Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 14 (2012), Issue 4   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Internet-Based Brief Personalized Feedback Intervention in a Non-Treatment-Seeking Population of Adult Heavy Drinkers: A Randomized Controlled Trial (e98)
Anders Hansen, Ulrik Becker, Anette Nielsen, Morten Grønbæk, Janne Tolstrup, Lau Thygesen. 4

Comparison of Two Internet-Based Interventions for Problem Drinkers: Randomized Controlled Trial (e107)
John Cunningham. 25

Public Health Interventions: Reaching Latino Adolescents via Short Message Service and Social Media (e99)
Amita Vyas, Megan Landry, Marisa Schnider, Angela Rojas, Susan Wood. 32

The Effects of Combining Web-Based eHealth With Telephone Nurse Case Management for Pediatric Asthma Control: A Randomized Controlled Trial (e101)
David Gustafson, Meg Wise, Abhik Bhattacharya, Alice Pulvermacher, Kathleen Shanovich, Brenda Phillips, Erik Lehman, Vernon Chinchilli, Robert Hawkins, Jee-Seon Kim. 42

Effects of Functional Interactivity on Patients' Knowledge, Empowerment, and Health Outcomes: An Experimental Model-Driven Evaluation of a Web-Based Intervention (e105)
Luca Camerini, Peter Schulz. 61

Use of Email and Telephone Prompts to Increase Self-Monitoring in a Web-Based Intervention: Randomized Controlled Trial (e96)
Mary Greaney, Kim Sprunck-Harrild, Gary Bennett, Elaine Puleo, Jess Haines, K Viswanath, Karen Emmons. 77

Preventing Smoking Relapse via Web-Based Computer-Tailored Feedback: A Randomized Controlled Trial (e109)
Iman Elfeddali, Catherine Bolman, Math Candel, Reinout Wiers, Hein de Vries. 88

Improving Diabetes Care for Young People With Type 1 Diabetes Through Visual Learning on Mobile Phones: Mixed-Methods Study (e111)
Dag Frøisland, Eirik Årsand, Finn Skårderud. 114

Online Advertising to Reach and Recruit Latino Smokers to an Internet Cessation Program: Impact and Costs (e116)
Amanda Graham, Ye Fang, Jose Moreno, Shawn Streiff, Jorge Villegas, Ricardo Muñoz, Kenneth Tercyak, Jeanne Mandelblatt, Donna Vallone.
1 6 1

Socioeconomic and Geographic Disparities in Health Information Seeking and Internet Use in Puerto Rico (e104)
Lila Finney Rutten, Bradford Hesse, Richard Moser, Ana Ortiz Martinez, Julie Kornfeld, Robin Vanderpool, Margaret Byrne, Guillermo Tortolero Luna. 174

Computer Skills and Internet Use in Adults Aged 50-74 Years: Influence of Hearing Difficulties (e113)
Helen Henshaw, Daniel Clark, Sujin Kang, Melanie Ferguson. 186

Speech Perception Benefits of Internet Versus Conventional Telephony for Hearing-Impaired Individuals (e102)
Georgios Mantokoudis, Patrick Dubach, Flurin Pfiffner, Martin Kompis, Marco Caversaccio, Pascal Senn. 200

Animated Graphics for Comparing Two Risks: A Cautionary Tale (e106)
Brian Zikmund-Fisher, Holly Witteman, Andrea Fuhrel-Forbis, Nicole Exe, Valerie Kahn, Mark Dickson. 212

Adoption, Acceptability, and Accuracy of an Online Clinical Trial Matching Website for Breast Cancer (e97)
Ellyn Cohen, Jeff Belkora, Joanne Tyler, Joan Schreiner, Mary Deering, Lakshmi Grama, Brenda Duggan, Julie Illi, Julia Pederson, Aprajita Anand, Alexandra Teng, Erin McCreary, Dan Moore, Debu Tripathy, Michael Hogarth, Morton Lieberman, John Park, Laura Esserman. 239

Impact of the Presence of Medical Equipment in Images on Viewers' Perceptions of the Trustworthiness of an Individual On-Screen (e100)
Moyez Jiwa, Stephan Millett, Xingqiong Meng, Vivien Hewitt. 253

P2P Watch: Personal Health Information Detection in Peer-to-Peer File-Sharing Networks (e95)
Marina Sokolova, Khaled El Emam, Luk Arbuckle, Emilio Neri, Sean Rose, Elizabeth Jonker. 265

Tutorial

Design Considerations in Developing a Text Messaging Program Aimed at Smoking Cessation (e103)
Michele Ybarra, Jodi Holtrop, A Ba ci Bosi, Salih Emri. 104

Viewpoints

Making Sense of Mobile Health Data: An Open Architecture to Improve Individual- and Population-Level Health (e112)
Connie Chen, David Haddad, Joshua Selsky, Julia Hoffman, Richard Kravitz, Deborah Estrin, Ida Sim. 127

Munchausen by Internet: Current Research and Future Directions (e115)
Andy Pulman, Jacqui Taylor. 150

Reviews

The Therapeutic Relationship in E-Therapy for Mental Health: A Systematic Review (e110)
Madalina Sucala, Julie Schnur, Michael Constantino, Sarah Miller, Emily Brackman, Guy Montgomery. 137

Are Personal Health Records Safe? A Review of Free Web-Accessible Personal Health Record Privacy Policies (e114)
Inmaculada Carrión Señor, José Fernández-Alemán, Ambrosio Toval. 225

Corrigenda and Addenda

Correction: Development and Validation of Filters for the Retrieval of Studies of Clinical Examination From Medline (e108)
Robert Badgett, Nader Shaikh, Mini Pi, Nancy Wilczynski, K. McKibbon, Andrea Ketchum, R. Haynes. 278

Original Paper

Internet-Based Brief Personalized Feedback Intervention in a Non-Treatment-Seeking Population of Adult Heavy Drinkers: A Randomized Controlled Trial

Anders Blædel Gottlieb Hansen[1], PhD; Ulrik Becker[1,2], MD, DMSc; Anette Søgaard Nielsen[1,3], PhD; Morten Grønbæk[1], MD, DMSc; Janne Schurmann Tolstrup[1], PhD, DMSc; Lau Caspar Thygesen[1], PhD

[1]National Institute of Public Health, Faculty of Health Sciences, University of Southern Denmark, Copenhagen, Denmark
[2]Department of Medical Gastroenterology, Copenhagen University Hospital, Capital Region, Hvidovre, Denmark
[3]The Treatmentcenter, Odense, Denmark

Corresponding Author:
Anders Blædel Gottlieb Hansen, PhD
National Institute of Public Health
Faculty of Health Sciences
University of Southern Denmark
Øster Farimagsgade 5 A
2nd floor
Copenhagen, 1353
Denmark
Phone: 45 6550 7777 ext 7729
Fax: 45 3920 8010
Email: ago@niph.dk

Abstract

Background: Internet-based interventions for heavy drinkers show promising results, but existing research is characterized by few studies in nonstudent adult populations and few comparisons with appropriate control groups.

Objective: To test whether a fully automated Internet-based brief personalized feedback intervention and a fully automated Internet-based personalized brief advice intervention in a non-treatment-seeking population of heavy drinkers would result in a reduced alcohol intake.

Methods: We conducted a 3-arm parallel randomized controlled trial in a general population-based sample of heavy drinkers. The 54,157 participants (median age of 58 years) were screened for heavy drinking. Of the 3418 participants who had a weekly alcohol consumption above 14 drinks for women and 21 drinks for men, 1380 (619 women) consented to take part in the trial and were randomly assigned to an Internet-based brief personalized feedback intervention group (normative feedback, n = 476), an Internet-based personalized brief advice intervention group (n = 450), or a nonintervention control group (n = 454). Follow-up after 6 and 12 months included 871 and 1064 participants, respectively, of all groups combined. The outcome measure was self-reported weekly alcohol consumption. We analyzed the data according to the intention-to-treat principle. To examine changes over time and to account for the multiple time measurements, we used a multilevel linear mixed model. To take attrition into account, we used multiple imputation to address missing data.

Results: The intervention effect of the Internet-based brief personalized feedback intervention, determined as the mean additional difference in changes in alcohol consumption in the Internet-based brief personalized feedback intervention compared with the control group, was –1.8 drinks/week after 6 months and –1.4 drinks/week after 12 months; these effects were nonsignificant (95% confidence interval –4.0 to 0.3 at 6 months, –3.4 to 0.6 at 12 months). The intervention effect of the Internet-based personalized brief advice intervention was –0.5 drinks/week after 6 months and –1.2 drinks/week after 12 months; these effects were nonsignificant (95% confidence interval –2.7 to 1.6 at 6 months, –3.3 to 0.9 at 12 months).

Conclusions: In this randomized controlled trial we found no evidence that an Internet-based brief personalized feedback intervention was effective in reducing drinking in an adult population of heavy drinkers.

Trial registration: ClinicalTrials.gov NCT00751985; http://clinicaltrials.gov/ct2/show/NCT00751985 (Archived by WebCite at http://www.webcitation.org/68WCRLyaP)

(J Med Internet Res 2012;14(4):e98) doi:10.2196/jmir.1883

KEYWORDS

Internet-based personalized feedback; normative feedback; alcohol; heavy drinking; adult; Internet-based personalized brief advice, brief intervention

Introduction

Heavy alcohol intake increases the risk of numerous chronic diseases, injuries, disabilities, and death [1]. Many drinkers who do not meet diagnostic criteria for alcohol dependence or harmful drinking nonetheless consume alcohol at a level or in a pattern that increases the risk of negative health and social consequences [2]. In Denmark, it has been estimated that 20% of the adult population are heavy drinkers [3]; hence, the need to detect and intervene in the early stages of heavy drinking is obvious. Face-to-face brief interventions, which are intended as an early intervention for non-treatment-seeking, non-alcohol-dependent drinkers, have proven to be effective and have been advocated as a strategy to curb heavy drinking [4]. However, problems with feasibility and barriers to implementation have been encountered, such as a limited number of professionals who administer them and the difficulty of contacting heavy drinkers [5,6]. As a consequence, there is a gap between need and access to interventions to reduce alcohol intake. It has been estimated that as many as 80% of problem drinkers do not receive help due to a combination of missed screening opportunities and the stigma associated with alcohol treatment [7,8]. Delivering brief interventions over the Internet may overcome some of these barriers, and Internet-based interventions can reach individuals who are otherwise unwilling or not motivated to seek help [9,10]. The increasing access to the Internet in the population, currently 58% for Europe, 78% for North America, and 30% worldwide [11], and the well-documented demand for Internet-based interventions in the general public [12] mean that, if delivered broadly, Internet-based interventions could have potentially major public health implications, the main argument being that Internet-based interventions combine the scalability of a public health intervention with the capacity to deliver a personalized approach [13].

Recent systematic reviews concluded that Internet-based interventions were more effective than minimally active comparator groups at reducing alcohol intake, with a mean difference of 2–3 drinks per week [14], and found small to medium effect sizes [15-17]. However, several methodological flaws in the reviewed trials caused the authors to state that the ability to generalize about the efficacy and utility of Internet-based interventions for alcohol use is impeded, and hence it is not possible to interpret the evidence with any degree of certainty [14,17]. Unresolved questions remain, such as the need to establish which components of Internet-based interventions are effective [18]. The provision of a personalized feedback intervention that compares one’s own drinking with peers’ actual drinking has been found to increase motivation to change drinking by making individuals aware of discrepancies between their personal alcohol consumption and social norms. This approach originates in self-regulation theory and builds on the assumption that change is triggered by creating an awareness of a perceived discrepancy. Therefore, if heavy drinkers find no such discrepancy, they would view their personal behavior as being normal rather than abnormal. This personal tailored approach, also termed normative feedback, which tries to create behavioral change by targeting normative misperceptions, is assumed to be more effective at reducing drinking than is delivering standardized feedback in the form of self-help material [16,19-22].

Some drawbacks of personalized feedback intervention studies are small sample sizes, short-term follow-up, the existence of few studies in nonstudent adult populations, few comparisons with appropriate control groups, and high rates of attrition [14,16,23]. In this study, we addressed these shortcomings by comparing a single-session, Internet-based, brief personalized feedback intervention with an Internet-based, personalized, brief advice intervention against a pure control group in the context of the Danish Health Examination Survey [24]. We sought to determine whether these single-session interventions would result in a decrease in alcohol use in a non-treatment-seeking population of adult heavy drinkers. As a secondary aim, we also sought to determine whether the Internet-based brief personalized feedback intervention would have any gender-specific effects, especially since differential effectiveness between genders in non-Internet brief interventions remains ambiguous [4,25,26] and because few Internet-based intervention studies present data separately for men and women.

Methods

Setting

The Danish Health Examination Survey was carried out by the National Institute of Public Health, University of Southern Denmark, in 13 municipalities in 2007/2008. The Danish Health Examination Survey focused primarily on diet, smoking, alcohol, and physical activity and consisted of an Internet-based questionnaire and a health examination. In this study, we used data from the Internet-based questionnaire from 12 of the 13 municipalities. All adult inhabitants in 12 municipalities were invited to complete the Internet-based questionnaire (n = 401,607). The sample was drawn from the adult Danish population (18 years or older) using the Danish Civil Registration System, which contains information on gender, age, address, citizenship, and marital status for each individual (each Danish resident has a unique personal identification number) [27]. The questionnaire was fully or partially completed by 54,157 participants, corresponding to 13.49% of all adults in the 12 municipalities [24].

Recruitment

Recruitment for the study began in September 2008. Follow-up started in February 2009 and ended in February 2010.

Study Design

The study was a 3-arm randomized controlled trial.

Participants

Invitees to the Danish Health Examination Survey received a letter inviting them to participate by completing an Internet-based questionnaire containing questions on their sociodemographic characteristics, self-reported health status, living conditions, and health behavior including alcohol consumption. The baseline questionnaire was completed at the respondent's home. In 7 of the 12 municipalities, the questionnaire was supplemented with questions to test the willingness of respondents to change their health behavior in four domains: weight, diet, smoking, and alcohol (n = 33,554 completed these questions). The alcohol questions were beverage specific (beer, wine, fortified wine, or spirits) and asked for amount consumed each day during a typical week. Additionally, the Alcohol Use Disorders Identification Test questions 3–10 were included (timeframe: preceding 12 months) [28]. Respondents who had provided an email address (75% of the population) and whose weekly alcohol consumption was above the recommended maximum drinking limit, as stated by the Danish National Board of Health (14 drinks = 168 g of alcohol for women, 21 drinks = 252 g for men), were eligible for the study. One standard drink corresponds to 12 g of pure alcohol [29]. Heavy drinking was defined as 168 g or more of alcohol/week for women and 252 g/week or more for men. Binge drinking was defined as drinking 5 or more drinks on a single occasion both for men and women.

Interventions

The Internet-based brief personalized feedback intervention was a fully automated, single-session intervention; it was displayed in a single screenshot and addressed to the participant by name. It consisted of a summary of the participant's weekly consumption, a comparison of the weekly consumption with the maximum drinking limit, and a graphical comparison of the participant's consumption with the average level in the municipality (gender specific). The Internet-based brief personalized feedback intervention also included information about the risks to health and social relationships linked to heavy drinking, as well as links for further self-help material and a local alcohol treatment facility (see Figure 1).

The Internet-based personalized brief advice intervention was a fully automated single-session intervention and was displayed in a single screenshot and addressed to the participant by name. It informed the participant that his or her alcohol consumption exceeded the recommended maximum drinking limit, followed by information about the health and social risks associated with heavy drinking, as well as links for further standardized self-help material and a local alcohol treatment facility (see Figure 2).What distinguishes the Internet-based brief personalized feedback intervention from the Internet-based personalized brief advice intervention is the normative component and the summary of the participant's weekly alcohol consumption. Common to the two interventions is the information on the adverse effects of heavy drinking, advice to cut down, and links for further material (see Multimedia Appendix 1).

