuser No-Show Rate (%)	HPV Enrollee No-Show Rate (%)	*P* [a]	OR Nonuser/ Enrollee (95% CI)
All appointments	10.6 (18,423/173,204)	5.0 (1245/24,915)	< .001	2.26 (2.2-2.33)
Race				
Asian	5.5 (277/5013)	4.1 (37/900)	.09	1.36 (1.14-1.63)
Black	16.4 (12,325/75,141)	8.4 (418/4949)	< .001	2.13 (2.02-2.24)
White	5.7 (4468/78,317)	4.1 (745/18,209)	< .001	1.42 (1.36-1.48)
Other	9.2 (1353/14,733)	5.3 (45/857)	< .001	1.82 (1.56-2.13)
Age group				
≥ 65	5.9 (1718/29,182)	3.5 (169/4862)	< .001	1.74 (1.6-1.89)
55-64	6.8 (1360/19,964)	4.1 (236/5701)	< .001	1.69 (1.57-1.82)
40-54	8.6 (2391/27,829)	5.2 (378/7238)	< .001	1.71 (1.61-1.81)
30-39	10.6 (1526/14,372)	5.5 (196/3537)	< .001	2.02 (1.87-2.19)
18-29	15.2 (2311/15,203)	7.1 (150/2117)	< .001	2.35 (2.15-2.57)
≤ 17	13.7 (9117/66,654)	8.0 (116/1460)	< .001	1.84 (1.67-2.02)
Sex				
Male	10.8 (7988/73,781)	4.7 (436/9184)	< .001	2.44 (2.32-2.56)
Female	10.5 (10,435/99,423)	5.1 (809/15,731)	< .001	2.16 (2.08-2.25)
Ethnicity				
Hispanic/Latino	10.4 (946/9128)	6.7 (20/299)	.04	1.61 (1.28-2.04)
Other	10.7 (17,477/164,076)	5.0 (1225/24,616)	< .001	2.28 (2.21-2.35)
Day of week				
Monday	11.0 (4012/36,611)	5.2 (283/5441)	< .001	2.24 (2.11-2.39)
Tuesday	10.2 (3712/36,564)	4.9 (291/5988)	< .001	2.21 (2.08-2.35)
Wednesday	10.6 (3663/34,667)	4.8 (225/4705)	< .001	2.35 (2.19-2.52)
Thursday	10.1 (3299/32,747)	4.5 (214/4738)	< .001	2.37 (2.20-2.55)
Friday	11.4 (3454/30,169)	5.8 (228/3922)	< .001	2.09 (1.95-2.25)
Saturday	11.6 (283/2446)	3.3 (4/121)	.003	3.83 (2.29-6.39)
Insurance				
Medicaid	19.0 (7926/41,689)	10.3 (76/737)	< .001	2.04 (1.81-2.31)
Medicare	7.2 (2166/30,165)	4.1 (182/4402)	< .001	1.79 (1.66-1.94)
Private	7.7 (7440/96,687)	4.9 (937/19,150)	< .001	1.62 (1.56-1.68)
Uninsured/self-pay	23.2 (586/2525)	10.4 (17/163)	< .001	2.60 (2.00-3.37)
Unknown	14.3 (305/2138)	7.1 (33/463)	< .001	2.17 (1.79-2.62)
Visit type				
Return visits	10.3 (15,509/149,974)	4.6 (964/20,904)	< .001	2.39 (2.31-2.47)
Study visits	12.7 (1970/15,541)	7.5 (236/3133)	< .001	1.78 (1.66-1.91)
Consults	13.5 (343/2539)	5.3 (20/375)	< .001	2.77 (2.19-3.51)
New visits	11.7 (601/5150)	5.0 (25/503)	< .001	2.53 (2.05-3.12)
Clinic				
Duke Children's Primary Care, clinic 1	6.1 (878/14,349)	5.4 (33/615)	.49	1.15 (0.96-1.38)

Parameter	Nonuser No-Show Rate (%)	HPV Enrollee No-Show Rate (%)	P [a]	OR Nonuser/ Enrollee (95% CI)
Duke Children’s Primary Care, clinic 2	16.4 (7360/45,022)	10.4 (122/1169)	< .001	1.68 (1.52-1.85)
Duke Children’s Primary Care, clinic 3	16.4 (465/2839)	2.3 (1/43)	.01	8.23 (2.99-22.66)
Duke Primary Care, clinic 1	8.2 (976/11,953)	5.4 (353/6542)	< .001	1.56 (1.46-1.66)
Duke Primary Care, clinic 2	13.1 (5102/38,817)	5.4 (307/5666)	< .001	2.64 (2.49-2.81)
Duke Primary Care, clinic 3	4.5 (1525/34,031)	2.6 (130/5086)	< .001	1.79 (1.63-1.96)
Hematology/oncology clinic	8.2 (2158/26,193)	5.3 (312/5794)	< .001	1.58 (1.49-1.68)

[a] P value by 2-tailed χ^2 test

Assessment of Portal Impact After Logistic Regression Modeling

In order to better understand the effect of HVP registration on appointment arrivals, a logistic regression model was created to obtain ORs adjusted for confounders that may affect this relationship. Table 3 shows the maximum likelihood estimates for the parameters of which the most parsimonious model is composed. The overall model was considered significant compared with the null hypothesis, which states the parameters have slopes of zero (-2 log likelihood = 128,176.7, $P < .001$). This model retains a significant term for HVP enrollment ($P < .001$, OR = 1.39, 95% CI 1.22-1.57) after the addition of other parameters and backwards elimination. We obtained a better model by identifying each of the seven clinics uniquely in the model as opposed to grouping them into classes as in Table 2. Only the interaction terms between HVP enrollee status and clinic remained significant after backward elimination. From these data, we conclude the adjusted odds of appointment arrival are increased 39.0% for portal enrollees over nonusers. Compared with the unadjusted OR (2.26), the decrease in the HVP OR in the adjusted model (OR = 1.39) captures the confounding effects the other predictors have on the relationship between HVP enrollment and appointment arrival.

Table 3. Multivariable logistic regression model for the outcome of arrived appointment

Variable	Coefficient (Beta)	SE Beta	Wald χ^2	*P*	OR	95% CI
Intercept	2.97	0.07	2009.7	< .001		
HVP status						
Nonuser (reference)						
Enrollee	0.33	0.07	25.1	< .001	1.39	1.22-1.57
Clinic						
Duke Primary Care, clinic 1 (reference)						
Duke Primary Care, clinic 2	−0.12	0.04	9.2	.003	0.89	0.82-0.96
Duke Primary Care, clinic 3	0.80	0.04	328.7	< .001	2.22	2.04-2.42
Duke Children’s Primary Care, clinic 1	0.31	0.06	27.9	< .001	1.36	1.21-1.52
Duke Children’s Primary Care, clinic 2	−0.25	0.05	28.2	< .001	0.78	0.71-0.85
Duke Children’s Primary Care, clinic 3	−0.16	0.07	5.5	.02	0.85	0.74-0.97
Hematology/oncology clinic	0.07	0.04	3.1	.08	1.08	0.99-1.17
Sex						
Male (reference)						
Female	0.06	0.01	14.4	< .001	1.06	1.03-1.10
Race						
White (reference)						
Asian	0.09	0.06	2.3	.13	1.10	0.97-1.24
Black	−0.82	0.02	1798.2	< .001	0.44	0.42-0.46
Other	−0.21	0.05	23.1	< .001	0.80	0.73-0.88
Age						
40-64 (reference)						
≤ 17	0.01	0.04	0.3	.60	1.02	0.95-1.09
18-29	−0.35	0.03	131.2	< .001	0.70	0.66-0.75
30-39	−0.25	0.03	65.8	<.001	0.78	0.73-0.83
≥ 65	0.43	0.04	133.5	< .001	1.54	1.43-1.66
Appointment type						
Return visits (reference)						
Consults	0.09	0.06	2.4	.12	1.10	0.98-1.23
New visits	−0.18	0.05	16.0	< .001	0.83	0.76-0.91
Clinical trial visits	−0.51	0.03	389.7	< .001	0.60	0.57-0.63
Ethnicity						
Hispanic (reference)						
Other	−0.18	0.05	11.5	< .001	0.84	0.75-0.93
Payor						
Private insurance (reference)						
Medicaid/indigent	−0.53	0.02	632.6	< .001	0.59	0.56-0.61
Medicare	−0.22	0.03	40.8	< .001	0.80	0.75-0.86
Uninsured	−0.93	0.05	354.9	< .001	0.39	0.36-0.43
Unknown	−0.28	0.06	20.5	< .001	0.76	0.67-0.85
Interaction with HVP enrollee (Duke Primary Care, clinic 1*HVP enrollee)						