Participants in the control group received a single screenshot that explained that being randomly selected for the control group implied no intervention and follow-up after 6 and 12 months (see Figure 3).

Individuals who consumed less than the maximum drinking limit did not receive any feedback or interventions.

Procedure

After completing the Internet-based Danish Health Examination Survey questionnaire, invitees were automatically screened and heavy drinkers were identified. Heavy drinkers received an email inviting them to participate in the intervention study. By clicking on a link in the email, invitees were directed to a secure website where, after entering a username and personal access code (provided in the Danish Health Examination Survey invitation letter), they were directed to another website that explained the study (see Multimedia Appendix 1). After providing their online consent, participants were automatically randomly assigned and directed to a new personalized website that presented one of the interventions or control, which was displayed immediately on the screen.

Randomization

Eligible persons were randomly assigned and enrolled into the Internet-based brief personalized feedback intervention, the Internet-based personalized brief advice intervention, or the control group by the Web server software, which was implemented by a technician who was not involved in the recruitment process. Blinding was not feasible. Participants did not know which of the two interventions was the intervention of interest. Prior to randomization, all three groups were informed about the purpose of the study and the nature of the control group (see Multimedia Appendix 1).

Outcome Measure and Follow-up

There was one planned primary analysis: overall reduction in alcohol use; and one post hoc secondary analysis: gender-specific reductions in alcohol use. The outcome measure was specified a priori and was based on self-reported drinking each day during a typical week and included beverage-specific questions (beer, wine, fortified wine, and spirits). The follow-up at 6 and 12 months contained the same alcohol items included in the baseline questionnaire and was conducted using an Internet-based questionnaire that participants accessed using a link provided in an email. The follow-up at 12 month was also supplemented by a letter containing the questionnaire, which the participants could answer if they did not respond to the email.

Power Estimates

The sample size was calculated based on a meta-analysis of Internet-based interventions, in which a mean difference of 2–3 drinks (26 g of alcohol) per week was found [14], and based on non-Internet-based interventions meta-analyses, where a 12% to 15% reduction in the previous week's alcohol consumption was found, relative to a baseline consumption of approximately 300 g/week [4,30]. We anticipated a decrease in alcohol consumption of approximately 15% for the Internet-based brief personalized feedback intervention, 10% for the Internet-based personalized brief advice intervention, and 5% for the control group. Decreases in control groups in Internet-based interventions and non-Internet-based interventions have been substantial. However, this decrease has not been quantified in meta-analyses due to the highly variable content of control groups in both Internet-based interventions and face-to-face brief interventions [31]. Assuming that the standard deviation was equal to a third of the expected baseline consumption, we estimated that 182 participants in each group would be needed to give the trial 80% power to detect an effect of the Internet-based brief personalized feedback intervention versus the control group of this size at the 5% level of significance. To detect an effect of the Internet-based personalized brief advice intervention versus the control group, we estimated that 726 participants in each group would be needed. A target sample size of more than 1200 enrollees was deemed necessary to allow for substantial attrition.

Statistical Analysis

The primary and secondary analyses were based on the intention-to-treat principle and concerned the mean difference in changes in alcohol consumption between the two intervention groups and the control group.

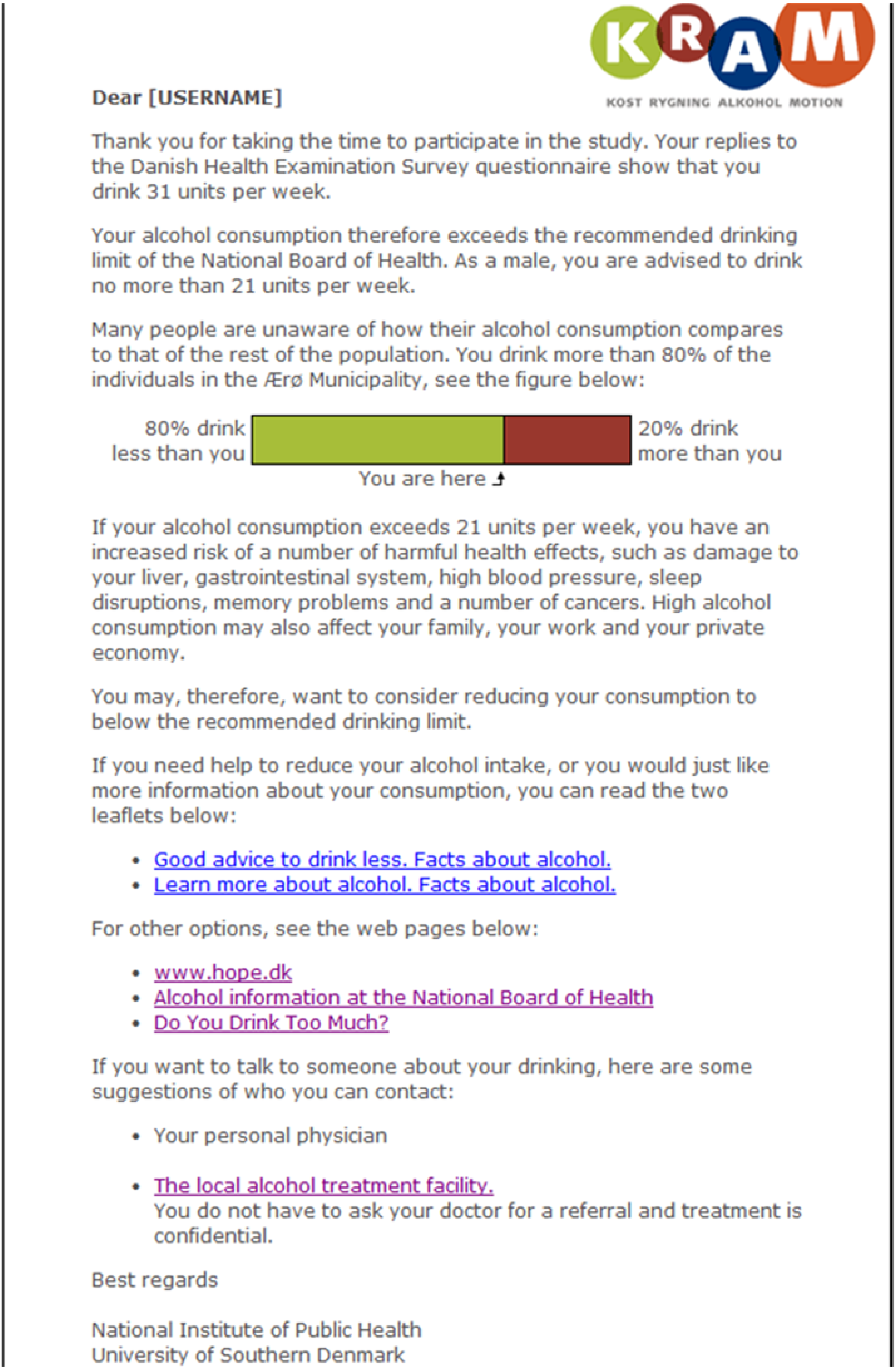
We carried out analyses using Stata version 11.2 (StataCorp LP, College Station, TX, USA). Quantitative variables were described by the mean and standard deviation, by the median and its interquartile range, or by its 95% confidence interval (CI). In all tests, we chose $P < .05$ as the level of significance.

The residuals were approximately normally distributed. Hence, to examine changes over time and to account for the multiple time measurements, we analyzed data by using a multilevel mixed model, using the *xtmixed* procedure. The model examined fixed effects for alcohol consumption, group, gender, and month and a random intercept to account for clustering within each participant. The model also included an interaction term between intervention group and month, allowing for differences in the intervention effect between follow-up assessments [32]. The fixed effect of most interest was the month × group interaction effect, which indicated the difference between intervention groups and the control group as a change in alcohol consumption over time.

The Kruskal-Wallis nonparametric test for continuous variables was used to compare the three groups in the secondary analyses.

For the loss to follow-up analysis, we used the chi-square test and the Kruskal-Wallis nonparametric test for continuous variables to compare the baseline characteristics between those followed up and those lost to follow-up. In a preplanned analysis, we used multiple imputation to take attrition into account for participants who did not complete the 6- or 12-month follow-up. Multiple imputation allows for the uncertainty about the missing data by creating several different plausible imputed data sets and appropriately combining results obtained from each (we generated 20 data sets), which often provides a more reliable approach than complete case analysis in the presence of missing data [23,33]. For this we used the *mi impute mvn* procedure, which uses multivariate normal regression for continuous data and assumes that data are missing at random [34]. As sensitivity analyses, we also report results from (1) an analysis of all available results without the imputation of missing data (completers-only analysis), and (2) an analysis with simple imputation (last observation carried forward) assuming that nonresponders had no change in their alcohol consumption.

Figure 1. Screenshot of the Internet-based brief personalized feedback intervention.



Dear [USERNAME]

Thank you for taking the time to participate in the study. Your replies to the Danish Health Examination Survey questionnaire show that you drink 31 units per week.

Your alcohol consumption therefore exceeds the recommended drinking limit of the National Board of Health. As a male, you are advised to drink no more than 21 units per week.

Many people are unaware of how their alcohol consumption compares to that of the rest of the population. You drink more than 80% of the individuals in the Ærø Municipality, see the figure below:

If your alcohol consumption exceeds 21 units per week, you have an increased risk of a number of harmful health effects, such as damage to your liver, gastrointestinal system, high blood pressure, sleep disruptions, memory problems and a number of cancers. High alcohol consumption may also affect your family, your work and your private economy.

You may, therefore, want to consider reducing your consumption to below the recommended drinking limit.

If you need help to reduce your alcohol intake, or you would just like more information about your consumption, you can read the two leaflets below:

- Good advice to drink less. Facts about alcohol.
- Learn more about alcohol. Facts about alcohol.

For other options, see the web pages below:

- www.hope.dk
- Alcohol information at the National Board of Health
- Do You Drink Too Much?

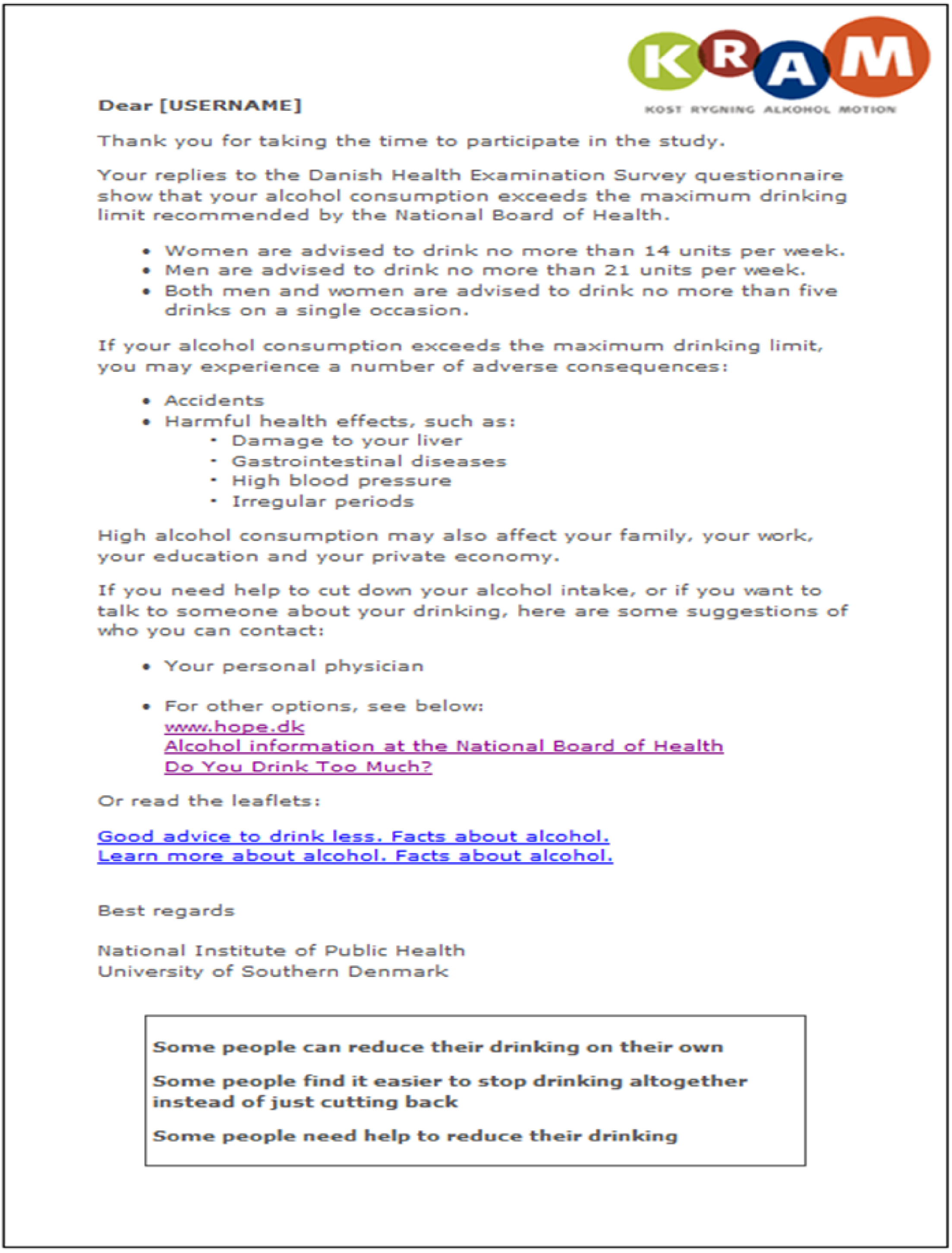
If you want to talk to someone about your drinking, here are some suggestions of who you can contact:

- Your personal physician
- The local alcohol treatment facility.
  You do not have to ask your doctor for a referral and treatment is confidential.

Best regards

National Institute of Public Health
University of Southern Denmark

Figure 2. Screenshot of the Internet-based personalized brief advice intervention.



Dear [USERNAME]

Thank you for taking the time to participate in the study.

Your replies to the Danish Health Examination Survey questionnaire show that your alcohol consumption exceeds the maximum drinking limit recommended by the National Board of Health.

- Women are advised to drink no more than 14 units per week.
- Men are advised to drink no more than 21 units per week.
- Both men and women are advised to drink no more than five drinks on a single occasion.

If your alcohol consumption exceeds the maximum drinking limit, you may experience a number of adverse consequences:

- Accidents
- Harmful health effects, such as:
  - Damage to your liver
  - Gastrointestinal diseases
  - High blood pressure
  - Irregular periods

High alcohol consumption may also affect your family, your work, your education and your private economy.

If you need help to cut down your alcohol intake, or if you want to talk to someone about your drinking, here are some suggestions of who you can contact:

- Your personal physician

- For other options, see below:
  www.hope.dk
  Alcohol information at the National Board of Health
  Do You Drink Too Much?

Or read the leaflets:

Good advice to drink less. Facts about alcohol.
Learn more about alcohol. Facts about alcohol.

Best regards

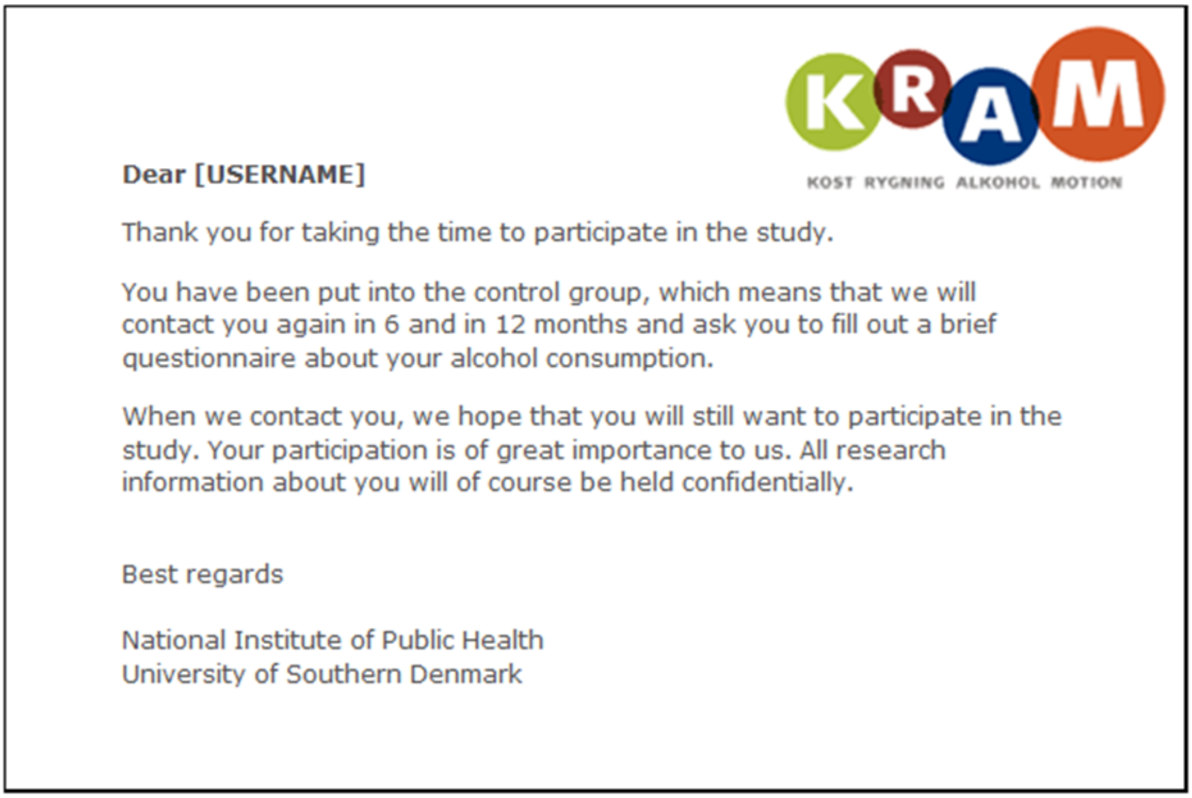
National Institute of Public Health
University of Southern Denmark

Some people can reduce their drinking on their own

Some people find it easier to stop drinking altogether instead of just cutting back

Some people need help to reduce their drinking

Figure 3. Screenshot of the control group condition.



Dear [USERNAME]

Thank you for taking the time to participate in the study.

You have been put into the control group, which means that we will contact you again in 6 and in 12 months and ask you to fill out a brief questionnaire about your alcohol consumption.

When we contact you, we hope that you will still want to participate in the study. Your participation is of great importance to us. All research information about you will of course be held confidentially.

Best regards

National Institute of Public Health
University of Southern Denmark

Results

Participant Flow

Of the 54,157 screened individuals, 3418 (6.31%) were heavy drinkers. Of these, 785 (23.0%) declined participation, and 1215 (35.6%) did not respond to the invitation email. In total, 1380 (40.37%) individuals accepted participation and were randomly assigned into the Internet-based brief personalized feedback intervention (n = 476), the Internet-based personalized brief advice intervention (n = 450), or the control group (n = 454). The 6-month follow-up was completed by 871 (63.1% of enrolled participants) individuals and the 12-month follow-up was completed by 1064 (77.10%) individuals (Figure 4).

Baseline Data

At baseline, men consumed a mean of 32 drinks/week and women 21 drinks/week. During the previous year, 49.9% (n = 380) of the men had been binge drinking once a week or more often, while among women the corresponding figure was 25.5% (n = 158) (Table 1). At baseline, 384 (46.2%) individuals answered “yes” or “yes, maybe” to the question “Do you want to cut down on your drinking?”, 319 (38.4%) answered “no”, and 128 (15.4%) did not respond (831 individuals received this question). The median age was 58 years, 55.1% (n = 761) were men, 51.7% (n = 714) had more than 15 years of education, 53.1% (n = 733) were employed, and 69.6% (n = 961) were married or cohabiting. Among the participants, 10% (n = 139) were daily smokers and 12% (n = 161) were heavy smokers (more than 15 cigarettes a day). There were no significant differences between randomized groups for any baseline characteristic.

Table 1. Baseline characteristics of participants randomly assigned to Internet-based brief personalized feedback intervention (PFI), Internet-based personalized brief advice intervention (PBA), or control group in The Danish Health Examination Survey 2008.

Characteristic	PFI	PBA	Control
Men			
No.	271	246	244
Age (years), median (IQR[a])	61 (50–66)	59 (49–65)	60 (51–65)
Alcohol intake (drinks/week)[b], mean (SD)	32.8 (16.9)	32.7 (14.0)	31.3 (10.3)
Binge drinking, n (%)[c]	137 (50.5%)	118 (48.0%)	125 (51.2%)
Education level (years), n (%)			
<10	11 (4%)	19 (8%)	10 (4%)
10–12	65 (24%)	55 (22%)	59 (24%)
13–14	45 (17%)	50 (20%)	55 (23%)
15+	149 (54.9%)	117 (47.5%)	118 (48.3%)
Employed, n (%)	146 (53.8%)	121 (49.1%)	144 (59.0%)
Smoking, n (%)			
Daily	31 (11%)	29 (12%)	23 (9%)
Heavy[d]	40 (15%)	29 (12%)	26 (11%)
Married or cohabiting, n (%)	198 (73.1%)	172 (69.9%)	185 (75.8%)
Motivated to reduce alcohol use, n (%)[e]			
Yes	19 (7%)	26 (11%)	22 (9%)
Yes, maybe	53 (20%)	41 (17%)	47 (19%)
No	56 (21%)	56 (23%)	57 (23%)
Women			
No.	205	204	210
Age (years), median (IQR)	54 (41–62)	56 (46–63)	56 (44–62)
Alcohol intake (drinks/week), mean (SD)	20.9 (7.0)	21.5 (9.0)	21.3 (8.2)
Binge drinking, n (%)	50 (24%)	55 (27%)	53 (25%)
Education level (years), n (%)			
<10	10 (5%)	10 (5%)	12 (6%)
10–12	55 (27%)	41 (20%)	45 (21%)
13–14	32 (16%)	39 (19%)	41 (20%)
15+	107 (52.2%)	112 (54.9%)	111 (52.9%)
Employed, n (%)	107 (52.2%)	116 (56.9%)	99 (47%)
Smoking, n (%)			
Daily	20 (8%)	19 (9%)	17 (8%)
Heavy	26 (13%)	18 (9%)	22 (10%)
Married or cohabiting, n (%)	129 (62.9%)	135 (66.2%)	142 (67.6%)
Motivated to reduce alcohol use, n (%)			
Yes	24 (12%)	21 (10%)	23 (11%)
Yes, maybe	33 (16%)	32 (16%)	43 (20%)
No	49 (24%)	48 (24%)	53 (25%)

[a] Interquartile range.

[b] Number of standard drinks in a typical week.

[c] Drinking 5 or more drinks per occasion at least once a week.

[d] Smoking more than 15 cigarettes a day.

[e] Numbers do not sum to 100% due to missing data.

Loss to Follow-up Analysis

We compared participants lost to follow-up (n = 509, 37% at 6 months and n = 316, 23% at 12 months) by intervention group with those who participated in the follow-up in terms of baseline characteristics.

Participants lost to follow-up were significantly more likely to be heavy smokers, less likely to have a high level of education (15+ years), and more likely to have a low level of education (10–12 years). Furthermore, participants lost to follow-up were more likely to be unmotivated to cut down their drinking (Table 2 and Table 3).

Table 2. Comparison of characteristics of participants randomly assigned to Internet-based brief personalized feedback intervention (PFI), Internet-based personalized brief advice intervention (PBA), or control group at baseline between those followed up after 6 months and those lost at 6-month follow-up.[a]

Characteristic	Followed up after 6 months			Lost to follow-up at 6 months					
	PFI (n = 288)	PBA (n = 280)	Control (n = 303)	PFI (n = 188)	*P* value[b]	PBA (n = 170)	*P* value[c]	Control (n = 151)	*P* value[d]
Men, n (%)	161 (55.9%)	155 (55.4%)	165 (54.5%)	110 (58.5%)	.57	91 (54%)	.71	79 (52%)	.67
Women, n (%)	127 (44.1%)	125 (44.6%)	138 (45.5%)	78 (42%)		79 (47%)		72 (48%)	
Age (years), median (IQR)[e]	58 (46–65)	58 (48–64)	60 (48–64)	58 (47–65)	.46	57 (47–65)	.89	56 (46–63)	.04
Alcohol intake, mean (SD)[f]	28.0 (16.7)	27.9 (12.9)	26.1 (9.6)	27.1 (11.1)	.85	27.0 (13.7)	.43	27.9 (12.9)	.35
Binge drinking, n (%)[g]	114 (39.6%)	117 (41.8%)	119 (39.3%)	73 (39%)	.85	56 (33%)	.05	59 (39%)	.98
Education level (years), n (%) [h]					.03		.28		.65
<10	13 (5%)	17 (6%)	15 (5%)	8 (4%)		12 (7%)		7 (5%)	
10–12	66 (23%)	52 (19%)	64 (21%)	54 (29%)		44 (26%)		40 (26%)	
13–14	38 (13%)	56 (20%)	66 (22%)	39 (21%)		33 (19%)		30 (20%)	
15+	170 (59.0%)	150 (53.6%)	156 (51.5%)	86 (46%)		79 (47%)		73 (48%)	
Employed, n (%)	155 (53.8%)	152 (54.3%)	157 (51.8%)	98 (52%)	.80	85 (50%)	.37	86 (57%)	.30
Smoking, n (%)					.02		.01		.12
Daily	31 (11%)	38 (14%)	28 (9%)	20 (11%)		10 (6%)		12 (8%)	
Heavy[i]	29 (10%)	21 (8%)	25 (8%)	37 (20%)		26 (15%)		23 (15%)	
Married or cohabiting, n (%)	203 (70.5%)	190 (67.9%)	219 (72.3%)	124 (66.0%)	.04	117 (68.8%)	.12	108 (71.5%)	.28
Motivated to reduce alcohol use, n (%) [h]					.02		.56		.47
"Yes" or "yes, maybe"	87 (56%)	84 (47%)	99 (48%)	42 (38%)		36 (38%)		36 (41%)	
"No"	55 (36%)	69 (39%)	76 (37%)	50 (45%)		35 (37%)		34 (39%)	

[a] *P* values for categorical variables by chi-square test and for continuous variables by Kruskal-Wallis test.

[b] Participants in the PFI group who participated in follow-up versus those lost to follow-up.

[c] Participants in the PBA group who participated in follow-up versus those lost to follow-up.

[d] Participants in the control group who participated in follow-up versus those lost to follow-up.

[e] Interquartile range.

[f] Number of standard drinks in a typical week.

[g] Drinking 5 or more drinks per occasion at least once a week.

[h] Numbers do not sum to 100% due to missing data.

[i] Smoking more than 15 cigarettes a day.

Table 3. Comparison of characteristics of participants randomly assigned to Internet-based brief personalized feedback intervention (PFI), Internet-based personalized brief advice intervention (PBA), or control group at baseline between those followed up after 12 months and those lost at 12-month follow-up.[a]

Characteristic	Followed up after 12 months			Lost to follow-up 12 months					
	PFI (n = 365)	PBA (n = 341)	Control (n = 358)	PFI (n = 111)	*P* value[b]	PBA (n = 109)	*P* value[c]	Control (n = 96)	*P* value[d]
Men, n (%)	209 (57.3%)	193 (56.6%)	196 (54.8%)	62 (56%)	.79	53 (49%)	.15	48 (50%)	.41
Women, n (%)	156 (42.7%)	148 (43.4%)	162 (45.3%)	49 (44%)		56 (51%)		48 (50%)	
Age (years), median (IQR)[e]	58 (47–65)	58 (48–65)	60 (49–65)	57 (47–64)	.69	55 (46–63)	.11	54 (44–61)	<.01
Alcohol intake, mean (SD)[f]	27.7 (15.6)	27.5 (13.5)	26.4 (9.9)	27.6 (11.5)	.72	28.0 (12.3)	.47	27.9 (12.8)	.80
Binge drinking, n (%)[g]	139 (38.1%)	133 (39.0%)	131 (36.6%)	48 (43%)	.34	40 (37%)	.70	47 (49%)	.03
Education level (years), n (%) [h]					<.01		.10		.09
<10	13 (4%)	22 (6)	19 (5)	8 (7)		7 (6)		3 (3)	
10–12	87 (24%)	66 (19)	77 (22)	33 (30)		30 (28)		27 (28)	
13–14	50 (14%)	63 (18)	70 (20)	27 (24)		26 (24)		26 (27)	
15+	213 (58.4%)	184 (54.0%)	190 (53.1%)	43 (39%)		45 (41%)		39 (41%)	
Employed, n (%)	196 (53.7%)	178 (52.2%)	184 (51.4%)	57 (51%)	.65	59 (54%)	.70	59 (61%)	.07
Smoking, n (%)					.02		.23		.44
Daily	38 (10%)	39 (11%)	35 (10%)	13 (12%)		9 (8%)		5 (5%)	
Heavy[i]	41 (11%)	30 (9%)	36 (10%)	25 (23%)		17 (16%)		12 (13%)	
Married or cohabiting, n (%)	258 (70.7%)	231 (67.7%)	258 (72.1%)	69 (62%)	.19	76 (70%)	.81	69 (72%)	.70
Motivated to reduce alcohol use, n (%) [h]					.02		.07		.26
“Yes” or “yes, maybe”	108 (53%)	91 (43%)	112 (48%)	21 (33%)		29 (47%)		23 (40%)	
“No”	74 (36%)	89 (42%)	85 (36%)	31 (49%)		15 (24%)		25 (44%)	

[a] *P* values for categorical variables by chi-square test and for continuous variables by Kruskal-Wallis test.

[b] Participants in the PFI group who participated in follow-up versus those lost to follow-up.

[c] Participants in the PBA group who participated in follow-up versus those lost to follow-up.

[d] Participants in the control group who participated in follow-up versus those lost to follow-up.

[e] Interquartile range.