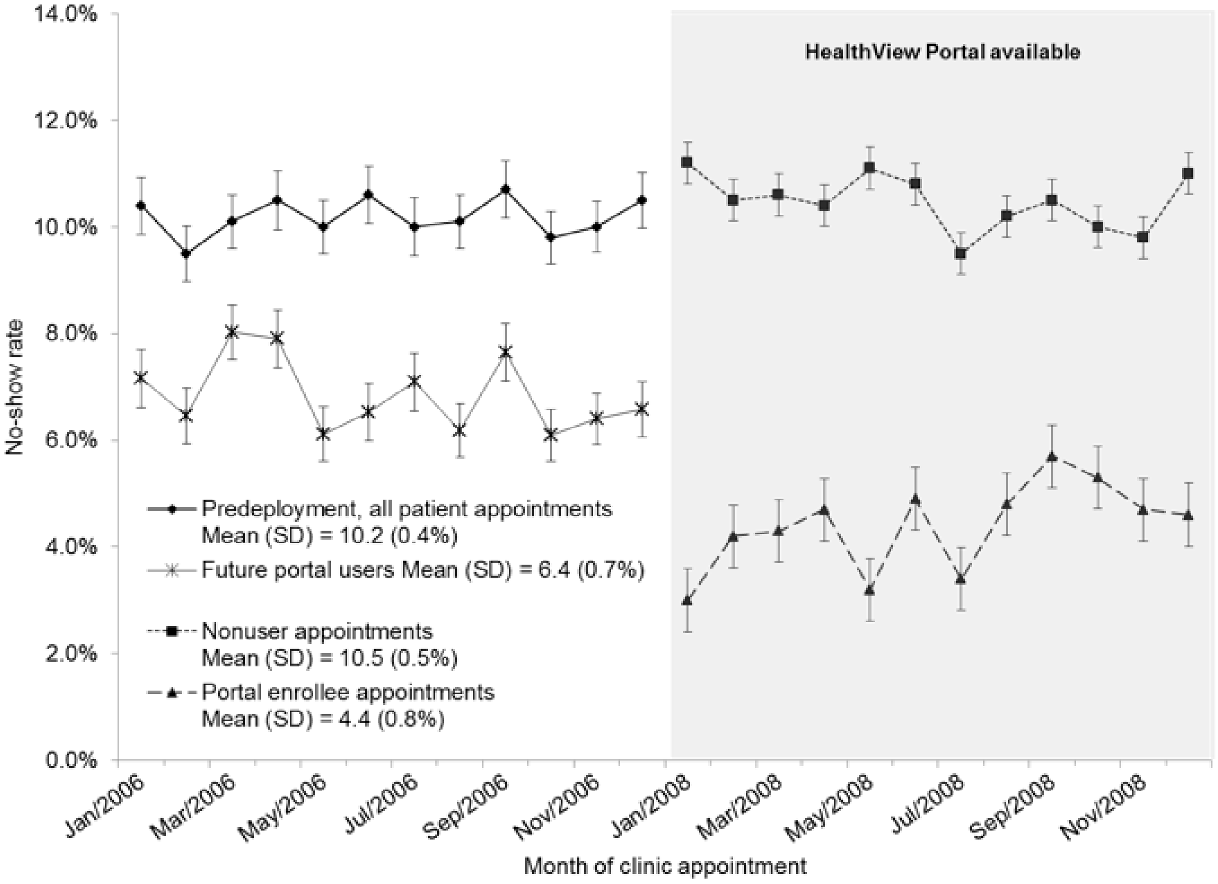
Variable	Coefficient (Beta)	SE Beta	Wald χ^2	*P*	OR	95% CI
Duke Primary Care - clinic 2	−0.21	0.20	1.2	.27	0.81	0.55-1.18
Duke Primary Care - clinic 3	−0.11	0.11	0.8	.37	0.90	0.71-1.13
Duke Children’s Primary Care - clinic 1	1.40	1.02	1.9	.17	4.03	0.55-29.70
Duke Children’s Primary Care - clinic 2	0.31	0.09	11.7	< .001	1.36	1.14-1.62
Duke Children’s Primary Care- clinic 3	0.13	0.11	1.3	.26	1.14	0.91-1.42
Hematology/oncology clinic	−0.05	0.09	0.3	.61	0.95	0.80-1.14

Analysis of No-Show Rates Over Time

In order to understand how patient no-show rates changed over time in this study, we examined the monthly trend in no-show rates prior to HVP deployment. Shown in Figure 2 are monthly appointment no-show rates over time for the seven DUHS clinics. All patients that had at least 1 appointment in both 2006 (pre-HVP deployment) and 2008 (HVP fully deployed as of March 2008) were included in the graph. Of the original 58,942 patients, 37,408 had 158,420 scheduled appointments at the seven study clinics during 2006. Of these, 35.8% (13,407/37,408) had 1 or more missed appointments in either 2006 or 2008. In the predeployment period, patients had an average monthly no-show rate of 10.2 (SD 0.4%) (Figure 2).

In the postdeployment period (2008), appointments were broken out by HVP registration status. The 2008 nonuser group had a mean monthly no-show rate of 10.5 (SD 0.5%), which was statistically indistinguishable from the mean monthly no-show rate prior to HVP deployment (*P*= .09). Relative to the 2006 predeployment phase, no-show rates among the HVP enrollee group fell to a monthly mean of 4.4 (SD 0.8%) (*P*< .001). These results show that the HVP enrollee group uniquely improved its appointment attendance rate. The mean no-show rate for patients in 2006 who would become enrollees by the end of the study was 6.4 (SD 0.7%), 2.0% greater than the no-show rate for appointments scheduled by these same patients and clinics in 2008 (*P*< .001).

Figure 2. Monthly appointment no-show rates over time for seven DUHS clinics



Discussion

Principal Results and Comparison With Prior Work

In this study, we developed a demographic profile of portal users at Duke Medicine clinics and found that monthly clinic no-show rates were reduced considerably among those patients having registered for portal usage, as they would have received email reminders of upcoming appointments. Portal enrollment was a significant predictor of appointment arrival even after adjustment for confounding factors such as patient demographics.

Although it is well-established that missed appointments can be costly to providers and decrease operational efficiency, there is little information in the literature regarding the effect of health information technology (HIT) applications on scheduled clinic appointment arrivals. Most reports focus on practices outside the United States—a context in which no-show rates vary considerably and the use of reminder methods such as text messaging have been shown to reduce no-show rates by amounts ranging from 0.8% to 14.5% [16-19]. Mailed reminders have been credited with reducing no-show rates by 10% [14]; telephone reminders have been shown to reduce rates by proportions that ranged from 5.8% to as much as 9.5% [20]. In our study, we saw a more modest but still highly significant 5.6% difference in no-show rates, as stratified by portal enrollment. Portal enrollment remains a significant independent predictor of appointment arrival even when adjusted by potential confounders in a multivariable regression model (OR = 1.39). Because all patients in our study received both telephone and mailed reminders, the improved arrival rate attributable to this HIT application may represent the next needed level of patient involvement beyond simpler reminder methods.

In our study, portal enrollees were typically white, female, middle-aged, and held private health insurance, findings consistent with demographic profiles reported by groups developing the PatientSite [6] and MyGroupHealth [4] portals. There is great concern that the overall impact of portal technologies will not transcend socioeconomic lines and age barriers and access will thus be restricted to younger, healthier, and wealthier persons [21]. We, too, see evidence of this “digital divide” based on socioeconomic status, given that Medicaid patients comprise only 1.7% of portal registrants.

Interestingly, the bivariate analysis of no-show rates demonstrates that the greatest improvement in appointment keeping is seen among traditionally disadvantaged populations who do use the portal. The no-show rate of Medicaid recipients and uninsured/self-pay patients fell by 8.7% and 12.8% respectively for portal enrollees relative to nonusers (each $P < .001$) whereas scheduled appointments belonging to Medicare and privately insured patients showed a 2.8% to 3.1% reduction in the no-show rate. Similarly, black patients who registered for the Duke Medicine HVP showed an 8.0% reduction in the rate of no-shows ($P < .001$) compared with a much more modest reduction among white patients (1.6%, $P < .001$).

Taken together, these data suggest that efforts to enroll patients in HVPs should focus on traditionally disadvantaged populations in order to achieve the greatest gains in appointment keeping. A recent report evaluated several cases of personal health record implementation and concluded that while disadvantaged populations tend not to be early adopters of such technology, patients with limited resources do use online health materials [22], a finding that lends further support for our approach. In fact, a survey of more than 17,000 Medicaid beneficiaries from Durham County, NC, found that 52% of beneficiaries had high-speed Internet access and 64.5% of beneficiaries would view health information through a portal at least once a year if given the option [23]. Although the digital divide also describes barriers for seniors, we found that 20.4% of persons aged 65 years of age or older in our study registered for the HVP by the end of 2008—a proportion that accounts for 14.5% of all enrollees and does not differ statistically from the proportion of seniors in the nonuser group ($P = .89$). However, the PatientSite study, using data gathered in 2004, noted that seniors made up only 7% of enrollees. Increasing access to technology over time may help account for these disparate findings.

With the exception of minors (who most often have portal accounts through parental surrogates), the no-show ORs contrasting nonusers with portal enrollees decreased as age increased to 65 years and older. Although more research is needed, we might infer that older persons are more independently active in managing their own health care and thus may benefit less from a portal in terms of its effect on keeping appointments. Similarly, young people may be healthier and possibly less responsible, meaning that portal-generated reminders are proportionately more helpful for keeping appointments.

Individual clinic was an important independent predictor of appointment arrival in the multivariable model; further, the no-show rate differences between enrollee groups varied by clinic class. Only DPC clinic 2 had statistically significant interaction with portal enrollment, indicating portal enrollment is more influential on no-show rates for this clinic relative to the other clinics. The reason for this difference is not clear and may be due to the unique environment of each clinic, differential responses among patient groups to portal marketing efforts, or the technology itself. Compared with the primary care clinics, the hematology/oncology specialty clinic saw the most modest no-show rate difference across enrollee groups (2.9%) despite having the same proportion of portal enrollees as the adult primary care clinics (17% of all appointments were scheduled by HVP enrollees in both groups). Patients with serious long-term illness have more pressing ongoing care needs; it is thus understandable that they have lower no-show rates. The impact of portal technologies across different types of patients may need to be evaluated in terms of different quality-of-care metrics. For example, a recent randomized controlled trial showed that diabetic patients who used an EHR-integrated portal had their treatment regimens adjusted more often than nonusers [24].

Our analysis of monthly no-show rates following HVP deployment illustrates that portal enrollees consistently demonstrate a 6.1% lower rate of appointment no-shows compared with nonusers. However, it is possible that HVP enrollment did not improve patient attendance so much as identify the most compliant and engaged set of patients. Our adjusted odds of appointment arrival for enrollees versus nonusers is 1.39, which is similar to other published regression models evaluating the effect of reminder methods on this outcome (eg, Parikh et al reported adjusted ORs of 1.58 to 1.98 for telephone calls relative to a control group receiving no reminders [20]). Moreover, when no-show rates from the period before deployment in 2006 were examined, rates for those who would later enroll were 2.0% higher across the same set of clinics, indicating the pattern of appointment attendance improved even among patients who were potentially more compliant.

Limitations

This study has several limitations. Our findings are drawn from only seven clinics within a large health system and describe a period of time when the HVP was a relatively new addition to patient care and thus may not be applicable to all health care settings. Due to the structure of the log files that described patient use of the portal, we were able to capture patient registration dates in order to identify enrollees but were unable to capture enrollee attrition. Ideally, detailed information regarding log-ins and page views would better define the profile of active enrollees and make a stronger case that active HVP use—not just registration and receipt of email appointment reminders—is an important component of patient arrivals. A project is underway to extract information on HVP use from application log files and load those data originating from July of 2009 forward into the organizational data warehouse. At the time of this writing, the HVP currently has more than 100,000 registered enrollees. We expect a future study analyzing enrollee demographics across all DUHS clinics will allow us to better estimate the effect of the technology on patient care management and allow development of a more rigorous predictive model that explains not only appointment arrivals, but also patient care outcomes such as emergency department utilization or long-term disease management. In this future work, we hope to identify patterns of clinic characteristics strongly associated with appointment keeping and portal enrollment, as was seen in this study with DPC clinic 2.

We did not collect data on advance cancellations. Hagerman and colleagues reported that patient reminders might increase the likelihood of cancellations [25], although this trend has not been observed elsewhere [14]. In our study, all patients were notified by telephone and mail. Only HVP enrollees received an additional reminder via email. However, even if an increase in advance cancellations was present in the HVP enrollee arm, this still represents a net benefit to study clinics, because staffing could be adjusted as needed and other patients could be accommodated on short notice. This question will be a focus of future analysis across a wider set of clinics.

Finally, we did not collect data on the projected costs of a no-show appointment. Given that our study was conducted among high-volume clinics served by a large number of providers, the wide variability in providers makes it difficult to determine the specific amount of revenue lost due to no-show appointments. Data on the cost of appointment no-shows are limited, and actual costs are likely to be highly clinic-dependent; thus, the issue of financial impact merits its own separate study. But given the effect of missed appointments on clinic operating costs and efficiency, such a study would help inform the decision of whether Duke Medicine should consider creating incentives for patients to use portal technologies.

Conclusions

This study developed a demographic profile for enrollees of the Duke Medicine HVP and described the relationship between portal registration and appointment arrival rates when email reminders are in place. As seen in other studies, historically disadvantaged groups are less likely to use the portal, providing further evidence of a digital divide. Monthly no-show rates across seven DUHS clinics were reduced considerably among those patients who registered for portal usage and received email reminders, suggesting this technology may have important beneficial effects on clinic operations, as nonusers would have only received mail and telephone reminders. Portal enrollment was a significant predictor of appointment arrival even when adjusted for known confounders such as payor, clinic, and race. The greatest proportional improvements in appointment attendance associated with portal enrollment were seen among historically disadvantaged groups. Further research that would examine actual log-in activity is needed to understand the relationship between different types of portal usage and more nuanced quality and safety outcomes in order to better elucidate the overall profile of portal effectiveness.

Acknowledgments

The authors thank the HealthView portal development team and the DHTS Data Warehouse Group for providing extracts, and Andrea Long and Heidi Cozart for critical review of this manuscript. The authors also wish to thank Morgan deBlecourt and Jonathan McCall for editorial assistance with this manuscript.

Conflicts of Interest

None declared

References

1. The Markle Foundation Connecting for Health. Americans want benefits of personal health records. URL: http://www.connectingforhealth.org/resources/phwg_survey_6.5.03.pdf [accessed 2010-11-22] [WebCite Cache ID 5uSbSgtnn]

2. Fox S. Online Health Search 2006. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. URL: http://www.pewinternet.org/Reports/2006/Online-Health-Search-2006.aspx [accessed 2010-11-21] [WebCite Cache ID 5uSbdfUcC]
3. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug 5;363(6):501-504 [FREE Full text] [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
4. Ralston JD, Carrell D, Reid R, Anderson M, Moran M, Hereford J. Patient web services integrated with a shared medical record: patient use and satisfaction. J Am Med Inform Assoc 2007;14(6):798-806. [doi: 10.1197/jamia.M2302] [Medline: 17712090]
5. Cimino JJ, Patel VL, Kushniruk AW. The patient clinical information system (PatCIS): technical solutions for and experience with giving patients access to their electronic medical records. Int J Med Inform 2002 Dec 18;68(1-3):113-127. [Medline: 12467796]
6. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95. [doi: 10.1197/jamia.M1833] [Medline: 16221943]
7. Carrell D, Ralston JD. Variation in adoption rates of a patient web portal with a shared medical record by age, gender, and morbidity level. AMIA Annu Symp Proc 2006:871. [Medline: 17238491]
8. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736. [doi: 10.1197/jamia.M2547] [Medline: 18756002]
9. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126. [doi: 10.1197/jamia.M2025] [Medline: 16357345]
10. Kilbridge PM, Classen DC. The informatics opportunities at the intersection of patient safety and clinical informatics. J Am Med Inform Assoc 2008;15(4):397-407. [doi: 10.1197/jamia.M2735] [Medline: 18436896]
11. Ralston JD, Rutter CM, Carrell D, Hecht J, Rubanowice D, Simon GE. Patient use of secure electronic messaging within a shared medical record: a cross-sectional study. J Gen Intern Med 2009 Mar;24(3):349-355. [doi: 10.1007/s11606-008-0899-z] [Medline: 19137379]
12. No authors listed. New statistics on outpatients' care confirm millions of missed appointments. New J (Inst Health Rec Inf Manag) 2006 Aug;47(4):17. [Medline: 17552489]
13. Capko J. The price you pay for missed appointments. J Med Pract Manage 2007;22(6):368. [Medline: 17612315]
14. Bigby J, Giblin J, Pappius EM, Goldman L. Appointment reminders to reduce no-show rates. A stratified analysis of their cost-effectiveness. JAMA 1983 Oct 7;250(13):1742-1745. [Medline: 6411944]
15. Wiljer D, Bogomilsky S, Catton P, Murray C, Stewart J, Minden M. Getting results for hematology patients through access to the electronic health record. Can Oncol Nurs J 2006;16(3):154-164. [Medline: 17523576]
16. da Costa TM, Salomão PL, Martha AS, Pisa IT, Sigulem D. The impact of short message service text messages sent as appointment reminders to patients' cell phones at outpatient clinics in São Paulo, Brazil. Int J Med Inform 2010 Jan;79(1):65-70. [doi: 10.1016/j.ijmedinf.2009.09.001] [Medline: 19783204]
17. Price H, Waters AM, Mighty D, Nixon J, Burton N, Picket J, et al. Texting appointment reminders reduces 'Did not Attend' rates, is popular with patients and is cost-effective. Int J STD AIDS 2009 Feb;20(2):142-143. [doi: 10.1258/ijsa.2008.008448] [Medline: 19182066]
18. Fairhurst K, Sheikh A. Texting appointment reminders to repeated non-attenders in primary care: randomised controlled study. Qual Saf Health Care 2008 Oct;17(5):373-376. [doi: 10.1136/qshc.2006.020768] [Medline: 18842978]
19. Downer SR, Meara JG, Da Costa AC, Sethuraman K. SMS text messaging improves outpatient attendance. Aust Health Rev 2006 Aug;30(3):389-396. [Medline: 16879098]
20. Parikh A, Gupta K, Wilson AC, Fields K, Cosgrove NM, Kostis JB. The effectiveness of outpatient appointment reminder systems in reducing no-show rates. Am J Med 2010 Jun;123(6):542-548. [doi: 10.1016/j.amjmed.2009.11.022] [Medline: 20569761]
21. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. Am J Prev Med 2007 May;32(5 Suppl):S131-S133. [doi: 10.1016/j.amepre.2007.02.012] [Medline: 17466818]
22. Kanaan SB. Safety-net providers bring patients online: Lessons from early adopters. Oakland, CA: California HealthCare Foundation; 2009 Apr. URL: http://www.chcf.org/~/media/Files/PDF/S/PDF%20SafetyNetPatientsOnline.pdf [accessed 2011-05-17] [WebCite Cache ID 5ymGworV2]
23. Willis JM, Macri JM, Simo J, Anstrom KJ, Lobach DF. Perceptions about use of a patient Internet portal among Medicaid beneficiaries. AMIA Annu Symp Proc 2006:1145. [Medline: 17238764]
24. Grant RW, Wald JS, Schnipper JL, Gandhi TK, Poon EG, Orav EJ, et al. Practice-linked online personal health records for type 2 diabetes mellitus: a randomized controlled trial. Arch Intern Med 2008 Sep 8;168(16):1776-1782 [FREE Full text] [doi: 10.1001/archinte.168.16.1776] [Medline: 18779465]
25. Hagerman GA. Testing the mailed appointment reminder in family practice. J Fam Pract 1978 Jul;7(1):199-201. [Medline: 670914]

Abbreviations

CI: confidence interval
DCPC: Duke Children's Primary Care
DHTS: Duke Health Technology Solutions
DPC: Duke Primary Care
DUHS: Duke University Health System
EHR: electronic health record
HIT: health information technology
HITECH: Health Information Technology for Economic and Clinical Health
HVP: HealthView portal
OR: odds ratio

Edited by G Eysenbach; submitted 01.12.10; peer-reviewed by J Cimino; comments to author 28.01.11; revised version received 16.02.11; accepted 19.03.11; published 26.05.11.

Please cite as:
Horvath M, Levy J, L'Engle P, Carlson B, Ahmad A, Ferranti J
Impact of Health Portal Enrollment With Email Reminders on Adherence to Clinic Appointments: A Pilot Study
J Med Internet Res 2011;13(2):e41
URL: http://www.jmir.org/2011/2/e41/
doi:10.2196/jmir.1702
PMID:21616784

Original Paper

Older Adults with Multi-Morbidity: Medication Management Processes and Design Implications for Personal Health Applications

Leah M Haverhals[1], MA; Courtney A Lee[2], MA; Katie A Siek[3], PhD; Carol A Darr[4], PhD; Sunny A Linnebur[5], PharmD; J Mark Ruscin[6], PharmD; Stephen E Ross[7], MD

[1]Colorado REAP (Research Enhancement Award Program) to Improve Care Coordination, Veterans Affairs Medical Center, Denver, CO, United States
[2]Department of Health and Behavioral Sciences, University of Colorado Denver, Denver, CO, United States
[3]Wellness Innovation and Interaction Lab, Department of Computer Science, University of Colorado Boulder, Boulder, CO, United States
[4]Colorado Health Outcomes Program, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States
[5]School of Pharmacy, University of Colorado Anschutz Medical Campus, Aurora, CO, United States
[6]School of Pharmacy, Southern Illinois University Edwardsville, Edwardsville, IL, United States
[7]Division of General Internal Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

Corresponding Author:
Stephen E Ross, MD
Division of General Internal Medicine
School of Medicine
University of Colorado Anschutz Medical Campus
Mail Stop B180
12631 E. 17th Avenue
Aurora, CO, 80045
United States
Phone: 1 303 724 2267
Fax: 1 303 724 2270
Email: steve.ross@ucdenver.edu

Abstract

Background: Older adults often have multiple chronic problems requiring them to manage complex medication regimens overseen by various clinicians. Personal health applications (PHAs) show promise assisting in medication self-management, but adoption of new computer technologies by this population is challenging. Optimizing the utility of PHAs requires a thorough understanding of older adults' needs, preferences, and practices.