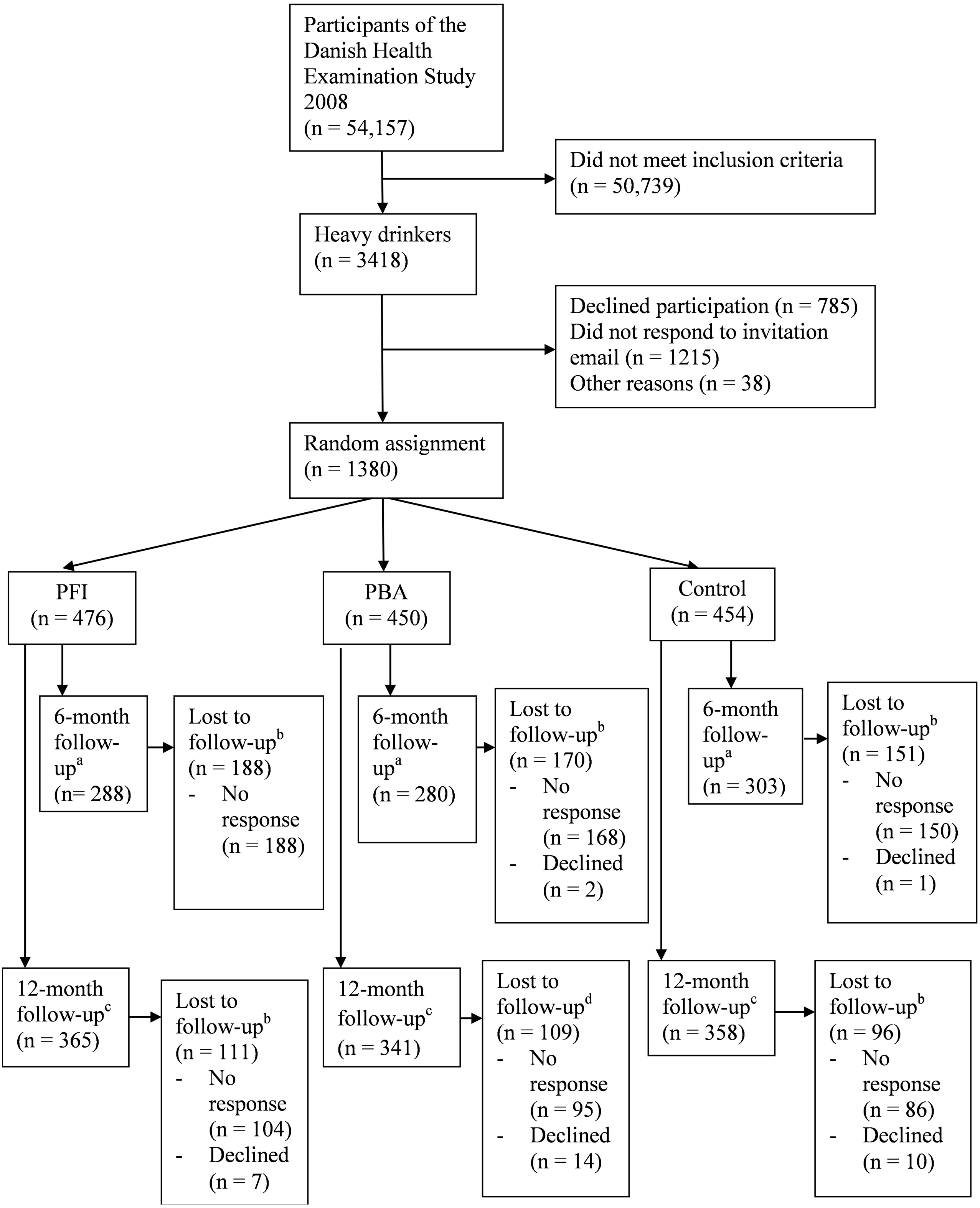
[f] Number of standard drinks in a typical week.

[g] Drinking 5 or more drinks per occasion at least once a week.

[h] Numbers do not sum to 100% due to missing data.

[i] Smoking more than 15 cigarettes a day.

Figure 4. Flow of participants through the study. PBA = Internet-based personalized brief advice intervention, PFI = Internet-based brief personalized feedback intervention. [a]Follow-up took place by means of two emails. [b]No response and declined are subsets of lost to follow-up. [c] Follow-up took place by means of two emails and two letters.



Outcomes

Table 4 and Table 5 present the intervention effects of the Internet-based brief personalized feedback intervention and the Internet-based personalized brief advice intervention, with and without imputation for missing values. The intervention effects indicate the additional difference in change in alcohol consumption in the intervention groups compared with the control group.

In the primary analysis, using multiple imputation, the intervention effects of the Internet-based brief personalized feedback intervention were –1.8 drinks/week after 6 months and –1.4 drinks/week after 12 months and were nonsignificant (95% CI –4.0 to 0.3 at 6 months, –3.4 to 0.6 at 12 months). The intervention effects of the Internet-based personalized brief advice intervention were –0.5 drinks/week after 6 months and –1.2 drinks/week after 12 months and were nonsignificant (95% CI –2.7 to 1.6 at 6 months, –3.3 to 0.9 at 12 months).

A sensitivity analysis without multiple imputation for missing values (completers-only analysis) showed that the intervention effects of the Internet-based brief personalized feedback intervention were –3.9 drinks/week after 6 months and –2.3 drinks/week after 12 months; these effects were significant (95% CI –5.8 to –2.0 at 6 months, –4.1 to –0.5 at 12 months). The Internet-based personalized brief advice intervention had no significant intervention effects.

A sensitivity analysis with simple imputation (last observation carried forward) yielded similar results to the completers-only analysis, but the differences were less pronounced.

For the control group, the overall difference between the baseline and 6-month follow-up was –4.6 drinks/week, and this difference was significant (95% CI –6.1 to –3.1). Corresponding figures for 12-month follow-up were –5.5 (95% CI –7.0 to –4.1). The two sensitivity analyses produced similar results (Table 4 and Table 5).

Table 4. Intervention effects on drinks/week based on random intercept model with and without imputation for missing values.

	With multiple imputation for missing values[a]			Without multiple imputation for missing values[b]		
	Drinks/week[c]	95% CI[d]	*P* value	Drinks/week	95% CI	*P* value
Intervention effects of the PFI [e] (month × group interaction)						
6 months	–1.8	–4.0 to 0.3	.09	–3.9	–5.8 to –2.0	<.001
12 months	–1.4	–3.4 to 0.6	.16	–2.3	–4.1 to –0.5	.01
Intervention effects of the PBA [f] (month × group interaction)						
6 months	–0.5	–2.7 to 1.6	.62	–1.4	–3.3 to 0.6	.17
12 months	–1.2	–3.3 to 0.9	.28	–1.5	–3.3 to 0.3	.10
Difference between baseline and follow-up for control group						
6 months	–4.6	–6.1 to –3.1	<.001	–4.8	–6.1 to –3.4	<.001
12 months	–5.5	–7.0 to –4.1	<.001	–5.8	–7.1 to –4.6	<.001

[a] Based on 20 imputed datasets.

[b] Based on 871 individuals after 6 months and 1064 after 12 months.

[c] Mean number of standard drinks in a typical week.

[d] Confidence interval.

[e] Internet-based brief personalized feedback intervention.

[f] Internet-based personalized brief advice intervention.

Table 5. Intervention effects on drinks/week based on random intercept model with simple imputation for missing values (last observation carried forward).[a]

	Drinks/week[b]	95% CI[c]	*P* value
Intervention effects of the PFI [d] (month × group interaction)			
6 months	–2.5	–4.0 to –1.0	<.001
12 months	–2.0	–3.4 to –0.5	.01
Intervention effects of the PBA [e] (month × group interaction)			
6 months	–0.8	–2.3 to 0.6	.27
12 months	–1.2	–2.7 to 0.3	.11
Difference between baseline and follow-up for control group			
6 months	–2.9	–4.0 to –1.9	<.001
12 months	–4.8	–5.9 to –3.8	<.001

[a] Based on 1380 individuals after 6 and 12 months.

[b] Mean number of standard drinks in a typical week.

[c] Confidence interval.

[d] Internet-based brief personalized feedback intervention.

[e] Internet-based personalized brief advice intervention.

Figure 5 and Figure 6 show secondary post hoc analyses comprising descriptive statistics for alcohol consumption, by gender and group, at baseline and at 6- and 12-month follow-ups using multiple imputation. Consumption among women decreased from a mean baseline level of 21.0 drinks/week to 16.7 drinks/week for the control group (95% CI 14.7–18.8), 16.0 drinks/week for the Internet-based brief personalized feedback intervention (95% CI 14.2–17.9), and 17.0 for the Internet-based personalized brief advice intervention (95% CI 14.6–19.5) after 6 months (Figure 5). Consumption among men decreased from a mean baseline level of 32.0 drinks/week to 26.7 drinks/week for the control group (95% CI 25.0–28.4), 25.1 drinks/week for the Internet-based brief personalized feedback intervention (95% CI 23.2–27.1), and 26.9 drinks/week for the Internet-based personalized brief advice intervention (95% CI 24.9–28.8) after 6 months (Figure 6). Figures for the 12-month follow-up were approximately similar (Figure 5 and Figure 6).

When analyzing only those who participated in follow-up (completers-only analysis), we observed significant differences between men and women. Figure 7 and Figure 8 show that the Internet-based brief personalized feedback intervention seemed to have a significant effect only on men, with a difference of 3.5 drinks/week between the Internet-based brief personalized feedback intervention and the control group at 6-month follow-up (P = .01) (Figure 8).

Figure 5. Alcohol consumption at baseline and at 6- and 12-month follow-ups for women based on multiple imputation. Error bars indicate 95% confidence interval. Drinks/week = mean number of standard drinks in a typical week, PBA = Internet-based personalized brief advice intervention, PFI = Internet-based brief personalized feedback intervention.

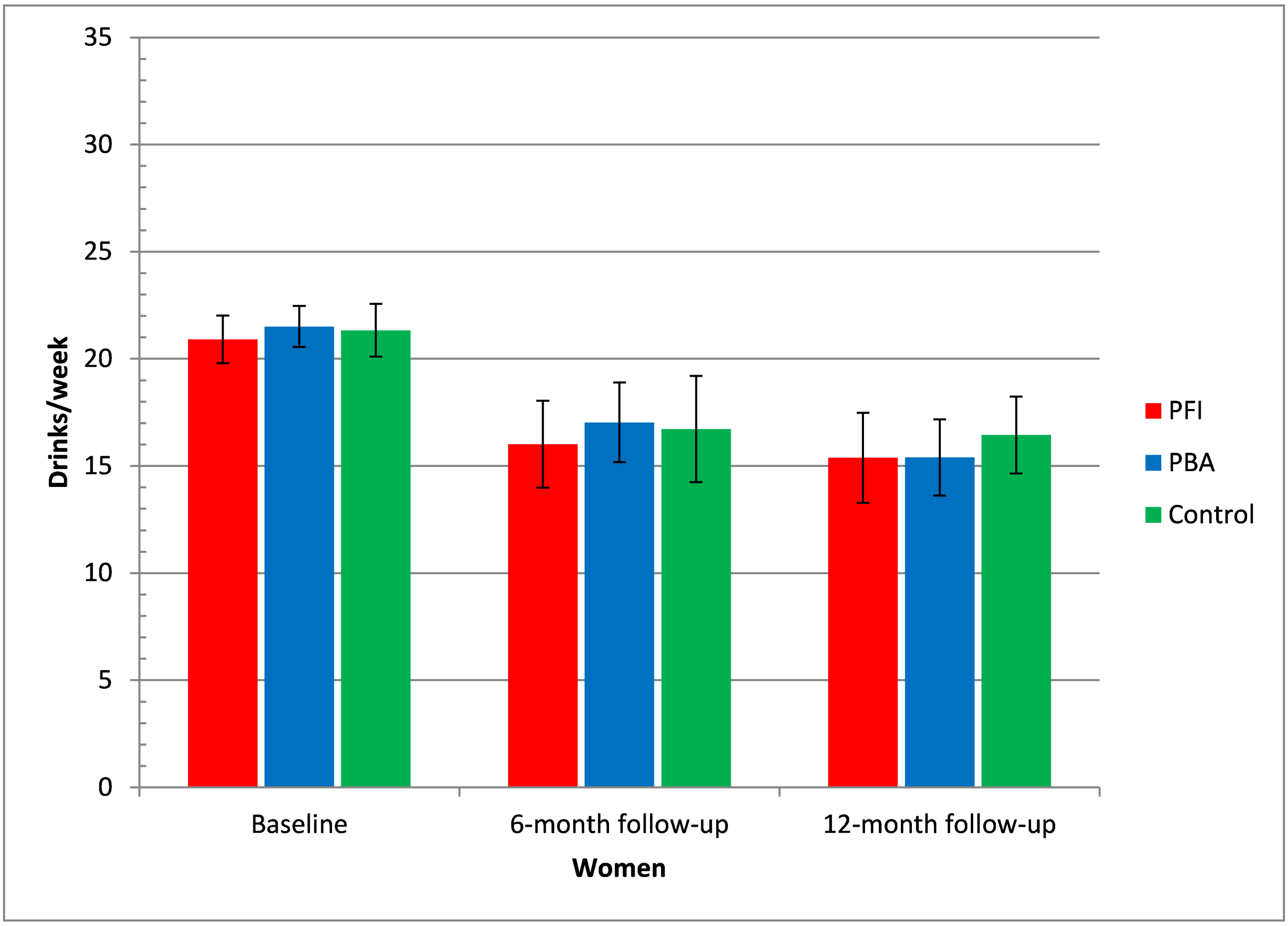


Figure 6. Alcohol consumption at baseline and at 6- and 12-month follow-ups for men based on multiple imputation. Error bars indicate 95% confidence interval. Drinks/week = mean number of standard drinks in a typical week, PBA = Internet-based personalized brief advice intervention, PFI = Internet-based brief personalized feedback intervention.

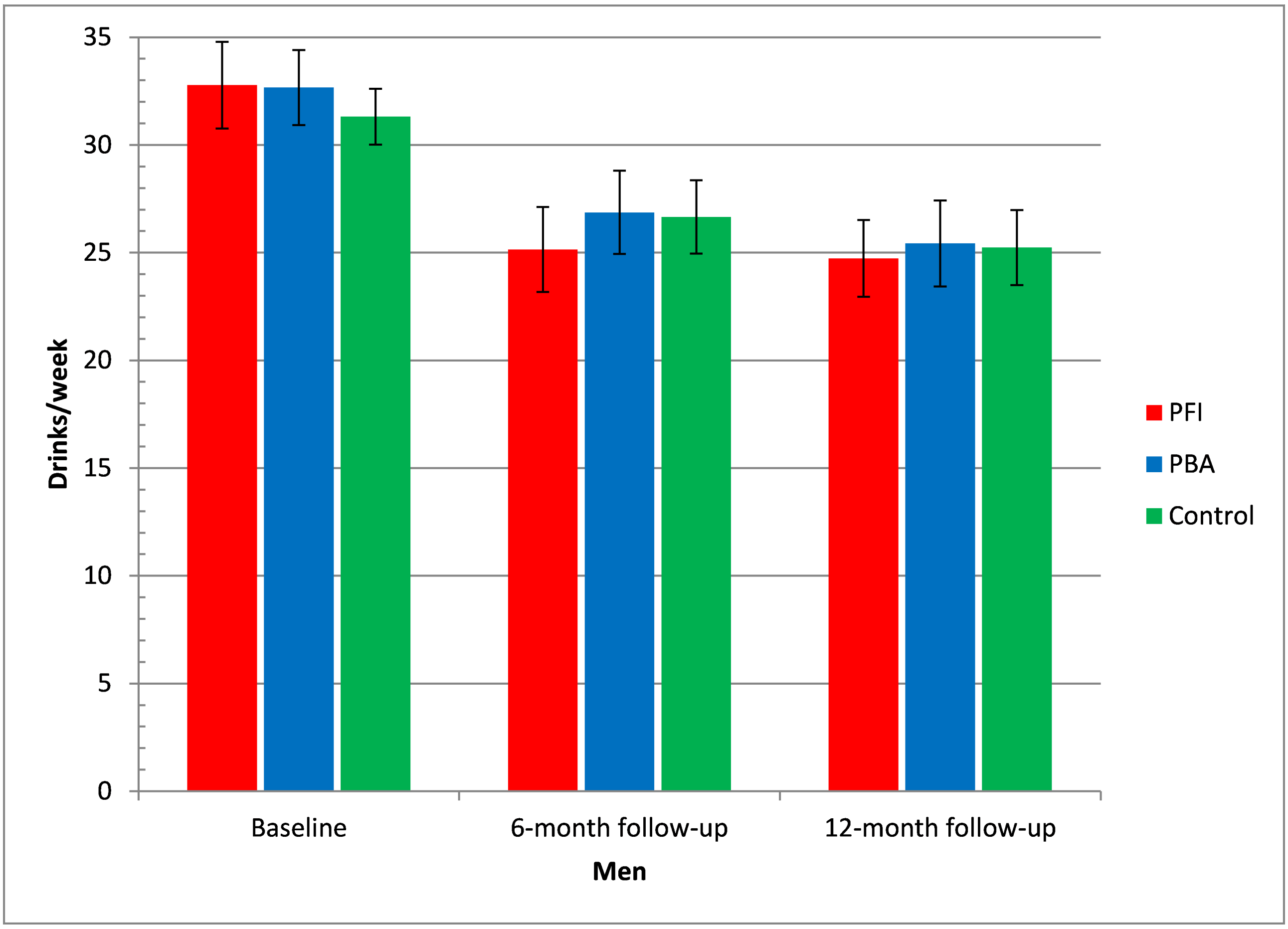


Figure 7. Alcohol consumption at baseline and at 6- and 12-month follow-ups for women based on completers-only analysis. n = 390 after 6 months and 466 after 12 months. Error bars indicate 95% confidence interval. Drinks/week = mean number of standard drinks in a typical week, PBA = Internet-based personalized brief advice intervention, PFI = Internet-based brief personalized feedback intervention.

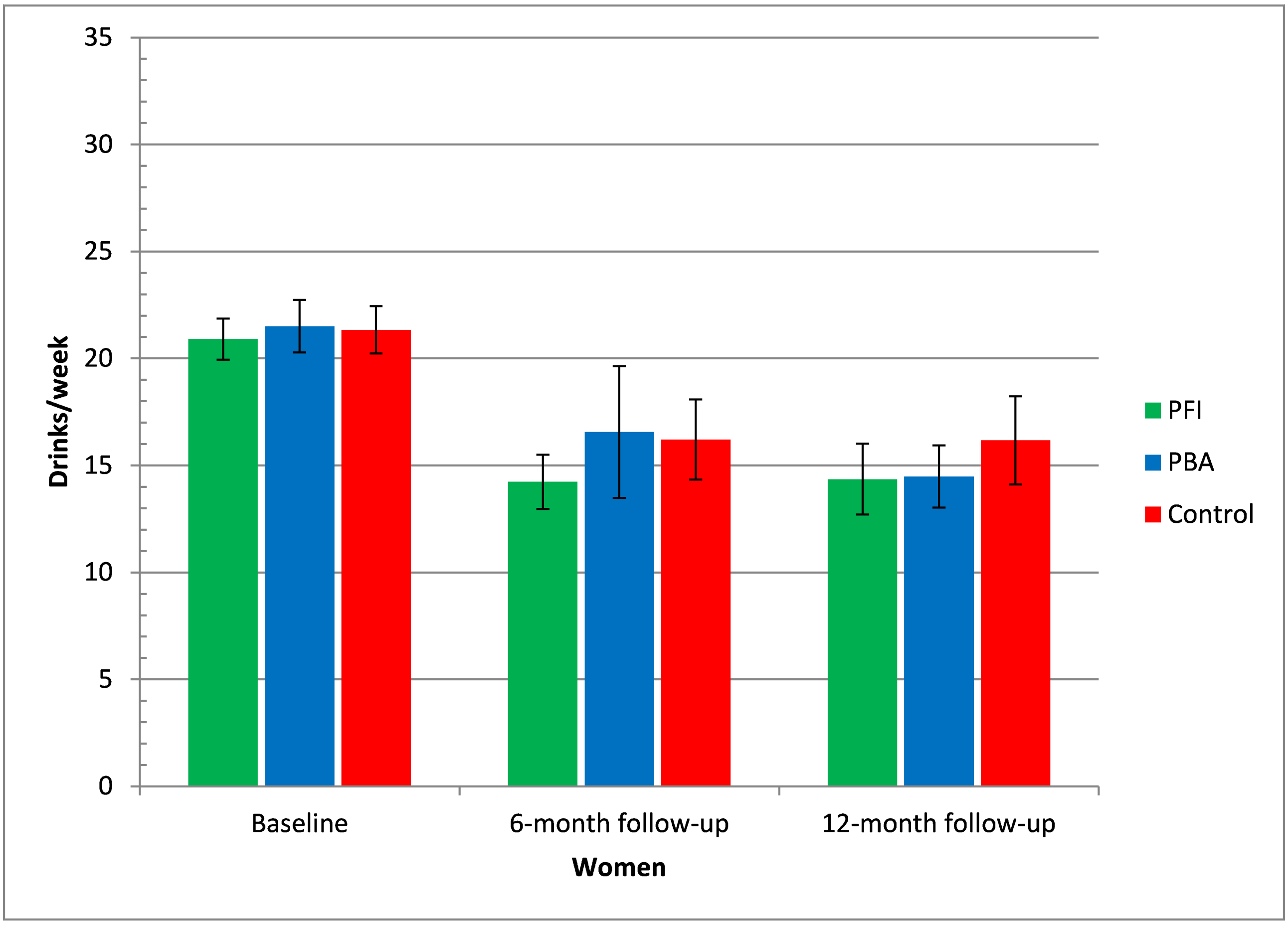
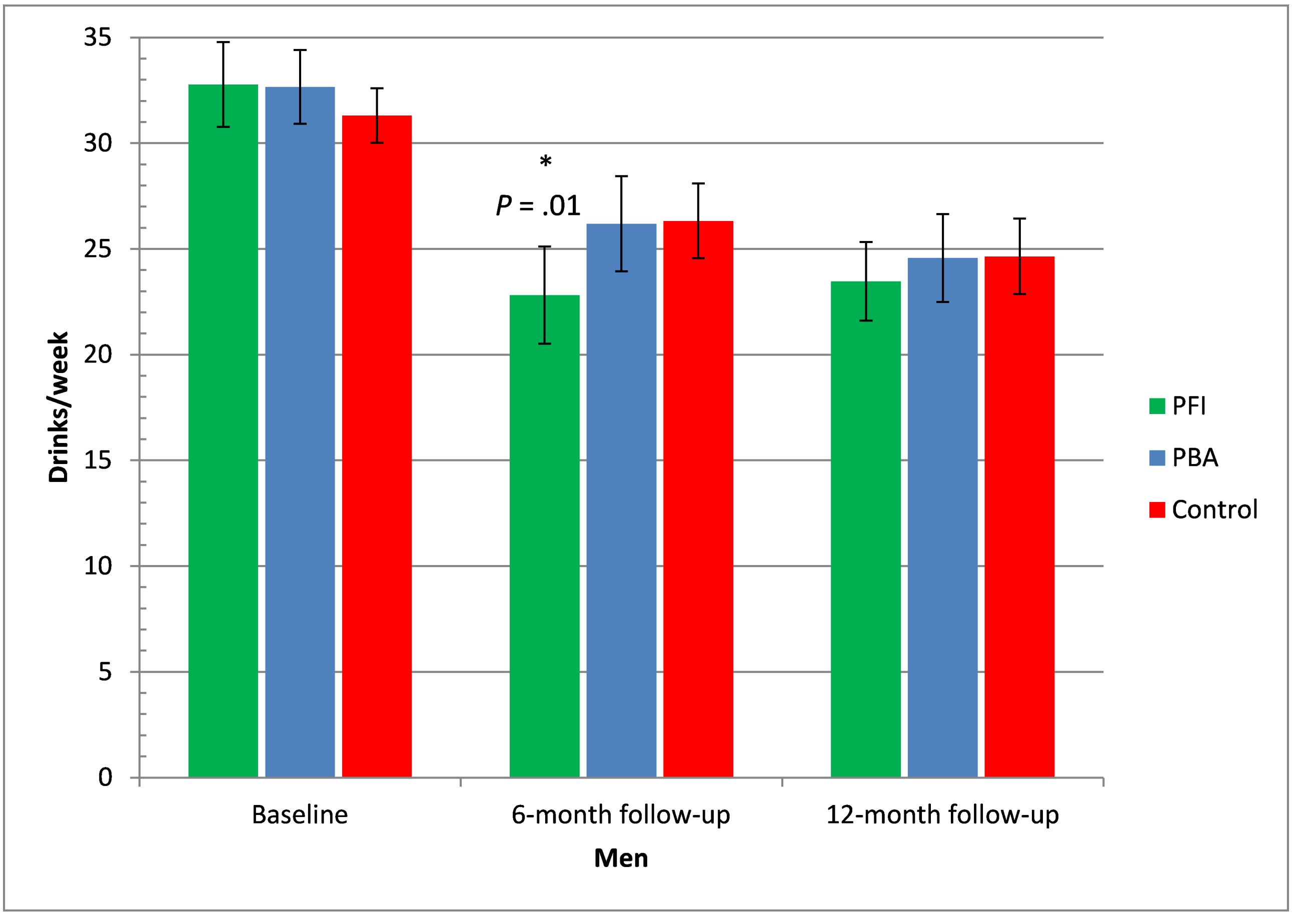


Figure 8. Alcohol consumption at baseline and at 6- and 12-month follow-ups for men based on completers-only analysis. n = 481 after 6 months and 598 after 12 months. Error bars indicate 95% confidence interval. Drinks/week = mean number of standard drinks in a typical week, PBA = Internet-based personalized brief advice intervention, PFI = Internet-based brief personalized feedback intervention. *P value for difference between PFI and control group (Kruskal-Wallis test).



Discussion

Key Findings

In this randomized controlled trial, the primary analysis provided no evidence that an Internet-based brief personalized feedback intervention was effective in reducing drinking in an adult population of heavy drinkers. The intervention effect of the Internet-based brief personalized feedback intervention was approximately 2 drinks/week and nonsignificant, but highly significant in the sensitivity analyses with an intervention effect of approximately 3 drinks/week for the Internet-based brief personalized feedback intervention. In a post hoc secondary completers analysis of men, we found a significant difference of 3.5 drinks/week between the Internet-based brief personalized feedback intervention and the control group at the 6-month follow-up. However, it must be stressed that when interpreting the results from the completers analysis, we are dealing with a self-selected sample and no longer an unbiased sample from a randomized trial, and hence no clear conclusions regarding the efficacy of the intervention can be drawn. From baseline to 6- and 12-month follow-ups, alcohol consumption declined significantly in both intervention groups and the control group by approximately 6 drinks/week.

Possible Mechanism and Explanations for the Findings

When interpreting these results, other factors must be borne in mind that could have contributed to the null findings. Participating in a health examination survey may have motivated participants to change their health behavior, which may have contributed to the decrease in alcohol consumption. Of particular interest in this context is the fact that 46% were motivated to change their alcohol consumption, while 38% were not (that is, of the subsample of 831 persons who were given questions about motivation). The nonblinded nature of the study, and hence the assessment effects (intervention effects of the research procedures), could also explain part of the significant reductions in all groups from the baseline to follow-up [35,36]. These reductions could also be related to regression to the mean, social desirability bias, and historical changes in alcohol consumption. The implications of the above-mentioned effects, if they occurred, are important because, when assessment has a therapeutic benefit or when regression to the mean occurs, the experimental contrast can blur. This is particularly important for brief interventions, where effect sizes are modest [37]. The above-mentioned effects may have biased the results towards the null and hence the intervention effect may be underestimated [38]. It should also be noted that the null finding for the Internet-based personalized brief advice intervention must be interpreted in light of the insufficient sample size to detect an effect of the Internet-based personalized brief advice intervention.

Results in Relation to Other Studies

The null finding in our study is not unusual and mirrors the findings in two recently published trials [39,40]. However, these trials did not include a pure control group, and their intervention websites were somewhat more extensive than the Internet-based brief personalized feedback intervention and Internet-based personalized brief advice intervention used in our study. In fact, many studies have used much more extensive interventions than the very brief Internet-based brief personalized feedback intervention we used. For example, Riper et al used a multicomponent, interactive self-help intervention with a recommended treatment period of 6 weeks [41]. Other studies, such as that of Cunningham et al, recruited participants from a general population telephone survey, which differs from our recruitment procedure by way of a health examination survey [42]. This is likely to have implications for the study population, with regard to generalizability, as a sample of problem drinkers recruited through a telephone survey could be hypothesized to display a broader spectrum of alcohol problems than would a sample from a health examination survey with an emphasis on lifestyle issues in relation to diet, smoking, alcohol, and physical activity [43]. The short duration of our interventions could explain why our study's findings differ from those of three recent meta-analyses, which concluded that brief interventions based on normative feedback are more effective than those that do not include these features [14-16]. Riper et al found an effect size (Cohen *d*) of 0.22 (95% CI 0.16–0.29) for brief, single-session personalized feedback interventions [16], and Webb et all observed small but significant effects on behavior for interventions that provided automated tailored feedback, with an effect size (Cohen *d*) of 0.18 (95% CI 0.07–0.28) [15]. In terms of amount, a systematic review found a mean difference of 26 g of alcohol between computer-based interventions and minimally active comparator groups [14].

Strengths and Limitations of the Study

In this rigorously conducted trial, we succeeded in implementing an Internet-based intervention in a general population-based sample of heavy drinkers. The naturalistic setting of the trial (ie, participants accessed the intervention in their own homes) increases confidence in the generalizability of the results [18,42]. Another strength of the naturalistic trial design is that it elucidates important feasibility aspects of reaching a non-treatment-seeking population of heavy drinkers by email. Knowing that 36% of the invitees did not respond to the invitation email and that 23% declined participation is applicable knowledge when designing and disseminating similar interventions.

Our aim was to investigate how minimal an Internet-based intervention can be while still having an impact on drinking. Thus, the interventions were displayed in a single screenshot immediately after the participants had provided their online consent, and we avoided the problem of knowing whether the participants randomly assigned to the interventions actually used the interventions [42].

Due to our design with two intervention groups and a pure control group, this study partially supports our hypothesis that the active component in our interventions is feedback regarding one's own drinking relative to normative standards, at least in the sensitivity analyses. This knowledge can be used in the design of future alcohol interventions. Knowing the mechanism of change would be an important contribution to the Internet-based interventions field, as the existing research in this area has been focused on college samples [44].

Attrition (37% at 6 months and 23% at 12 months) is an area of concern in our study, as it could introduce a selection bias, thereby causing imbalance among the previously randomized groups and threatening internal validity. This was partly confirmed by our analysis, which revealed differential attrition. Participants lost to follow-up from the Internet-based brief personalized feedback intervention group were more likely to be unmotivated to cut down on their drinking. Furthermore, participants lost to follow-up were more likely to be heavy smokers and to have a low level of education. The importance of dealing with this complicated picture of differential attrition is underlined by the sensitivity analyses, which revealed that an analysis of completers only, or an analysis using last observation carried forward, will probably overestimate the treatment effects. By using multiple imputation in our main analysis, we have provided a plausible estimate of the possible result if no attrition had occurred. The generalizability of the present findings is restricted due to the underrepresentation of individuals with the lowest level of education, unmarried individuals, and younger individuals in our population. However, generalizability is a common problem in much brief intervention research that deals with populations that are not representative of the population of heavy drinkers [37,45]. This was confirmed by our results, which showed that in the Danish Health Examination Survey population, 6% were heavy drinkers, compared with a 20% prevalence estimated for the Danish population as a whole [3]. Due to our recruitment of participants from a sample in which almost everyone had Internet access, Internet and computer literacy were high in our sample. When generalizing to the Danish population, the fact that 86% of Danes have Internet access should be borne in mind. The results must also be interpreted in consideration the possibility that the use of a health examination survey to proactively enlist heavy drinkers (who were not seeking help) may have resulted in a preponderance of heavy drinkers with low levels of alcohol-related harm.