Objective: The objective of our study was to understand the medication self-management issues faced by older adults and caregivers that can be addressed by an electronic PHA.

Methods: We conducted a qualitative analysis of a series of individual and group semistructured interviews with participants who were identified through purposive sampling.

Results: We interviewed 32 adult patients and 2 adult family caregivers. We identified 5 core themes regarding medication self-management challenges: seeking reliable medication information, maintaining autonomy in medication treatment decisions, worrying about taking too many medications, reconciling information discrepancies between allopathic and alternative medical therapies, and tracking and coordinating health information between multiple providers.

Conclusions: This study provides insights into the latent concerns and challenges faced by older adults and caregivers in managing medications. The results suggest that PHAs should have the following features to accommodate the management strategies and information preferences of this population: (1) provide links to authoritative and reliable information on side effects, drug interactions, and other medication-related concerns in a way that is clear, concise, and easy to navigate, (2) facilitate communication between patients and doctors and pharmacists through electronic messaging and health information exchange, and (3) provide patients the ability to selectively disclose medication information to different clinicians.

(J Med Internet Res 2011;13(2):e44) doi:10.2196/jmir.1813

KEYWORDS

Older adults; medication management; health records, personal

Introduction

Medication self-management is essential to drug safety but remains a challenging issue to address. Many patients, particularly older adults, have problems understanding their medication regimens or remembering to take their medications despite the use of patient information sheets and pillboxes [1]. Older adults experience more comorbidity, which increases the complexity of medication regimens, the risk of nonadherence, and the likelihood of fragmented care [2,3]. As a result, preventable adverse drug events in ambulatory care are more common among older adults [4].

Paper-based personal health records (PHRs) improve medication self-management for older adults in care transitions [5]. In theory, a PHA could provide on-demand, personalized, authoritative information on health issues, enhance management of medication information, and improve communication with health professionals and caregivers wherever a person may be [6,7]. However, privacy [8-13], usability [14-16], and updating [16] issues are potential barriers to the adoption of PHAs by older adults and caregivers. To better understand these issues, as part of a user-centered design process, we employed qualitative research techniques to elucidate the medication self-management needs and strategies of older adults and their adult caregivers that could be addressed through effective PHA design.

Methods

This study was conducted under the review and oversight of the Colorado Multiple Institutional Review Board.

Inclusion Criteria

All participants had to meet the following inclusion criteria: English fluency, the ability to pass a brief cognitive screen (provide name, year of birth, age, and telephone number), and their agreement that they would consider use of a computer application to manage health information. Patients of interest were defined as persons 65 years old or older, taking at least 3 prescription medications as an outpatient, with 2 or more outpatient visits in the last year, and 1 or more chronic medical conditions. Participants could either be “patients of interest” or “adult family caregivers” (caregivers for spouses, family members, or friends) of patients of interest.

Recruitment

We used a purposive sampling strategy to capture variations in education, ethnicity, income levels, and living arrangements. Participants were recruited from 4 sites in the metropolitan areas of Denver and Boulder, Colorado: (1) an academic hospital-based ambulatory geriatric clinic in Aurora, Colorado serving adults living in private homes, (2) a municipal senior citizen center in Denver, Colorado catering to the surrounding working-class community, (3) a small, independent-living residential facility in suburban metro Denver with middle-class clientele, and (4) an assisted/independent-living facility in Boulder, Colorado serving wealthier and more highly educated older adults. Recruitment was coordinated with facility directors, and research team members visited sites in person to explain the study and distribute flyers about it. Participants were recruited via these flyers or through word of mouth and were given a gift card for $25 to a local grocery store for participation in the study.

Data collection

First, 2 exploratory focus groups were conducted (15 participants in total) to broadly survey medication management challenges and to refine the topic guide for subsequent individual interviews. Minor refinements to the topic guide were made after a pilot interview with 2 older adults. We conducted 1- to 2-hour interviews in the homes of 16 participants, and one interview took place in a hospital. Researcher observations were recorded and digital photographs were taken to document patient arrangements for medication management techniques. After qualitative analysis of the individual interviews, 2 confirmatory focus groups were conducted to test and refine the themes and conclusions that were derived. All interviews and focus groups used semistructured topic guides, and were audio-recorded using a digital recorder, transcribed for analysis by a transcriptionist, and reviewed for accuracy by the research assistant who conducted the interviews and focus groups. During all phases of data collection and analysis, an advisory board of health professionals reviewed protocols, topic guides, and interim findings. All interview and focus group guides are available in the multimedia appendices.

Qualitative analysis

Data collection activities provided data in the form of field notes, photographs, and interview and focus group transcriptions. A member of the research team then systematically coded these using ATLAS.ti (6.0, Scientific Software Development GmbH, Berlin, Germany) using both a deductive approach based on the initial agreed-upon themes to look for and an inductive approach that allowed new themes to emerge from the data. Each primary code and its associated quotations were then reviewed and discussed by 3 members of the multidisciplinary research team. Discrepancies in analysis were discussed, revised, and synthesized into a core set of themes. Members of the research team identified quotes during analysis that were particularly illustrative of these themes.

Results

Participants in focus groups and interviews were 32 older adult patients and 2 family caregivers. From this pool of participants, 15 participated in the exploratory focus groups, 17 participated in individual interviews, and 10 participated in the confirmatory focus groups. Five older adults participated in both exploratory focus groups and individual interviews; 3 older adults participated in both individual interviews and confirmatory focus groups. Demographic information was collected from individual interview participants only (Table 1).

Table 1. Demographics of interview participants[a]

	Older patients (n = 15)	Family caregivers (n = 2)
Age (mean years, range)	82, 73–90	53, 48–57
Race: white (n, %)	13, 87%	1, 50%
Ethnicity: Hispanic (n, %)	2, 13%	1, 50%
Gender: female (n, %)	9, 60%	2, 100%
Has computer access (n, %)	10, 67%	2, 100%
Use of computer (n, %)		
Regular	5, 33%	2, 100%
Rare	3, 20%	0, 0%
None	7, 47%	0, 0%
Has Internet access (n, %)	8, 53%	2, 100%
Use of Internet (n, %)		
Regular	4, 27%	2, 100%
Rare	3, 20%	0, 0%
None	8, 53%	0, 0%

[a] The only demographic data captured from focus groups was gender and is not included here.

There were 13 individual interviews, composed of 15 older patient participants and 2 family caregivers, which yielded more detailed information about medication management. All of the 15 older patient participants managed their own medications, and 5 also assisted a friend or spouse in managing their medications. In 4 of the 13 interviews, 2 participants were present, and in 3 of the 13 interviews, a participant was interviewed in the presence of the cognitively impaired person they cared for.

Our qualitative analysis identified 5 key concerns of older adults surrounding medication self-management: seeking reliable medication information, maintaining autonomy in medication treatment decisions, worrying about taking too many medications, reconciling information discrepancies between allopathic and alternative medical therapies, and tracking and coordinating health information with and between multiple providers.

Seeking Reliable Medication Information

Participants most often sought medication information from pharmacists and clinicians, followed by the insert that comes with medications and the Internet. Other less frequent medication information sources cited were family, friends, reference books, and nurses. Information sources were selected based on accessibility and the type of information needed. The most readily available information (eg, the Internet and the medication insert) was not always seen as the most credible or useful. Participants tended to use various information-seeking strategies to get what they needed.

Pharmacists were generally the most trusted source of medication information, based on their knowledge of drugs and interactions. Participants who ordered prescriptions by mail felt that pharmacist consultation by toll-free number was trustworthy and convenient.

A majority of participants also named their doctor as a source for medication information. However, while doctors were highly valued for questions about general health issues, many participants were concerned that doctors chose medications without fully considering drug interactions or costs. Doctors were least likely to be consulted if a timely response was required. About half the participants said, unprompted, that they felt their doctor was too busy to address their medication information needs. Overall, while some participants described long-term, trusting relationships with doctors, many felt that answers to medication questions during rushed visits were unclear. For follow-up questions, many were frustrated about not being able to talk to their doctor on the phone. Although office staff or nurses could be intermediaries for medication questions by telephone, participants preferred to speak directly to the doctor. If this was not possible, some participants said that they would schedule an appointment just to ask questions about their medications.

Many participants also used the insert that comes with medications as a resource, but it was most often used when new medications were prescribed, not when questions arose in ongoing medication use. Those who used inserts primarily looked for information on side effects or interactions. Some kept inserts in elaborate files of personal health information (Figure 1, Figure 2), but rarely consulted them after they were filed.

Figure 1. Example of one participant’s personal health information management area

Figure 2. Example of another participant’s personal health information management area

Of the 9 participants who used the Internet, most used it to search for medication information with the Google search engine. The Internet was used for finding quick information, but was usually used in conjunction with other information sources.

> *A lot of times I ask the pharmacist [questions about my medications] because the pharmacist often times knows some of the connections that maybe the physician is not quite up on ...If I have a real question I would certainly ask my doctor... And as I said, I do go on the Internet if I just want some information. For instance, this doctor had prescribed medication for me and I started taking it and I was disturbed because I woke up in the middle of the night and my mouth was so dry and my throat was so dry and now I was having trouble with urination and I thought you know, that medication is doing this to me. And I'm stopping taking it. So I immediately went on the Internet and even though I had information— well just the brochure that was in it... And I just went on the Internet to find out. And if I ever question something I want to know right away that's what I do.* [Interview participant 1014]

Some felt overwhelmed by the number of sources available on the Internet. Others were skeptical about the reliability of websites or did not feel comfortable using the Internet in general.

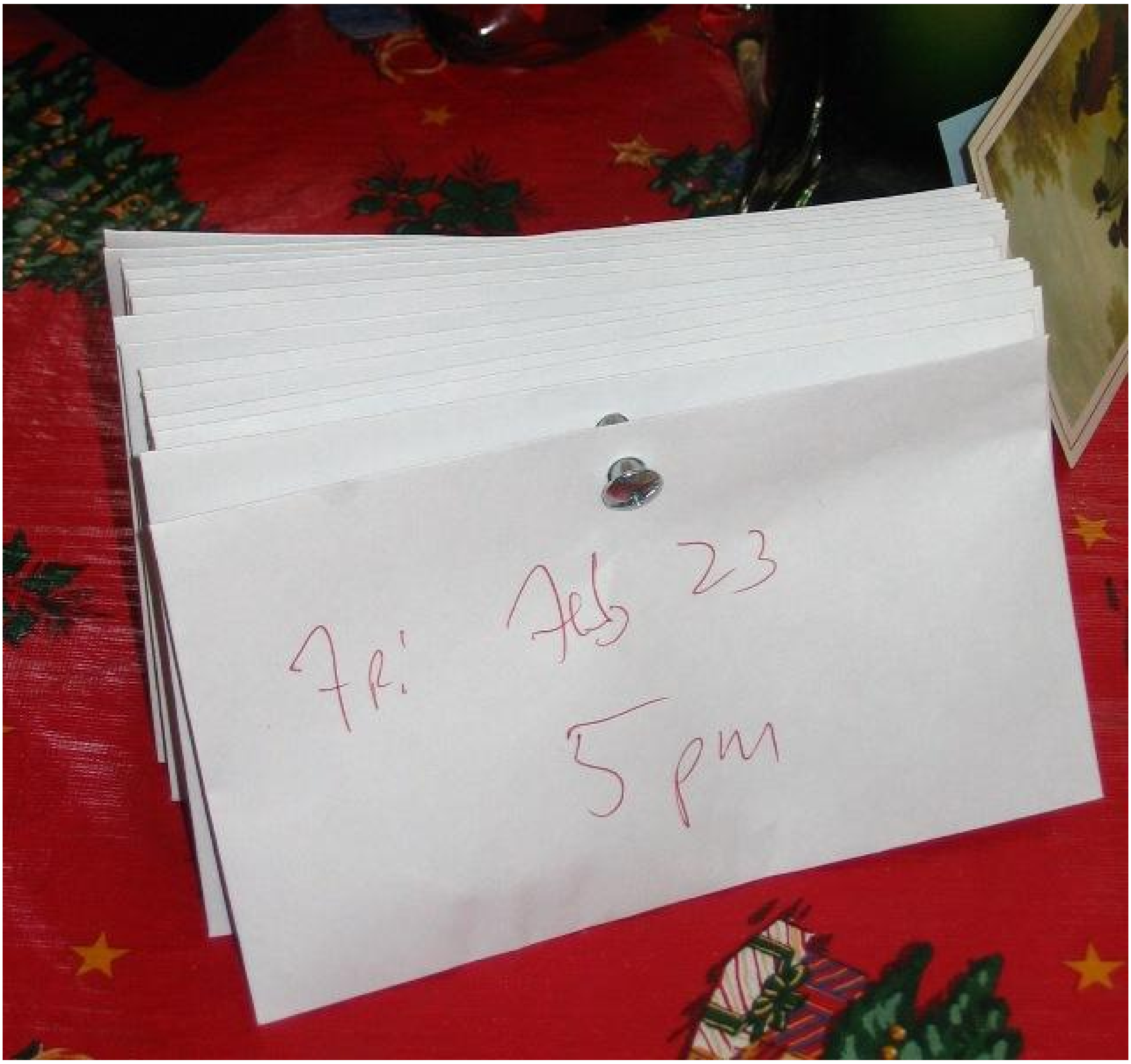
Maintaining Autonomy in Medication Treatment Decisions

Participants expressed a desire to remain independent and in control of their medication treatment decisions, despite acknowledging a decline in their mental and physical abilities and sometimes feeling overwhelmed by complicated medication regimens. While some participants followed their doctor's recommendations without question, others wanted to play a more participatory role in treatment decisions. The latter were more assertive in seeking information, actively monitoring their health, and altering their medication regimens accordingly. For these participants, maintaining autonomy in medication management appeared to be part of a broader desire to maintain autonomy in various aspects of daily life.

Certain medications were considered a threat to autonomy because they were perceived as detrimental to cognition or health. Common reasons participants stopped or altered medication regimens included experiencing negative side effects, feeling they know their own bodily responses to medications (or those of the individual they care for) better than the doctor, or feeling they were on too many medications (a theme that is expanded upon in the next section). When deciding to stop medications, some participants discussed this and negotiated with their doctors while others acted unilaterally. A few made decisions to alter their regimen without consulting with their doctor because they felt their input was not well received in clinic visits.

When dosing intervals conflicted with daily activities, the regimen was occasionally perceived as a threat to an individual's autonomy. Mid-day doses were particularly inconvenient, and several participants mentioned that they frequently forgot them. Others developed elaborate systems with pill boxes, envelopes (Figure 3), or spatial and temporal orderings of medications in multiple locations to integrate medication reminders into everyday life. For instance, participants stored pills in locations where they were likely to come across them at the scheduled dosing time, such as living rooms where they watch television for evening pills, the kitchen for mid-day pills, or the bedroom for nighttime pills.

Figure 3. A caregiver created this medication management system to provide some autonomy to her older adult mother. Her mother could access these envelopes fastened together with a pin and pull off the appropriate envelope to receive the correct dosage. This system was also used to carry her mid-day pills with her if she was out of the house



Worrying about Taking Too Many Medications

In every interview and focus group we conducted, participants expressed concerns about taking too many medications. They were skeptical of the rationale for and purported benefits of taking multiple medications, particularly for the same condition (eg, hypertension). Most participants desired to reduce the amount of medications they were taking and either approached their doctor, or made unilateral decisions to wean themselves or their loved one from some medications:

> *I have a lot of hesitancy with respect to medications and I feel like she is getting too much medication and more than likely some of her problems are caused by medication. So I've been attempting to modify her medications somewhat. Like weaning her off of prescribed depression medication and using the homeopathic medicine. So in a sense I am being somewhat of a diagnostician and dispenser.* [Interview participant 1010]

Apprehension about taking too many medications had many dimensions. There was a prevalent view that the potency of medication regimens prescribed to older adults was often inappropriate:

> *My biggest frustration with the doctors is overdosing...Here a couple years ago I ended up in the hospital because I was getting bad dizzy spells and I couldn't breathe and after they checked me out they found out they had overdosed me with medications— the doctor had. So they had to readjust my blood work and medications all over again.* [Exploratory focus group 1 participant]

Another common worry was that older adults' bodies cannot handle multiple medicines:

> *As a caregiver, [I feel that] doctors give too many medications to older people and their bodies— our bodies— can't process them well and all kinds of disastrous things happen. I've seen this many, many times. It's just too much. The liver doesn't work like it did when you are younger...It's just here take this, take this, take this, and then the person suffers.* [Confirmatory focus group 2 participant]

Participants were concerned about drug interactions and perceived that doctors tended to "layer on" medications, rather than simplifying regimens. Several questioned the concept of taking a medication indefinitely, fearing the medications may overtax vital organs or lead to a dependency. Some participants also voiced concerns about the number of medications they took because, even with medical insurance, it added unnecessary costs to their health care expenses. They felt doctors did not consider cost when prescribing medications.

Although overmedication was the largest area of concern for participants, many did not think their doctors shared their concerns, and as a result the issue was insufficiently addressed. Many shared stories about doctors unwilling to work with them to simplify medication regimens, address their worries, or explain the rationale behind polypharmacy.

Reconciling Information Discrepancies between Allopathic and Alternative Medical Therapies

Participants consistently distinguished prescription medications from nonprescription medications (eg, vitamins and supplements) as 2 distinct categories. When participants were asked what medications they took, they typically named all prescription medications, and often had to be prompted to list supplements and nonprescription medications. Likewise, nonprescription medications were often separated or omitted from medication lists that they kept on hand and were stored in different locations from prescription medications. Many participants did not feel it necessary to discuss the nonprescription medications they were taking with their doctor. Others had questions about the value of nonprescription medications, vitamins, and supplements in their complex drug regimens, but lacked information or direction because they typically did not think it was appropriate to talk to their doctor about them.

Several participants in the study visited practitioners of alternative forms of health care that included nutritionists, chiropractors, acupuncturists, or homeopaths. Many of these participants were frustrated about conflicting information that they received from their primary care doctor and alternative care providers regarding their health, the etiology of their illnesses, and the safety and effectiveness of their conventional and complementary medication regimens. In general, these participants did not feel that their allopathic doctors supported alternative care and therefore did not always tell them about alternative medications or therapies that they were pursuing. For 2 participants in particular, this was a source of confusion and stress regarding medication management and decision making:

> *The biggest problem in my mind for my personal planning and decision making is the kind of conflict between my primary care doctor, traditional medicine man and this alternative [practitioner] — who happens to be a chiropractor— but he has done a big study of supplements. Consequently I take 20 or 25 pills a day. My primary care physician just doesn't care a hoot about all those supplements ... Most people do not take all the pills I take, I've discovered. They take prescription pills, but I take magnesium and calcium and some brain pills and all kinds of stuff, and those are very important to the alternative medicine person in my life. But they are just uh ... minimized totally by the primary care doctor so I feel like I'm fighting two sides.* [Confirmatory focus group 2 participant]

Another participant asked:

> *If you don't follow the regimen, do you feel any different?*

Response:

> *"Yes, I do...But then I don't know which one to blame or credit!"*

Tracking and Coordinating Health Information between Multiple Providers

Participants described 2 dimensions of personal responsibility for coordinating health information: first, keeping track of health information through record keeping, creating medication lists, and compiling medication information, and, second, acting as a coordinator through assisting in the transfer of health information between multiple doctors and caregivers. Most kept paper medical files for bills, insurance papers, and medication information inserts. One couple, who split residency between 2 states and had doctors in both, made photocopies of all their medical paperwork and physically moved the papers between states in portable file cabinets. Almost all participants kept medication lists to assist with medical appointment paperwork and in cases of emergency, although some were outdated or illegible.