We relied on measuring outcome using self-reports of alcohol consumption, which is a method that has demonstrated reasonable levels of accuracy [37,46]. We tried to minimize the bias of underreporting by asking beverage-specific questions, which has been shown to yield higher volumes of alcohol consumption than questions that only ask for total alcohol consumption [47].

Conclusions

In this Internet-based study, we compared the efficacy of personalized brief advice and personalized normative feedback against a pure control group in a non-treatment-seeking population of adult heavy drinkers. The main analysis lends no support to the efficacy of personalized normative feedback or personalized brief advice. However, on the grounds of the sensitivity analyses, we cautiously conclude that the personalization in conjunction with the normative feedback enhanced attention to the message in the Internet-based brief personalized feedback intervention and thus gave an indication of decreased alcohol consumption in the Internet-based brief personalized feedback intervention group. It seems that the potential of encouraging people to become more aware of the level and consequences of their drinking, and how their drinking behaviors compare with those of others in a similar social or demographic group, is an applicable insight, in both medical and public health settings, when it comes to reducing heavy drinking.

Acknowledgments

The study was approved by the Danish local ethics committee and the Danish Data Protection Agency, and all local confidentiality and privacy requirements have been met.

The study was funded by the National Board of Health, Denmark. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interactive website used for the study, including information about the study, online consent, interventions and control conditions, and screenshots of additional material linked to from the interventions. Website can be found at URL: http://d1m.dk/DANHES/ [No login needed, translated from Danish].

[PPT File (Microsoft PowerPoint Presentation), 2MB - jmir_v14i4e98_app1.ppt]

Multimedia Appendix 2

CONSORT-Ehealth Checklist V1.6.1 [48].

[PDF File (Adobe PDF File), 572KB - jmir_v14i4e98_app2.pdf]

References

1. Rehm J, Baliunas D, Borges GL, Graham K, Irving H, Kehoe T, et al. The relation between different dimensions of alcohol consumption and burden of disease: an overview. Addiction 2010 May;105(5):817-843. [doi: 10.1111/j.1360-0443.2010.02899.x] [Medline: 20331573]
2. Tolstrup JS, Jensen MK, Tjønneland A, Overvad K, Grønbaek M. Drinking pattern and mortality in middle-aged men and women. Addiction 2004 Mar;99(3):323-330. [doi: 10.1111/j.1360-0443.2003.00667.x] [Medline: 14982545]
3. Hansen AB, Hvidtfeldt UA, Grønbæk M, Becker U, Nielsen AS, Tolstrup JS. The number of persons with alcohol problems in the Danish population. Scand J Public Health 2011 Mar;39(2):128-136. [doi: 10.1177/1403494810393556] [Medline: 21382854]
4. 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]
5. Beich A, Gannik D, Malterud K. Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. BMJ 2002 Oct 19;325(7369):870. [Medline: 12386040]
6. Nilsen P. Brief alcohol intervention--where to from here? Challenges remain for research and practice. Addiction 2010 Jun;105(6):954-959. [doi: 10.1111/j.1360-0443.2009.02779.x] [Medline: 20121717]
7. 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]
8. Harris K, McKellar J. Demand for alcohol treatment. Frontlines 2003;June:1-2.
9. Lieberman DZ, Huang SW. A technological approach to reaching a hidden population of problem drinkers. Psychiatr Serv 2008 Mar;59(3):297-303. [doi: 10.1176/appi.ps.59.3.297] [Medline: 18308911]
10. 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. [doi: 10.1176/appi.ajp.162.12.2393] [Medline: 16330613]
11. Internet World Stats. Miniwatts Marketing Group. 2011 Jun 22. World Internet Usage and Population Statistics March 31, 2011 URL: http://www.internetworldstats.com/stats.htm [accessed 2011-06-23] [WebCite Cache ID 5zec0twEe]
12. 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]

13. 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]
14. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. Addiction 2011 Feb;106(2):267-282. [doi: 10.1111/j.1360-0443.2010.03214.x] [Medline: 21083832]
15. 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]
16. 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]
17. 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]
18. Cunningham JA, Van Mierlo T. Methodological issues in the evaluation of Internet-based interventions for problem drinking. Drug Alcohol Rev 2009 Jan;28(1):12-17. [doi: 10.1111/j.1465-3362.2008.00001.x] [Medline: 19320670]
19. Collins SE, Carey KB, Sliwinski MJ. Mailed personalized normative feedback as a brief intervention for at-risk college drinkers. J Stud Alcohol 2002 Sep;63(5):559-567. [Medline: 12380852]
20. Walters ST, Neighbors C. Feedback interventions for college alcohol misuse: what, why and for whom? Addict Behav 2005 Jul;30(6):1168-1182. [doi: 10.1016/j.addbeh.2004.12.005] [Medline: 15925126]
21. Wild TC, Cunningham JA, Roberts AB. Controlled study of brief personalized assessment-feedback for drinkers interested in self-help. Addiction 2007 Feb;102(2):241-250. [doi: 10.1111/j.1360-0443.2006.01682.x] [Medline: 17222278]
22. Neighbors C, Larimer ME, Lewis MA. Targeting misperceptions of descriptive drinking norms: efficacy of a computer-delivered personalized normative feedback intervention. J Consult Clin Psychol 2004 Jun;72(3):434-447. [doi: 10.1037/0022-006X.72.3.434] [Medline: 15279527]
23. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
24. Eriksen L, Grønbæk M, Helge JW, Tolstrup JS, Curtis T. The Danish Health Examination Survey 2007-2008 (DANHES 2007-2008). Scand J Public Health 2011 Mar;39(2):203-211. [doi: 10.1177/1403494810393557] [Medline: 21257645]
25. 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]
26. Ballesteros J, González-Pinto A, Querejeta I, Ariño J. Brief interventions for hazardous drinkers delivered in primary care are equally effective in men and women. Addiction 2004 Jan;99(1):103-108. [Medline: 14678068]
27. Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011 Jul;39(7 Suppl):22-25. [doi: 10.1177/1403494810387965] [Medline: 21775345]
28. Babor T, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. 2nd edition. Geneva: World Health Organization; 2001.
29. Mørch LS, Becker U, Olsen J, Tjønneland AM, Grønbaek MN. [Should the sensible drinking limits for adults be changed?]. Ugeskr Laeger 2005 Oct 3;167(40):3777-3779. [Medline: 16221393]
30. Bertholet N, Daeppen JB, Wietlisbach V, Fleming M, Burnand B. Reduction of alcohol consumption by brief alcohol intervention in primary care: systematic review and meta-analysis. Arch Intern Med 2005 May 9;165(9):986-995. [doi: 10.1001/archinte.165.9.986] [Medline: 15883236]
31. Jenkins RJ, McAlaney J, McCambridge J. Change over time in alcohol consumption in control groups in brief intervention studies: systematic review and meta-regression study. Drug Alcohol Depend 2009 Feb 1;100(1-2):107-114. [doi: 10.1016/j.drugalcdep.2008.09.016] [Medline: 19041196]
32. Finucane MM, Samet JH, Horton NJ. Translational methods in biostatistics: linear mixed effect regression models of alcohol consumption and HIV disease progression over time. Epidemiol Perspect Innov 2007;4:8 [FREE Full text] [doi: 10.1186/1742-5573-4-8] [Medline: 17880699]
33. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393. [Medline: 19564179]
34. Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. Am J Epidemiol 2010 Mar 1;171(5):624-632 [FREE Full text] [doi: 10.1093/aje/kwp425] [Medline: 20106935]
35. McCambridge J, Day M. Randomized controlled trial of the effects of completing the Alcohol Use Disorders Identification Test questionnaire on self-reported hazardous drinking. Addiction 2008 Feb;103(2):241-248. [doi: 10.1111/j.1360-0443.2007.02080.x] [Medline: 18199302]

36. Kypri K, Langley JD, Saunders JB, Cashell-Smith ML. Assessment may conceal therapeutic benefit: findings from a randomized controlled trial for hazardous drinking. Addiction 2007 Jan;102(1):62-70. [doi: 10.1111/j.1360-0443.2006.01632.x] [Medline: 17207124]
37. Kypri K. Methodological issues in alcohol screening and brief intervention research. Subst Abus 2007;28(3):31-42. [doi: 10.1300/J465v28n03_04] [Medline: 18077301]
38. Bernstein JA, Bernstein E, Heeren TC. Mechanisms of change in control group drinking in clinical trials of brief alcohol intervention: implications for bias toward the null. Drug Alcohol Rev 2010 Sep;29(5):498-507. [doi: 10.1111/j.1465-3362.2010.00174.x] [Medline: 20887573]
39. Wallace P, Murray E, McCambridge J, Khadjesari Z, White IR, Thompson SG, et al. On-line randomized controlled trial of an internet based psychologically enhanced intervention for people with hazardous alcohol consumption. PLoS One 2011;6(3):e14740 [FREE Full text] [doi: 10.1371/journal.pone.0014740] [Medline: 21408060]
40. 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 [FREE Full text] [doi: 10.2196/jmir.1695] [Medline: 21719412]
41. 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]
42. 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]
43. Feveile H, Olsen O, Hogh A. A randomized trial of mailed questionnaires versus telephone interviews: response patterns in a survey. BMC Med Res Methodol 2007;7:27 [FREE Full text] [doi: 10.1186/1471-2288-7-27] [Medline: 17592653]
44. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
45. Edwards AG, Rollnick S. Outcome studies of brief alcohol intervention in general practice: the problem of lost subjects. Addiction 1997 Dec;92(12):1699-1704. [Medline: 9581002]
46. 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]
47. Rehm J. Measuring quantity, frequency, and volume of drinking. Alcohol Clin Exp Res 1998 Apr;22(2 Suppl):4S-14S. [Medline: 9603301]
48. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CI: confidence interval

Edited by G Eysenbach; submitted 08.07.11; peer-reviewed by K Kypri, J Cunningham, Anonymous; comments to author 13.09.11; revised version received 04.10.11; accepted 28.03.12; published 30.07.12.

Please cite as:
Hansen ABG, Becker U, Nielsen AS, Grønbæk M, Tolstrup JS, Thygesen LC
Internet-Based Brief Personalized Feedback Intervention in a Non-Treatment-Seeking Population of Adult Heavy Drinkers: A Randomized Controlled Trial
J Med Internet Res 2012;14(4):e98
URL: http://www.jmir.org/2012/4/e98/
doi:10.2196/jmir.1883
PMID:22846542

Original Paper

Comparison of Two Internet-Based Interventions for Problem Drinkers: Randomized Controlled Trial

John Alastair Cunningham[1], PhD

[1] Social and Epidemiological Research, Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
John Alastair Cunningham, PhD
Social and Epidemiological Research
Centre for Addiction and Mental Health
33 Russell St
Toronto, ON, M5S 2S1
Canada
Phone: 1 416 535 8501 ext 4682
Fax: 1 416 595 6899
Email: john_cunningham@camh.net

Abstract

Background: Alcohol problems are a serious public health concern, and few problem drinkers ever seek treatment. The Internet is one means of promoting access to care, but more research is needed to test the best types of interventions to employ. Evaluation of Internet-based interventions that contain a variety of research-validated cognitive-behavioral tools, which have been shown to be helpful to those with more severe alcohol concerns, should be a priority.

Objective: To evaluate whether providing access to an extended Internet intervention for alcohol problems offers additional benefits in promoting reductions in alcohol consumption compared with a brief Internet intervention. The hypothesis for the current trial was that respondents who were provided with access to an extended Internet intervention (the Alcohol Help Center [AHC]) would display significantly improved drinking outcomes at 6-month follow-up, compared with respondents who were provided with access to a brief Internet intervention (the Check Your Drinking [CYD] screener).

Methods: A single-blinded randomized controlled trial with a 6-month follow-up. A general population sample of problem drinkers was recruited through newspaper advertisements in a large metropolitan city. Baseline and follow-up data were collected by postal mail.

Results: A volunteer sample of problem drinkers of legal drinking age with home access to the Internet were recruited for the trial. Of 239 potential respondents recruited in 2010, 170 met inclusion criteria (average age 45 years; 101/170, 59.4% male; average Alcohol Use Disorders Identification Test [AUDIT] score of 22). Follow-up rates were 90.0% (153/170) with no adverse effects of the interventions reported. A repeated-measures multivariate analysis of variance of the outcome measures using an intent-to-treat approach found a significantly greater reduction in amount of drinking among participants provided access to the AHC than among participants provided access to the CYD ($P = .046$).

Conclusions: The provision of the AHC gave additional benefit in the short term to problem drinkers over that seen from the research-validated CYD, indicating the benefits of promoting access to these interventions as one means of helping people with problem drinking concerns.

Trial Registration: ClinicalTrials.gov NCT01114919; http://clinicaltrials.gov/ct2/show/NCT01114919 (Archived by WebCite at http://www.webcitation.org/68t1dCkRZ)

(J Med Internet Res 2012;14(4):e107) doi:10.2196/jmir.2090

KEYWORDS

Randomized controlled trial; problem drinking; alcohol abuse; Internet-based intervention; eHealth; brief intervention

Introduction

Alcohol is the third-leading cause of preventable death [1]. Unfortunately, the majority of people with drinking problems will never seek specialized addictions treatment [2]. Brief interventions in primary care settings have been identified as one means of addressing this important health problem [3]. However, given the prevalence of drinking problems and the resource restrictions in primary care settings, there is a need to also find alternative means of helping those with drinking problems.

Internet-based interventions have been identified as one promising option. Several reviews have concluded that there is a fast-developing evidence base for the efficacy of these interventions [4-8], particularly among college students, where the majority of these trials have been conducted. The evaluation of the efficacy of Internet-based interventions in general population samples is important if these brief interventions are to be promoted as helpful to anyone other than problem drinking, young adult college students. In addition, the majority of these Internet-based interventions have consisted of brief, personalized feedback interventions, which are thought to be useful to those with less severe alcohol problems. Evaluation of extended interventions that contain a variety of research-validated cognitive-behavioral tools, which have been shown to be helpful to those with more severe alcohol concerns, should also be a priority.

In this paper we report results of a randomized controlled trial testing the added benefit of providing access to such an extended Internet-based intervention, which contained an extensive array of cognitive-behavioral tools for problem drinkers (the Alcohol Help Center [AHC]), over the provision of a brief, personalized feedback Internet-based intervention (the Check Your Drinking [CYD] screener). Both of these Internet-based interventions are available free of charge on the Internet, making evaluations of their use of immediate benefit to problem drinkers. In addition, the CYD has already been subjected to four randomized controlled trials, in which the provision of this brief intervention yielded reductions in alcohol consumption among participants in a variety of different settings [9-12]. In the one study that employed a general population sample of problem drinkers [9], being provided access to the CYD resulted in an average reduction of 6 drinks at 3- and 6-month follow-up as compared with a no-intervention control group. Thus, the CYD brief intervention is an excellent comparator to evaluate whether providing access to an extended intervention (the AHC) would have additional benefit in promoting reductions in alcohol consumption. Finally, as there are few freely accessible, extended Internet-based interventions available [13,14] and as the evaluation of the efficacy of these extended Internet-based interventions has yielded mixed results [15], it is important to conduct further research in this area. The hypothesis for the current trial was that respondents in the extended Internet intervention condition (the AHC) would display significantly improved drinking outcomes at 6-month follow-up, compared with respondents in the brief Internet intervention condition (the CYD).