Although all participants in the study partook in some form of record-keeping behavior, the extent to which participants felt that they needed to review or transfer this information depended largely on whether their multiple doctors were in a closed system. When doctors were outside the umbrella of a single hospital, health maintenance organization, or other network of care, participants did not assume that information would be exchanged automatically or in a timely manner, though they expressed a desire for this to happen. In such cases, participants felt that they needed to take a more proactive role in knowing their health information and acting as a medium between doctors:

> *I see two different doctors on a regular schedule and the only thing I really worry about is if one of them changes my medicine I want the other one to know about it because I get medicine from each of them and I don't want to add something or take away*

> *something that is going to cause problems.* [Exploratory focus group 2 participant]

Conversely, when doctors did fall within the same network, participants assumed that they were using the same computer system and could view health and medication information from other doctors. In this case, participants tended not to take as prominent a role in managing their health information because they felt confident their records were well maintained, information was being shared, and care was well coordinated.

Personal factors also played a role in the degree to which participants actively coordinated health information. Some participants reported being compulsive about record keeping in general, while others seemed unwilling or unable to actively manage their health information because it was overwhelming or beyond their cognitive abilities.

Discussion

As part of a user-centered design process for a PHA for older adults, we identified 5 core concerns that older adults have with regard to medication self-management. These findings, many of which corroborate findings from previous studies, suggest functional requirements for PHAs to assist older adults in medication management.

Participants' desire for reliable medication information is consistent with previous research on consumers' use of health websites [17], as well as research on keeping certain medication information, such as inserts, in case they are needed [18]. Current PHAs commonly link drug names on a medication list to information about that drug. These links can provide useful information about medication indications and potential side effects, particularly if the information is written for a lay audience (rather than simply replicating the medication insert).

However, more could be done to address older adults' common latent concerns about overmedication. When polypharmacy is necessary, patients and caregivers want to understand the rationale for it and want to be reassured that long-term use of multiple medications will not damage the body in some way. Current PHAs typically provide information about individual medicines but provide limited information about the medication regimen as a whole. When information on potential drug interactions is provided, the gravity of interactions is unclear. PHAs that allow patients to document troubling symptoms that could indicate side effects and explore whether medications could be the cause may be beneficial in addressing these patient concerns. Ideally, PHAs could help patients assess appropriateness of the drug regimen as a whole through automated analysis or consultative services. While current PHAs focus on electronic messaging with physicians, the ability to share information and receive advice from pharmacists could be even more valuable, since patients find pharmacists to be the most credible resource for medication questions. This is consistent with the vision that PHAs should empower patients by providing new avenues of access to useful and necessary health services [19].

PHAs should also be flexible to accommodate older adults' desire for control over their regimens. Older adults who self-medicate or experiment with medications to get a desired effect or lessen side effects or lifestyle conflicts are unlikely to be satisfied with a PHA that provides a medication list owned and managed by a physician or practice (as in patient portals, also known as tethered PHRs) [15]. These patients are likely to be more satisfied with a PHA that lets them edit medication lists to reflect changes they have made to their medication regimens. On the other hand, patient portals have the advantage of autopopulating medication information [20]. More flexible interoperable PHAs may need to avoid overtaxing older adults with medication input and reconciliation tasks.

PHAs can also help address the common challenges older adults face in tracking and coordinating care. Previous research has demonstrated the appeal of sharing electronic personal health information among multiple providers to improve care coordination [20,21]. Our findings suggest that PHAs should support care coordination not only among allopathic practitioners and pharmacists, but also with practitioners of complementary medicine. PHAs should allow and encourage nonprescription and alternative medications to be added to medication lists. Authoritative, nonjudgmental information on these medications, and their appropriateness in the medication regimen as a whole, should be addressed. Additionally, because some patients are wary of disclosing their use of alternative medications to allopathic doctors, it may be useful for PHAs to provide patients with the ability to selectively disclose medication information to different practitioners. Given the distinction that older adults and caregivers make between keeping records and coordinating care, PHAs may also need to accommodate multiple views: (1) a more inclusive and comprehensive view for home use, and (2) a more streamlined or compact view appropriate for assessment by practitioners and for mobile use during clinic or emergency visits.

Our findings support and extend previous research demonstrating that medication self-management is a common and challenging issue for older adults and their caregivers [1,3,5]. Strengths of this study include a stepwise analytic approach, with refinement of topic guides based on exploratory focus groups, investigator triangulation in the analysis of transcripts, and the use of focus groups for member checking of findings derived from individual interviews. The research team included expertise in internal medicine, pharmacy, information technology, medical anthropology, and communications. This variety provided the ability to examine data from several perspectives. The qualitative nature of this study also allowed participants to elucidate their primary medication management concerns in an unrestricted manner and gave us more detail about the nature of these concerns and how they could be addressed. The primary limitation of this study is the small sample size, particularly of caregivers. Responses also could have been affected in those instances where caregivers were interviewed with patients as a dyad and when they participated in a focus group with patients [22]. Because of the small sample size and purposive sampling strategy, these findings from participants in the Denver-Boulder metropolitan area may not be fully generalizable to other regions.

This study was undertaken as part of a larger project to develop a prototype interoperable PHA. Our findings have elucidated specific medication management challenges that will inform development of PHAs that are appropriate for the increasingly important but hard to reach older population [20,23]. Since the original qualitative analysis, our team has developed a prototype PHA responsive to these findings [24]. The public release of interoperable platforms to support PHAs that can be tailored for different user populations (eg, Dossia, Google Health, and Microsoft HealthVault) should facilitate development and dissemination of PHAs responsive to the particular needs of older adults.

Acknowledgments

This research was funded by the Robert Wood Johnson Foundation Project Health Design Grant RWJ59880. Carolyn T Nowels, MSPH, and Debbi S Main, PhD, provided guidance on qualitative analysis techniques. Eric A Coleman, MD, advised the authors on key issues in transitional care. We would like to thank our participants and advisory board for their valuable feedback.

Conflicts of Interest

None declared

Multimedia Appendix 1

Interview guide:

[PDF file (Adobe PDF File), 115 KB - jmir_v13i2e44_app1.pdf]

Multimedia Appendix 2

Participant characteristics survey

[PDF file (Adobe PDF File), 78 KB - jmir_v13i2e44_app2.pdf]

Multimedia Appendix 3

Focus group facilitation guide

[PDF file (Adobe PDF File), 119 KB - jmir_v13i2e44_app3.pdf]

References

1. Manias E, Claydon-Platt K, McColl GJ, Bucknall TK, Brand CA. Managing complex medication regimens: perspectives of consumers with osteoarthritis and healthcare professionals. Ann Pharmacother 2007 May;41(5):764-771. [doi: 10.1345/aph.1H623] [Medline: 17456543]
2. Steinman MA, Hanlon JT. Managing medications in clinically complex elders: "There's got to be a happy medium". JAMA 2010 Oct 13;304(14):1592-1601 [FREE Full text] [doi: 10.1001/jama.2010.1482] [Medline: 20940385]
3. Wilson IB, Schoen C, Neuman P, Strollo MK, Rogers WH, Chang H, et al. Physician-patient communication about prescription medication nonadherence: a 50-state study of America's seniors. J Gen Intern Med 2007 Jan;22(1):6-12. [doi: 10.1007/s11606-006-0093-0] [Medline: 17351835]
4. Gurwitz JH, Field TS, Harrold LR, Rothschild J, Debellis K, Seger AC, et al. Incidence and preventability of adverse drug events among older persons in the ambulatory setting. JAMA 2003 Mar 5;289(9):1107-1116 [FREE Full text] [Medline: 12622580]
5. Coleman EA, Parry C, Chalmers S, Min SJ. The care transitions intervention: results of a randomized controlled trial. Arch Intern Med 2006 Sep 25;166(17):1822-1828 [FREE Full text] [doi: 10.1001/archinte.166.17.1822] [Medline: 17000937]
6. Brennan PF, Downs S, Casper G, Kenron D. Project HealthDesign: stimulating the next generation of personal health records. AMIA Annu Symp Proc 2007;Oct 11:70-74. [Medline: 18693800]
7. Siek KA, Khan DU, Ross SE. A usability inspection of medication management in three personal health applications. In: Kurosu M, editor. Human Centered Design: First International Conference, HCD 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. New York, NY: Springer; 2009:129-138.
8. Angst W. Privacy safeguards in PHR adoption. Health Manag Technol 2008 Apr;29(4):40, 39. [Medline: 18468218]
9. Brenner B. Is the provision of laboratory results via the Internet acceptable to patients? A survey of private patients in a large, specialist gynaecology practice. N Z Med J 2003 Dec 12;116(1187):U711. [Medline: 14752539]
10. Denton IC. Will patients use electronic personal health records? Responses from a real-life experience. J Healthc Inf Manag 2001;15(3):251-259. [Medline: 11642143]

11. Grimes-Gruczka T, Gratzer C. Institute For the Future. 2000. Ethics Survey of Consumer Attitudes About Health Web Sites URL: http://www.chcf.org/~/media/Files/PDF/E/PDF%20Ethics2ndEdition.pdf [accessed 2011-03-15] [WebCite Cache ID 5xD4i9CNe]
12. Lafky DB, Horan TA. Health Status and Prospective PHR Use. AMIA Annu Symp Proc 2008;Nov 6:1016. [Medline: 18999066]
13. Nelson R. The personal health record. Am J Nurs 2007 Sep;107(9):27-28. [doi: 10.1097/01.NAJ.0000287504.92887.de] [Medline: 17721144]
14. Kim MI, Johnson KB. Personal health records: evaluation of functionality and utility. J Am Med Inform Assoc 2002;9(2):171-180. [Medline: 11861632]
15. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126. [doi: 10.1197/jamia.M2025] [Medline: 16357345]
16. Personal Health Working Group. Markle Foundation. 2003. Personal Health Working Group: Final Report URL: http://www.providersedge.com/ehdocs/ehr_articles/The_Personal_Health_Working_Group_Final_Report.pdf [accessed 2011-03-15] [WebCite Cache ID 5xD52Zykh]
17. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
18. Marquard JL, Moen A, Brennan PF. Photographic data--An untapped resource to explore complex phenomena such as health information management in the household (HIMH). Stud Health Technol Inform 2006;122:58-62. [Medline: 17102218]
19. Bonander J, Gates S. Public health in an era of personal health records: opportunities for innovation and new partnerships. J Med Internet Res 2010;12(3):e33 [FREE Full text] [doi: 10.2196/jmir.1346] [Medline: 20699216]
20. Kim EH, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, et al. Challenges to using an electronic personal health record by a low-income elderly population. J Med Internet Res 2009;11(4):e44 [FREE Full text] [doi: 10.2196/jmir.1256] [Medline: 19861298]
21. Weitzman ER, Kaci L, Mandl KD. Acceptability of a personally controlled health record in a community-based setting: implications for policy and design. J Med Internet Res 2009;11(2):e14 [FREE Full text] [doi: 10.2196/jmir.1187] [Medline: 19403467]
22. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302 [FREE Full text] [Medline: 7633241]
23. Horan TA, Botts NE, Burkhard RJ. A multidimensional view of personal health systems for underserved populations. J Med Internet Res 2010;12(3):e32 [FREE Full text] [doi: 10.2196/jmir.1355] [Medline: 20685644]
24. Siek KA, Ross SE, Khan DU, Haverhals LM, Cali SR, Meyers J. Colorado Care Tablet: the design of an interoperable Personal Health Application to help older adults with multimorbidity manage their medications. J Biomed Inform 2010 Oct;43(5 Suppl):S22-S26. [doi: 10.1016/j.jbi.2010.05.007] [Medline: 20937480]

Abbreviations

PHA: personal health application
PHR: personal health record

Edited by G Eysenbach; submitted 22.03.11; peer-reviewed by S Reti, J Wald; comments to author 18.04.11; revised version received 27.04.11; accepted 11.05.11; published 29.06.11.

Please cite as:
Haverhals LM, Lee CA, Siek KA, Darr CA, Linnebur SA, Ruscin JM, Ross SE
Older Adults with Multi-Morbidity: Medication Management Processes and Design Implications for Personal Health Applications
J Med Internet Res 2011;13(2):e44
URL: http://www.jmir.org/2011/2/e44/
doi:10.2196/jmir.1813
PMID:21715286

Original Paper

Parent Satisfaction With the Electronic Medical Record in an Academic Pediatric Rheumatology Practice

Paul Rosen[1*], MD MPH; Steven J Spalding[2*], MD; Michael J Hannon[3*]; Robert M Boudreau[4*], PhD; C Kent Kwoh[5*], MD

[1]Children's Hospital of Pittsburgh of UPMC, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States
[2]The Cleveland Clinic, Cleveland, OH, United States
[3]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States
[4]Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States
[5]Department of Medicine, University of Pittsburgh, Pittsburgh VA Healthcare System, Pittsburgh, PA, United States
[*]all authors contributed equally

Corresponding Author:
Paul Rosen, MD MPH
Children's Hospital of Pittsburgh of UPMC
University of Pittsburgh School of Medicine
One Children's Hosptial Drive
4401 Penn Avenue
Pittsburgh, PA, 15224
United States
Phone: 1 412 692 3294
Fax: 1 412 692 5054
Email: Paul.Rosen@chp.edu

Abstract

Background: Patient satisfaction has not been widely studied with respect to implementation of the electronic medical record (EMR). There are few reports of the impact of the EMR in pediatrics.

Objective: The objective of this study was to assess the impact of implementation of an electronic medical record system on families in an academic pediatric rheumatology practice.

Methods: Families were surveyed 1 month pre-EMR implementation and 3 months post-EMR implementation.

Results: Overall, EMR was well received by families. Compared with the paper chart, parents agreed the EMR improved the quality of doctor care (55% or 59/107 vs 26% or 26/99, $P < .001$). More parents indicated they would prefer their pediatric physicians to use an EMR (68% or 73/107 vs 51% or 50/99, $P = .01$).

Conclusions: Transitioning an academic pediatric rheumatology practice to an EMR can increase family satisfaction with the office visit.

(J Med Internet Res 2011;13(2):e40) doi:10.2196/jmir.1525

KEYWORDS

Electronic medical record; pediatric rheumatology; ambulatory care

Introduction

Implementation of the electronic medical record (EMR) is currently taking place across industrialized countries. In 1991, the Institute of Medicine (IOM) issued a report concluding that computer-based patient records were an “essential technology” for health care and in 1997 called for the widespread adoption of a computer-based patient record over the next 10 years [1]. Little work has been done to study the direct influence of information technology on patient-physician relationships [2].

The 2001 IOM report, *Crossing the Quality Chasm: A Hew Health System for the 21st Century*, targeted six areas of health care that required significant improvement: safety, efficacy, timeliness, efficiency, equality, and patient-centeredness [3]. The report delineated how health information technology (HIT) was necessary to achieve all six aims. The United States government has recently reinforced the call issued for further adoption of HIT in the IOM report with the allocation of more than US $20 billion in the American Recovery and Reinvestment Act of 2009. However, some have argued that the use of HIT can hurt the delivery of patient-centered care [4-7].

Widespread EMR adoption has been slow. In 2008, DesRoches and colleagues surveyed over 2700 physicians and found that the implementation rate of a fully functional electronic medical record (EMR) system was 4%. Another 13% of physicians were using a more basic electronic medical system with limited functionality [8].

Few studies have examined the impact of EMR implementation in the pediatric setting, and, to our knowledge, none have looked at its impact on parent satisfaction and the doctor-patient relationship in ambulatory subspecialty pediatrics. At our academic children's hospital, ambulatory offices started conversion to an EMR in 2005. While all 20 ambulatory practices converted to the EMR by 2009, we were the first practice to make the conversion. Our study was designed to determine if the transition to EMR was associated with changing parent satisfaction.

Methods

This study was conducted at an academic pediatric rheumatology practice in Pittsburgh, Pennsylvania. The practice conducts over 95% of its patient visits in the ambulatory setting and has a small inpatient hospital service. The outpatient practice adopted a complete electronic medical record (Cerner PowerChart Office, Cerner Corporation, Kansas City, MO)

Prior to the transition to the EMR, the physicians worked with Cerner programming staff to design a universal rheumatology template that could be used for all outpatient rheumatology visits. The components of the computerized note template that were unique to pediatric rheumatology included a complete joint examination in the physical examination section. There was also a screening section that served as a reminder to the physicians to document routine pediatric rheumatology screening such as completion of eye examinations for uveitis, influenza vaccinations for immunosuppressed patients, and purified protein derivative (PPD) status. Physicians worked with the technical staff to develop the template and test the final note prior to implementation. The differences in the workflow between the pre-EMR paper charts and the post-EMR implementation are illustrated in Table 1.

Table 1. Comparison of office practices pre- and post-EMR implementation

Office Practice	Pre-EMR (Paper Charts)	EMR
Office visit letters	Physician dictates via telephone	Physician creates Cerner PowerNote via computer
Lag time for letter to be sent to referring physician	Letter faxed 3 to 5 business days after dictation	EMR note faxed immediately upon completion and signature
Prescription format	Hand written	EMR-generated via Cerner EZScript
Billing	Hand written on a form by physician and then entered electronically by secretary	Entered electronically by physician
Orders for laboratory tests, diagnostic tests, radiographs, consultations, etc	Hand written on forms or prescriptions	Entered electronically by physician with EMR-generated paper form for patient
Messaging between medical staff members	Paper slip attached to paper chart	Electronic messaging in patient's chart via electronic in-box
Laboratory, pathology, and radiograph result endorsement by physician	Result on paper initialed	Electronic endorsement via electronic in-box
Results or other medical records from outside hospitals	Papers manually added to paper charts	Electronically scanned into EMR

Computers were placed in all six patient examination rooms. In all, three computers were placed in the common work area for the use of 3 attending physicians, 2 nurses, and trainees (fellows and residents). Computers and printers were installed with existing office furniture. The layout of the examination rooms was not changed to accommodate the new workflow.

Physicians received three training sessions (one hour each) on EMR usage and specifically on how to integrate computers into the visit. On-site technical support was available during the 6-week implementation phase. Support by telephone was provided after that. The patient schedule was reduced by 50% for the first 2 weeks of EMR implementation. The schedule was then reduced by 25% for the following 4 weeks of EMR implementation. During the 6 weeks that the schedule was reduced, there was no change in time allotment for each patient appointment.

Patients

Participants were the parents of the patients of two of the authors. The patients were children with diagnoses typical of a pediatric rheumatology practice including juvenile idiopathic arthritis, fibromyalgia, systemic lupus erythematosus, dermatomyositis, scleroderma, vasculitis, and other chronic autoimmune conditions. This was a convenience sample, and all families coming to the office for their routine follow-up visits were asked to complete a parent survey after their visit was completed. All of the parents approached by the investigators agreed to complete the surveys.

For the 1-month period prior to transitioning to the EMR, we conducted surveys of the parents of all of our follow-up patients regarding the paper medical record. Starting 3 months after adoption of the EMR, for the next month families coming for their routine follow-up visits were surveyed. Names of the parents completing the surveys were not recorded so the number of patients who overlapped for the pre- and post-EMR implementation surveys could not be determined.

The parent surveys included 12 statements about the medical record (paper or electronic depending on the time point), quality of care received, family satisfaction with the office visit, and patient safety. The parents recorded their level of agreement with each statement on a 5-point Likert scale from *strongly agree* to *strongly disagree*. Of the 12 statements, 2 assessed computer usage (on a 5-point scale from *never* to 5 hours a day) and skill (on a 5-point scale from *poor* to *excellent*).

The study was submitted to the University of Pittsburgh institutional review board (IRB). The study was reviewed by the IRB, and it was determined that the project was primarily a quality assurance activity. The study was sent to two quality committees for review. The study was approved as a quality improvement study by the Children's Hospital of Pittsburgh Total Quality Council and by the University of Pittsburgh Medical Center Quality Assurance Committee.

Statistical Analysis

Because of the distribution of the responses, the survey responses were dichotomized into two groups: strongly agree/agree and strongly disagree/disagree/neutral prior to analysis. Fisher's exact test was used to compare responses for pre- and post-EMR implementation. Spearman correlation coefficients (ρ's) were calculated to evaluate the family's view of quality of care and the doctor-patient relationship after EMR implementation. Correlations were based on post hoc observations and analyses; they were not hypothesis driven. Correlation coefficients of .3 to .5 were considered weak; .5 to .7, moderate; and greater than .7, strong. Statistical Analysis Software (SAS Institute Inc, Cary, NC) was used to perform statistical analysis.

Results

Overall, families reported greater satisfaction with the EMR compared with the paper chart (Table 2).