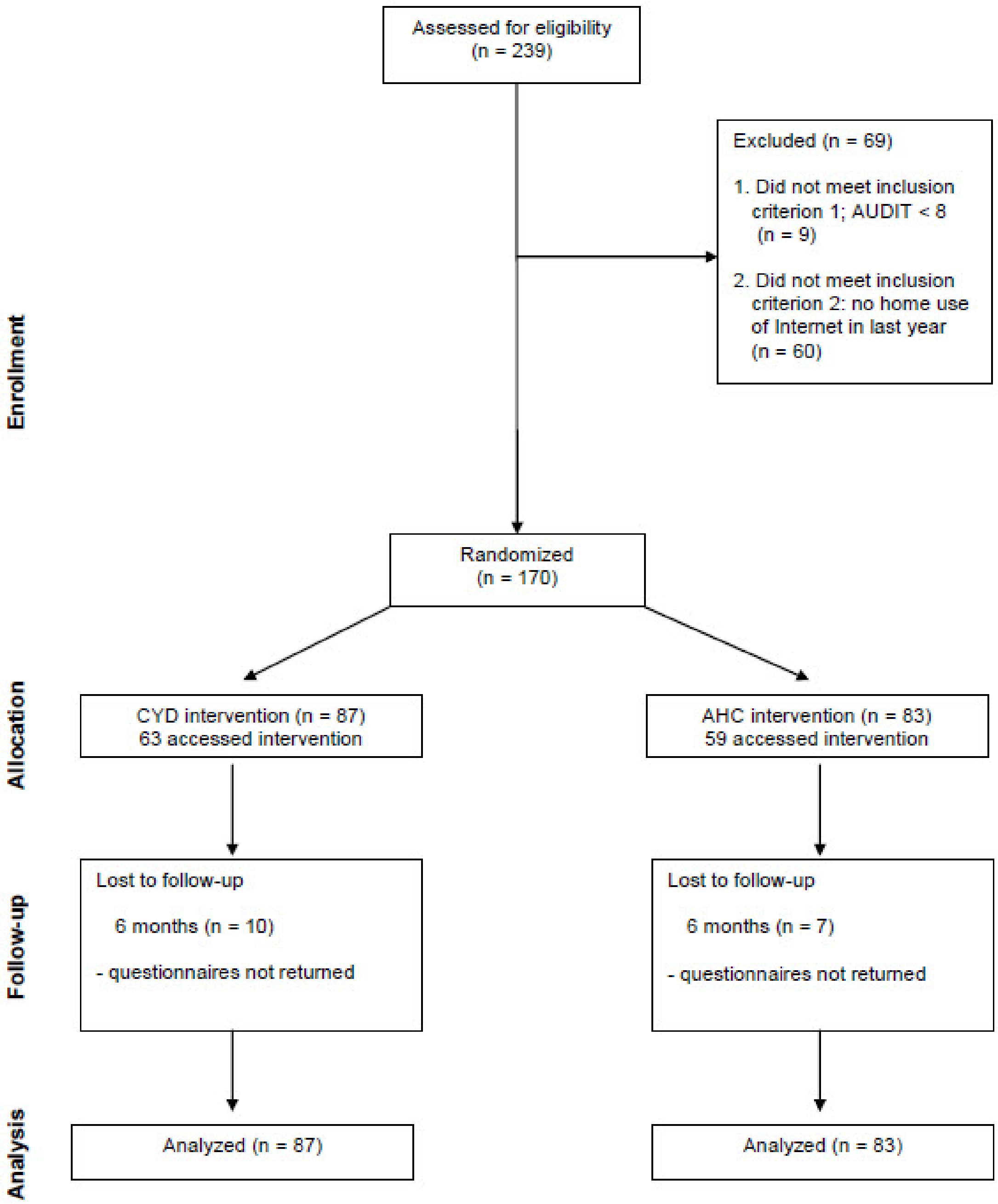
Methods

Participants were recruited through newspaper advertisements in metropolitan Toronto, Canada (May to September 2010) asking for current drinkers interested in helping “in the process of developing and evaluating Internet-based interventions for alcohol users.” Interested potential participants called the telephone number provided and left their name and address to be sent a consent form and a baseline questionnaire by postal mail. The questionnaire contained a graphic describing a standard drink (note that a standard drink in Canada contains 13.6 g of alcohol).

Those returning the signed consent form and the completed baseline questionnaire were randomly assigned into one of two conditions: to be provided access to the brief Internet-based CYD brief personalized feedback intervention or to be provided access to the extended Internet-based AHC. Access was provided by sending each participant a unique password to be entered into a study-specific online portal (password sent in a letter by postal mail). Random numbers where generated in blocks of two using an online random numbers generator (Research Randomizer, Social Psychology Network; http://www.randomizer.org/) by the author. The random numbers list was used sequentially by a research assistant, who sent out the relevant intervention letter to participants in the order that their consent forms were returned. Participants were blind to the different experimental conditions in the study. Participants were followed up at 6 months and were compensated Can $20 for completing the follow-up questionnaire (questionnaire and check sent by postal mail). Follow-ups were conducted from November 2010 to June 2011. If participants did not return their original follow-up questionnaire within 1 month, they were sent a second questionnaire by postal mail. The research assistant in charge of follow-up was not blind to experimental condition, as the follow-up questionnaires were slightly different (outcome variables were identical and were asked about first, but some questions about impressions of the websites were different at the end of the questionnaire). The trial was approved by the standing research ethics committee for the Centre for Addiction and Mental Health.

As per protocol, the primary outcome measures were number of drinks consumed in a typical week, highest number of drinks on one occasion in the last 6 months, frequency of drinking, typical quantity of drinking on one occasion, and frequency of drinking 5 or more drinks on one occasion (the questions were the same for male and female participants). These last 3 items were combined to form the Alcohol Use Disorders Identification Test-Consumption subscale (AUDIT-C) [16]. For participants who did not return the follow-up questionnaire, missing data on the outcome measures were replaced with their respective baseline values. In addition, variables were examined for distributional properties and were Winsorized to deal with outliers. Exclusion criteria were being less than 19 years old (the legal drinking age in Ontario, Canada), having an AUDIT-C score of less than 8 (a score of 8 or more indicates current problem drinking) [17,18], and not having home access to the Internet. Lack of home Internet access was included in an attempt to include only regular Internet users in the trial. See Figure 1 for a CONSORT diagram of the trial.

Figure 1. CONSORT diagram of participant recruitment. AHC = Alcohol Help Center, AUDIT = Alcohol Use Disorders Identification Test-C, CYD = Check Your Drinking.



Statistical Analyses and Power Analysis

Results from a related self-help trial [19] were used to estimate that the addition of the extended intervention to the brief intervention would explain 3% more of the variance at 6 months. An increase of 3% of the explained variance is a medium effect size and corresponds to a mean decrease of approximately 3 drinks per week. Following the convention that studies should be designed to have a statistical power of at least 80%, and that hypotheses be tested at the .05 level of significance, SamplePower 1.0 [20] was used to estimate the required sample size. These specifications result in a final sample (required after attrition) of N = 170 (85 respondents per condition). Analyses were conducted using a repeated-measures multivariate analysis of variance with intervention condition (CYD vs AHC) as the between-participants variable and time (baseline vs 6 months) as the within-participants variable (including all outcome variables: typical weekly drinking, highest amount on one occasion, AUDIT-C score).

Interventions

CYD Brief Intervention

The CYD screener (CheckYourDrinking.net) is a brief and personalized 18-item screener that has been designed to assess quantity and frequency of drinking, and the severity of drinking problems [21]. Following completion of the 18-item screener, the user is provided with a personalized final report that compares the person's drinking with that of others in the general population of the same age, sex, and country of origin (for Canada, the United States, and the United Kingdom). The CYD is intended for quick completion in a single session, although participants could access it as often as they wanted.

AHC Extended Internet-Based Intervention

The AHC (AlcoholHelpCenter.net) is a website that has been developed to contain the cognitive-behavioral, motivational, and relapse prevention components that have been found to be effective in well-validated self-help books and other brief interventions. Specific sources for the content include (1) components found in the self-help book *DrinkWise* [22], (2) exercises from the self-help booklet *Alcohol and You*, developed by the principal investigator of this study [23], and (3) common relapse prevention exercises used in brief treatment modalities [24]. The main component of the AHC is divided into three sections: getting started (10 exercises focused on initiating change), dealing with difficulties (6 exercises covering some of the key issues that often occur as a problem drinker works on change), and maintenance (4 exercises designed to help participants maintain their change). In addition, there is a series of interactive tools that are useful throughout the change process, such as a drinking diary, where the participant is encouraged to track his or her drinking, and a blood alcohol calculator. Further, there are several elements that provide additional support to the participant: (1) an online support group that is moderated by health educators [25], (2) an email messaging system that provides the participant with encouragement and tips to deal with drinking concerns, and (3) a text messaging program for interested participants to be sent tips on how to deal with drinking problems. In summary, the AHC is a well-designed Internet-based intervention that contains many research-validated elements that have been shown in self-help books and other brief interventions to help problem drinkers. Participants have the option of completing whatever exercises they choose in whatever order they like. The AHC is intended for repeated use over an extended time period.

Results

A total of 170 participants met inclusion criteria for this trial. Bivariate comparisons compared baseline demographic and drinking characteristics between experimental conditions (CYD, n = 87; AHC, n = 83) and found no significant differences between them ($P > .05$). Participants' mean age was 45.2 (SD 12.2), 59.4% (101/170) were male, 58.8% (100/170) had some postsecondary education, 39% (67/170) were married or living with a partner, 55% (94/170) were employed full- or part-time, and more than half (92/170, 54%) had a yearly household income of Can $30,000 or more. Baseline drinking was heavy, with participants reporting an average AUDIT score of 22.1 (SD 7.6), typical weekly consumption of 31.7 (SD 18.4) drinks, and highest number of drinks consumed on one occasion of 13.7 (SD 7.2). A total of 71.8% (122/170) of participants used the password and logged in to their respective online intervention; however, all participants were included in the analysis in an intent-to-treat approach. Follow-up rates were good, with 90% (n = 153) completing the 6-month follow-up. There were no significant differences ($P > .05$) between experimental conditions, demographic characteristics, or drinking characteristics in the proportion of participants who completed the 6-month follow-up.

A multivariate analysis of variance was conducted and found a main effect for time (baseline vs 6 months; $F_{3,163} = 15.0$, $P < .001$, partial eta squared = 0.22). In addition, there was a significant interaction between time and condition ($F_{3,163} = 2.7$, $P = .046$, partial eta squared = 0.05). Further univariate tests found that the time × condition interaction was not significant for the number of drinks in a typical week variable ($F_{1,165} = 0.4$, $P = .53$), approached significance for the AUDIT-C variable ($F_{1,165} = 3.6$, $P = .06$, partial eta squared = 0.021), and was significant for the highest number of drinks on one occasion variable ($F_{1,165} = 5.7$, $P = .02$, partial eta squared = 0.034). Table 1 displays the means for the three variables at baseline and 6-month follow-up by condition. Participants in the AHC condition reduced the highest amount they consumed on one occasion from baseline to 6-month follow-up more than did participants in the CYD condition.

Table 1. Mean (SD) alcohol consumption and AUDIT-C[a] scores at baseline and 6-month follow-up by intervention condition.

Variable	Baseline		6-month follow-up	
	Check Your Drinking (n = 87)	Alcohol Help Center (n = 83)	Check Your Drinking (n = 87)	Alcohol Help Center (n = 83)
No. drinks in a typical week	30.8 (19.2)	32.6 (17.9)	25.9 (18.7)	26.3 (20.7)
AUDIT-C score[b]	8.6 (2.0)	9.0 (1.5)	8.1 (2.5)	8.0 (2.6)
Highest number of drinks on one occasion[c]	13.1 (7.0)	13.9 (7.3)	11.5 (6.1)	10.3 (5.4)

[a] Alcohol Use Disorders Identification Test-Consumption subscale is a composite measure that consists of respondents' scores on frequency of drinking, drinks per drinking day, and frequency of 5 or more drinks on one occasion. Scores range from 0 to 12.

[b] Time × condition, $P = .06$.

[c] Time × condition, $P = .02$.

Discussion

The current trial found a small but significant additional reduction in drinking among participants who were provided with access to the extended Internet-based intervention (the AHC) as compared with participants receiving the brief Internet-based intervention (the CYD) at a 6-month follow-up. The difference specifically had to do with reductions in amount of alcohol consumed during the participants' heaviest drinking occasion. Given the increased risks associated with heavy drinking situations [26], this reduction is encouraging, although levels of consumption were still high in both conditions at follow-up. These results are also strengthened by the excellent follow-up rate in this trial (90%), something that has been put forth as a specific challenge in the conduct of eHealth intervention research [27]. The high follow-up rate could have resulted from the Can $20 honorarium or because much of the communication associated with the trial was conducted by postal mail rather than through email and online surveys. In addition, there was a trend ($P = .06$) toward an additional benefit of the AHC over the CYD on the AUDIT-C scores, a scale that is a composite measure of hazardous drinking. Finally, participants in both groups reduced their drinking in a typical week substantially (by 4.9 drinks per week in the CYD condition and 6.3 drinks per week in the AHC condition). These reductions are similar in scope to that seen in an earlier randomized controlled trial comparing the CYD with a no-intervention control [9]. However, there was no significant difference between the CYD and the AHC conditions in amount of drinking in a typical week in the current trial, so no claims regarding the efficacy of the AHC in reducing typical weekly drinking can be made based on these results (ie, there is no evidence that these results could not have been due to regression to the mean).

A strength, and a simultaneous limitation, of this trial was the severity of current drinking problems in the participants (an AUDIT score of 20 or more is taken as a reasonable proxy of alcohol dependence, and the mean AUDIT score in the current trial was 22) [17]. Thus, while it is encouraging to see an additional impact of the AHC over the CYD among this sample of people with relatively severe alcohol consumption, realistically, this may not be the ideal target population for an Internet-based intervention (even an extended one) [17,28]. Hybrid versions of Internet-mediated interventions where therapists interact with clients through an Internet portal is a promising alternative for treating those with more severe alcohol problems if face-to-face options do not exist [29,30]. Future research trials evaluating this intervention, or other extended Internet-based interventions of a similar nature, would be well served to use recruitment methods that emphasize those with hazardous, but not severe, alcohol problems. Other limitations that could be addressed in a systematic replication of this trial are a lack of a no-interventions control group (which would allow us to make statements about the efficacy of both Internet-based interventions to reduce typical weekly consumption), data reporting on the degree of engagement with the AHC, and a long-term follow-up that would allow statements to be made about any extended impact of these interventions. Finally, the results of this trial relied on self-report of the participants. While there is good evidence that self-reports of drinking are generally reliable in this type of trial [31] and, more important, there is no reason to expect that there would be differences between experimental conditions in the accuracy of participants' self-reports, further replications of this research would benefit from validation of participants' self-reports.

Despite these limitations, the results of this trial are important, as they demonstrate the added benefit of an extended Internet-based intervention over a brief, personalized feedback intervention for problem drinkers. Systematic replications of these findings would allow the development of an adequate research base supporting this highly accessible and cost-effective means of providing assistance to problem drinkers who might otherwise receive no help for this important health concern.

Acknowledgments

Stand-alone versions of the AHC and the CYD were licensed under paid contract from Evolution Health Systems (the owners of these websites) for use in this study. Support to the Centre for Addiction and Mental Health for salary of scientists and infrastructure has been provided by the Ontario Ministry of Health and Long-Term Care. The views expressed in this article do not necessarily reflect those of the Ministry of Health and Long-Term Care. John Cunningham is supported as the Canada Research Chair in Brief Interventions for Addictive Behaviours. JAC had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-ehealth V.1.6.1 checklist [32].