Table 2. Parent survey of medical record pre- and 3 months post-EMR implementation

Statement	Paper Chart, *Strongly Agree* and *Agree* (n = 99) n (%)	Electronic Medical Record, *Strongly agree* and *agree* (n = 107) n (%)	*P* value[a]
1. The current medical record system increases the time the doctor spends with my child.	58 (59)	62 (58)	.99
2. I would like more of my child's physicians to use an electronic medical record system.	50 (51)	73 (68)	.01
3. I would *not* miss the current charting system if it was no longer available.	53 (54)	10 (9)	< .001
4. I am dissatisfied with the current charting system used by my child's doctor.	6 (6)	2 (2)	.16
5. I worry that my child's private medical chart may be seen by others.	22 (22)	23 (21)	.99
6. The charting system allows me to better communicate with my child's doctor.	45 (45)	56 (53)	.33
7. The current charting system helps me to understand my child's medical tests.	33 (33)	51 (48)	.047
8. The current medical record system improves the quality of care provided by my child's doctor.	26 (26)	59 (55)	< .001
9. I feel that the current medical record system distances me from my child's doctor.	5 (5)	4 (4)	.74
10. I feel that the current medical record system adequately prevents medical errors.	14 (14)	14 (13)	.84
11. The information in my child's chart is kept current.	69 (70)	89 (83)	.03
12. The staff have adequately addressed my concerns about the current medical record system.	39 (39)	73 (68)	< .001

[a] Based on Fisher's exact test

Parents rated the quality of care provided by the doctor using the EMR higher compared with the paper chart ($P < .001$). Parents felt that the information in the EMR was up-to-date compared with the paper chart ($P = .03$); this aspect of the EMR had the highest percentage of parents responding *agree/strongly agree* (83% or 89/107). In addition, parents felt that the EMR helped them to understand their child's medical tests compared with the paper chart ($P = .05$).

Parents reported that they would like more of their child's physicians to use an EMR ($P = .01$). The parents' desire for other physicians to use EMR correlated with reporting more time spent with the doctor ($\rho = 0.52$, $P < .001$), better communication with the doctor ($\rho = 0.50$, $P < .001$), better understanding of medical tests ($\rho = 0.42$, $P < .001$), and higher quality of care ($\rho = 0.41$, $P < .001$). Family satisfaction with physician communication correlated with reporting better understanding of medical tests ($\rho = 0.67$, $P < .001$) and improved quality of care ($\rho = 0.51$, $P < .001$).

The majority of parents reported better than average computer skills and more than 5 hours of computer usage a month.

Discussion

This study demonstrates the impact on families of implementation of an electronic medical record in an academic pediatric subspecialty practice. Parents indicated greater satisfaction with the EMR compared with paper charts. Parents reported a preference for their child's other physicians to use an EMR.

When planning a conversion to an EMR, physicians may be concerned that the quality of the visit with families will be compromised. A 2006 report from Israel focusing on the EMR in the outpatient setting showed that physicians spent between 25% and 42% of the visit gazing at the computer screen [6]. In this observational study, monitor gazing decreased physician psychosocial questioning and emotional responsiveness to the patient. In addition, physician keyboarding decreased both physician and patient contribution to the medical dialogue [6]. In our study, physician computer usage in the examination room did not result in a negative patient perception.

Other reports have documented beneficial effects of an EMR. Using videotaped encounters, Arar et al showed that EMR use enhanced patient-physician communication and safety by facilitating medication reconciliation [9]. The group also found that use of an EMR, compared with a paper chart, was more likely to result in documentation of a diagnosis, of advice given, and of a referral ordered. In 2007, Simon et al surveyed a random sample of over 1000 physicians in Massachusetts and assessed physicians' perceptions of the EMR in medical practice. They found that compared with physicians not using an EMR, physicians using an EMR reported greater patient-physician communication [10].

Our findings of improved family satisfaction with an EMR are similar to those reported in a longitudinal quantitative study on the impact of computers in ambulatory care on patient-physician interactions by Hsu et al in 2005 [11]. The group surveyed patients 2 months prior to EMR implementation, and 1 and 7 months after EMR implementation. By 7 months, the investigators found improvements in the following areas of patient satisfaction: overall visit, physician's level of familiarity with the patient, communication about medical issues, and degree of comprehension with decisions made during the visit. In addition, the patients did not feel that there was less time for discussion about psychosocial issues with the computers in the exam room. There were no decreases in patient satisfaction areas such as physician's personal manner, level of concern for the patient, and attention to listening [11].

In 2006, Kemper et al reported 21% of 1000 pediatricians responding to a survey had an electronic health record in their practice [12]. The group listed the perceived barriers to implementation among general pediatricians without an EMR. Barriers included: expense of implementation, lack of EMR that meets the needs of a pediatric practice, physician resistance, increase in physician workload, inadequate computer skills of providers, lack of improvement in patient care, interference with doctor-patient relationship, and concerns about patient confidentiality.

Limitations

Our study faced several potential limitations. First, this study was conducted in an academic pediatric subspecialty office at a tertiary care children's hospital thus limiting the application to the general pediatric physician workforce. Because the surveys were anonymous, we were not able to track participant overlap in pre- and post-EMR implementation surveys or compare response rate pre- and post-EMR for the exact same groups of patients. In addition, as early adopters of the technology, the physicians in our practice may not be representative of other pediatric subspecialists. In our study, the physicians caring for the patients also recruited the patients for the study. Thus, response bias among participants toward a parent's perception of how the physicians felt about the EMR may have been introduced. In addition, our use of convenience sampling may have lead to biased results. Because our surveys included the families of just two physicians, our results may not generalize to larger groups of physicians. Our results need to be replicated across broader settings with random survey recruitment. Future recruitments should separate the roles of clinician and study recruiter.

As part of a large institution, the rheumatology physicians did not experience the financial impact of the reduction in patient scheduling. In addition, there was no up-front capital investment required from the physicians in order to purchase, implement, and maintain the technology. This would not hold true for the majority of practicing pediatricians.

Conclusions

Health care leaders, government officials, and policy makers have been calling for a paperless medical record system for over 2 decades. Despite a call for universal adoption of an EMR for all patients, implementation has been limited. In pediatrics, the challenges of implementing an EMR in a small practice are real. In our academic, hospital-supported, subspecialty practice we were able to transition to an EMR with a subsequent increase in patient satisfaction. More study in the area of EMR implementation in pediatrics is warranted.

Acknowledgments

The authors would like to thank Jim Levin, MD, Jonathan Bickel, MD, Anthony Fiorillo, MD, Daniel Martich, MD, Anna Rusztyn-Fitz, Jacque Dailey, and Christopher Gessner for their insights on the electronic medical record. This article is dedicated to the memory of Eugene S Wiener, MD.

Conflicts of Interest

None declared

References

1. Institute of Medicine. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington, DC: National Academy Press; 1997.
2. Weiner M, Biondich P. The influence of information technology on patient-physician relationships. J Gen Intern Med 2006 Jan;21(Suppl 1):S35-S39. [doi: 10.1111/j.1525-1497.2006.00307.x] [Medline: 16405708]
3. Institute of Medicine. Crossing the Quality Chasm: A New Health Care System for the 21st Century. Washington, DC: National Academy Press; 2001.
4. Kagan AR. The lost connection: the patient-doctor relationship. Am J Clin Oncol 2007 Jun;30(3):325-326. [doi: 10.1097/01.coc.0000258097.50836.20] [Medline: 17551314]
5. Weiner M, Gress T, Thiemann DR, Jenckes M, Reel SL, Mandell SF, et al. Contrasting views of physicians and nurses about an inpatient computer-based provider order-entry system. J Am Med Inform Assoc 1999;6(3):234-244. [Medline: 10332656]
6. Margalit RS, Roter D, Dunevant MA, Larson S, Reis S. Electronic medical record use and physician-patient communication: an observational study of Israeli primary care encounters. Patient Educ Couns 2006 Apr;61(1):134-141. [doi: 10.1016/j.pec.2005.03.004] [Medline: 16533682]
7. Booth N, Robinson P, Kohannejad J. Identification of high-quality consultation practice in primary care: the effects of computer use on doctor-patient rapport. Inform Prim Care 2004;12(2):75-83. [Medline: 15319059]
8. DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care--a national survey of physicians. N Engl J Med 2008 Jul 3;359(1):50-60 [FREE Full text] [doi: 10.1056/NEJMsa0802005] [Medline: 18565855]
9. Arar NH, Wen L, McGrath J, Steinbach R, Pugh JA. Communicating about medications during primary care outpatient visits: the role of electronic medical records. Inform Prim Care 2005;13(1):13-22. [Medline: 15949171]
10. Simon SR, Kaushal R, Cleary PD, Jenter CA, Volk LA, Orav EJ, et al. Physicians and electronic health records: a statewide survey. Arch Intern Med 2007 Mar 12;167(5):507-512 [FREE Full text] [doi: 10.1001/archinte.167.5.507] [Medline: 17353500]
11. Hsu J, Huang J, Fung V, Robertson N, Jimison H, Frankel R. Health information technology and physician-patient interactions: impact of computers on communication during outpatient primary care visits. J Am Med Inform Assoc 2005;12(4):474-480. [doi: 10.1197/jamia.M1741] [Medline: 15802484]
12. Kemper AR, Uren RL, Clark SJ. Adoption of electronic health records in primary care pediatric practices. Pediatrics 2006 Jul;118(1):e20-e24 [FREE Full text] [doi: 10.1542/peds.2005-3000] [Medline: 16818534]

Abbreviations

EMR: electronic medical record
IOM: Institute of Medicine
IRB: institutional review board
HIT: health information technology
SAS: Statistical Analysis Software
PPD: purified protein derivative

Edited by G Eysenbach; submitted 20.03.10; peer-reviewed by S Garde, J Ralston; comments to author 25.05.10; revised version received 28.01.11; accepted 14.03.11; published 27.05.11.

Please cite as:
Rosen P, Spalding SJ, Hannon MJ, Boudreau RM, Kwoh CK
Parent Satisfaction With the Electronic Medical Record in an Academic Pediatric Rheumatology Practice
J Med Internet Res 2011;13(2):e40
URL: http://www.jmir.org/2011/2/e40/
doi:10.2196/jmir.1525
PMID:21622292

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org