[PDF File (Adobe PDF File), 562KB - jmir_v14i4e107_app1.pdf]

References

1. Rehm J, Patra J, Popova S. Alcohol-attributable mortality and potential years of life lost in Canada 2001: implications for prevention and policy. Addiction 2006 Mar;101(3):373-384. [doi: 10.1111/j.1360-0443.2005.01338.x] [Medline: 16499510]

2. 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]
3. 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]
4. 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]
5. Tait RJ, Christensen H. Internet-based interventions for young people with problematic substance use: a systematic review. Med J Aust 2010 Jun 7;192(11 Suppl):S15-S21. [Medline: 20528701]
6. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. Addiction 2011 Feb;106(2):267-282. [doi: 10.1111/j.1360-0443.2010.03214.x] [Medline: 21083832]
7. 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]
8. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
9. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
10. 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]
11. Doumas DM, Haustveit T. Reducing heavy drinking in intercollegiate athletes: Evaluation of a web-based personalized feedback program. Sport Psychol 2008;22:213-229.
12. Doumas DM, McKinley LL, Book P. Evaluation of two Web-based alcohol interventions for mandated college students. J Subst Abuse Treat 2009 Jan;36(1):65-74. [doi: 10.1016/j.jsat.2008.05.009] [Medline: 18657941]
13. 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]
14. Linke S, McCambridge J, Khadjesari Z, Wallace P, Murray E. Development of a psychologically enhanced interactive online intervention for hazardous drinking. Alcohol Alcohol 2008;43(6):669-674 [FREE Full text] [doi: 10.1093/alcalc/agn066] [Medline: 18693217]
15. Wallace P, Murray E, McCambridge J, Khadjesari Z, White IR, Thompson SG, et al. On-line randomized controlled trial of an internet based psychologically enhanced intervention for people with hazardous alcohol consumption. PLoS One 2011;6(3):e14740 [FREE Full text] [doi: 10.1371/journal.pone.0014740] [Medline: 21408060]
16. Dawson DA, Grant BF, Stinson FS, Zhou Y. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. Alcohol Clin Exp Res 2005 May;29(5):844-854. [Medline: 15897730]
17. Babor TF, De La Fuente MF, Saunders JB, Grant M. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care. Geneva, Switzerland: World Health Organization; 1989.
18. 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]
19. Cunningham JA, Humphreys K, Koski-Jännes A, Cordingley J. Internet and paper self-help materials for problem drinking: is there an additive effect? Addict Behav 2005 Sep;30(8):1517-1523. [doi: 10.1016/j.addbeh.2005.03.003] [Medline: 15893433]
20. Borenstein M, Rothstein H, Cohen J. SamplePower. Chicago, IL: SPSS Inc 1997.
21. Cunningham JA, Humphreys K, Kypri K, van Mierlo T. Formative evaluation and three-month follow-up of an online personalized assessment feedback intervention for problem drinkers. J Med Internet Res 2006;8(2):e5 [FREE Full text] [doi: 10.2196/jmir.8.2.e5] [Medline: 16867968]
22. Sanchez-Craig M. DrinkWise: How to Quit Drinking or Cut Down. Toronto, ON: Addiction Research Foundation; 1996.
23. Cunningham JA, Sdao-Jarvie K, Koski-Jännes A, Breslin FC. Using self-help materials to motivate change at assessment for alcohol treatment. J Subst Abuse Treat 2001 Jun;20(4):301-304. [Medline: 11672647]
24. Herie MA, Watkin-Merek L. Structured Relapse Prevention: An Outpatient Conselling Approach. 2nd edition. Toronto, ON: Centre for Addiction and Mental Health; 2006.
25. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. Patient Educ Couns 2008 Feb;70(2):193-198. [doi: 10.1016/j.pec.2007.10.003] [Medline: 18022340]
26. Room R, Bondy SJ, Ferris J. The risk of harm to oneself from drinking, Canada 1989. Addiction 1995 Apr;90(4):499-513. [Medline: 7773113]

27. 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]
28. National Institute on Alcohol Abuse and Alcoholism. Helping Patients Who Drink Too Much: A Clinician's Guide. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 2005.
29. 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]
30. Blankers M, Koeter MW, Schippers GM. Internet therapy versus internet self-help versus no treatment for problematic alcohol use: A randomized controlled trial. J Consult Clin Psychol 2011 Jun;79(3):330-341. [doi: 10.1037/a0023498] [Medline: 21534652]
31. Babor TF, Steinberg K, Anton R, Del Boca F. Talk is cheap: measuring drinking outcomes in clinical trials. J Stud Alcohol 2000 Jan;61(1):55-63. [Medline: 10627097]
32. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AHC: Alcohol Help Center
AUDIT: Alcohol Use Disorders Identification Test
AUDIT-C: Alcohol Use Disorders Identification Test-Consumption subscale
CYD: Check Your Drinking

Edited by B Bewick; submitted 21.02.12; peer-reviewed by H Riper, D Newcombe; comments to author 22.03.12; revised version received 11.04.12; accepted 10.05.12; published 01.08.12.

Please cite as:
Cunningham JA
Comparison of Two Internet-Based Interventions for Problem Drinkers: Randomized Controlled Trial
J Med Internet Res 2012;14(4):e107
URL: http://www.jmir.org/2012/4/e107/
doi:10.2196/jmir.2090
PMID:22954459

Original Paper

Public Health Interventions: Reaching Latino Adolescents via Short Message Service and Social Media

Amita N Vyas[1], MHS, PhD; Megan Landry[2], MPH; Marisa Schnider[3], MPH; Angela M Rojas[1], MS, MPH; Susan F Wood[4], PhD

[1]Maternal and Child Health, Department of Prevention and Community Health, George Washington University, Washington, DC, United States
[2]Health Behavior, Department of Prevention and Community Health, George Washington University, Washington, DC, United States
[3]Program Evaluation, Cassandra Drennon & Associates, Athens, GA, United States
[4]Jacobs Institute for Women's Health, Department of Health Policy, George Washington University, Washington, DC, United States

Corresponding Author:
Amita N Vyas, MHS, PhD
Maternal and Child Health
Department of Prevention and Community Health
George Washington University
Suite 700
2175 K Street NW
Washington, DC, 20007
United States
Phone: 1 2029943695
Fax: 1 2029943601
Email: avyas@gwu.edu

Abstract

Background: Adolescents are substantial users of short message service (SMS) and social media. The public health community now has more opportunities to reach this population with positive youth development and health messages through these media. Latinos are a growing and youthful population with significant health risks and needs. This population may benefit from SMS and social media health interventions.

Objective: To examine (1) SMS and social media utilization and behavior among Latino youth, and (2) how SMS and social media can be effectively used as a component of public health interventions focused on decreasing sexual risk taking among Latino youth.

Methods: A mixed-methods approach, using both quantitative survey data and qualitative interview data, was used to provide a robust understanding of SMS and social media use and behavior for public health interventions. We recruited 428 ninth and tenth grade, self-identifying Latino adolescents to participate in a quantitative survey. Additionally, we conducted five key informant interviews with staff and 15 youth.

Results: We found that 90.8% (355/391) of respondents had access to a mobile phone either through having their own or through borrowing or sharing one. Of those who had access to a mobile phone, 94.1% (334/355) used SMS, with 41.1% (113/275) sending and receiving more than 100 text messages per day. Of 395 respondents, 384 (97.2%) had *at least one* social media account, and the mean number of accounts was 3.0 (range 0–8). A total of 75.8% (291/384) of adolescents logged in to their account daily. Of those with a social media account, 89.1% (342/384) had a Facebook account. Youth who took the survey in English were significantly more likely than those who took it in Spanish to have access to a mobile phone ($\chi^2_1 = 5.3$; 93.3% vs 86.3%; $P = .02$); to be high-volume texters ($\chi^2_2 = 16.8$; 49.4% vs 25.3%; $P < .001$); to use the Internet daily ($\chi^2_1 = 5.0$; 76.6% vs 66.0%; $P = .03$); to have a Facebook account ($\chi^2_1 = 9.9$; 90.9% vs 79.7%; $P = .002$); and to have a greater mean number of social media accounts ($t_{387} = 7.9$; 3.41 vs 2.07; $P < .001$).

Conclusions: SMS and social media are pervasive among Latino youth. Program staff and youth perceive these as credible and essential methods of communication in the context of public health programs. Public health interventions must continue to innovate and maximize new ways to reach young people to reinforce public health messages and education.

(J Med Internet Res 2012;14(4):e99) doi:10.2196/jmir.2178

KEYWORDS

Public health interventions; SMS; short message service; social media; health behavior; Latinos; adolescents

Introduction

The rapid expansion of access to and reach of mobile technology among adolescents gives the public health community new opportunities for delivering critical and timely interventions to young people that can promote positive behavior change. It is projected that by 2025, one-quarter of the adolescent US population will be Latino [1] and that Latino youth are at greater risk for engaging in unhealthy sexual behaviors leading to sexually transmitted diseases and teen pregnancy [2,3]. Nationally, the Latino community has both higher teen pregnancy and higher birth rates than the overall US population, with 53% of Latinas getting pregnant in their teens, approximately twice the national average [1]. Although dozens of evidence-based programs aimed at reducing sexual risk taking have been identified nationally, effective behavioral interventions to reduce sexual risk taking among Latino adolescents are sparse. Thus, as mobile technology becomes increasingly adopted to promote health, the need for innovative programs dedicated to the Latino adolescent population that also incorporate short message service (SMS) and social media is significant [4-10].

Using SMS, also known as text messaging, for health promotion is particularly appealing for reaching individuals not regularly in contact with health services, and for behaviors that may be socially sensitive, such as sexual health and teen pregnancy, as they offer a confidential means of communication [11]. According to the Pew Hispanic Center, Latinos are high users of mobile technology, with 55% of Latino youth using text messages as their main method to communicate with friends [12].

In addition to using SMS, over 150 health promotion activities for sexual health use social media websites that target young people [13]. Most of these focus on sexual health in general or on the human immunodeficiency virus [14]. Facebook, Twitter, and MySpace are the most popular social media sites used for health promotion. Some of the capabilities that make these websites attractive for health promotion are public and private messaging, photo and video sharing, live updates, and the possibility of forming groups pages and including applications such as games and polls [14].

Social media, including SMS, to supplement a program, or as a stand-alone program, have been useful in improving health outcomes, increasing adherence to medication schedules and appointments, and improving patient–provider communication, health information communication, data collection, and access to health records [4,6,15]. Using SMS to reinforce key prevention concepts displayed in a program is important to broadening the effects a face-to-face intervention has on the program’s participants [15,16]. Periodic prompts and reminders have also been an effective method for encouraging and reinforcing healthy behaviors [4,17]. Increasing communication, accountability, and reinforcement created through SMS and social media can increase participants’ ability to remember changes they have learned in the program and that they should be making [4].

Although a handful of studies have focused on SMS and social media use among Latino youth, relatively few have focused on how this behavior may be integrated into public health interventions. This study used data from a larger ongoing research and demonstration project, Empowering Latino Youth Program (ELYP). The objective of this analysis was to examine (1) SMS and social media utilization and behavior among Latino youth, and (2) how SMS and social media can be effectively used as a component of public health interventions focused on decreasing sexual risk taking among Latino youth. This study was reviewed and approved by the George Washington University Internal Review Board (IRB# 011217).

Methods

We used a mixed-methods approach, using both quantitative survey data and qualitative interview data, to provide a robust understanding of SMS and social media use and behavior for public health interventions.

Quantitative Phase

Sample and Procedures

We recruited ninth and tenth grade, self-identifying Latino adolescents from 12 public high schools in Maryland, USA to participate in the ELYP. ELYP is an after-school intervention aimed at reducing sexual risk taking among Latino youth and includes both an intervention and a control arm. This is a demonstration project incorporating both quantitative and qualitative measures at baseline and three points of follow-up. The present study used baseline data from 428 youth participating in the first year of ELYP. Active parental consent was obtained prior to data collection, and youth assent was read and reviewed with youth prior to survey administration. To ensure privacy and reduce reporting bias, we administered surveys via individual laptop computers with audio-capability to youth with low literacy levels so that they could listen to each question and have it read to them if desired. Study participants had the choice to complete the survey in English or Spanish and were given US $10 gift cards for their efforts in completing the survey. On survey completion, the data were stored in an encrypted file to be read only by the survey design software (Snap Surveys Ltd, Portsmouth, NH, USA).

Instruments and Measures

We adapted most of the questions used in this study from the National Longitudinal Study of Adolescent Health [18], Youth Risk Behavior Survey [19], and the PEW Internet Research Center’s 2011 teen survey [20]. Our survey included demographic variables such as age, grade, gender, US-born or years in the United States, family structure, and parents' educational level (Multimedia Appendix 1).

The survey asked a series of questions to assess SMS and social media behavior. Youth self-reported access, utilization, reasons for use, and frequency of using a mobile phone, texting, the Internet, and social media sites. The following response categories were used to assess frequency of overall *daily SMS*: 1–10 text messages per day, 11–20 per day, 21–50 per day, 51–100 per day, 101–200 per day, and more than 200 text messages per day. Finally, participants were asked on which social media sites they had an account: Facebook, MySpace, Twitter, Yahoo!, YouTube, myYearbook, Tumblr, Google Buzz, Flickr, Ustream, other, or “I don't have my own profile or account on a social media site.”

Analysis

We used SPSS version 20 (IBM Corporation, Somers, NY, USA) for all statistical analyses. Variables were recoded and collapsed to increase the saliency of the data. Patterns of missing data were analyzed and, although data were missing for key variables, it was for less than 10% of respondents. Therefore, the data presented include all participants (n = 428), and we used pairwise deletion for the bivariate analysis to retain as many respondents as possible in the analysis. Quantitative results were analyzed to provide summary descriptive statistics and cross-tabulations. We calculated chi-square and *t* test statistics to examine differences in social media utilization by selected demographic characteristics including gender, US-born, and survey language. Only significant findings of the bivariate analysis are presented.

Qualitative Phase

Sample and Procedures

The aim of the key informant interviews and youth triads was to assess staff and youth perceptions regarding feasibility of an SMS and Facebook component in the ELYP after-school program. The qualitative component consisted of (1) key informant interviews with staff delivering the *pilot* ELYP after-school program, and (2) youth triads with ninth and 10th grade Latino youth. We recruited and interviewed 5 staff members (3 women and 2 men). A convenience sample of 15 youth who were involved in the ELYP pilot after-school program were recruited and participated in the triads. Of the 15 youth in the triads, most were female. Youth who participated in the triads received incentives such as gift cards and food for their time and cooperation.

Instruments and Measures

We conducted key informant interviews using a semistructured field guide. The field guide contained 7 predetermined questions with probes. Each interview took approximately 45 minutes to complete (Multimedia Appendix 2). Youth triads were conducted using a semistructured field guide consisting of 13 questions with additional probes (Multimedia Appendix 3). The triads lasted between 45 and 90 minutes. We conducted two triads in Spanish and three in English.

Analysis

All interviews and triads were audio-recorded and transcribed verbatim, and 2 people from the research team read the staff interview and youth triad transcripts. The team convened before and after analyzing the transcripts to agree on a series of thematic codes to describe various categories and subcategories addressed throughout the interviews and triads. The codes were then entered and applied to the transcripts using NVivo software (version 9, QSR International Pty Ltd, Southport, UK).

Results

Quantitative Phase

The study sample consisted of 428 self-identifying Latino adolescents with a mean age of 15.7 years (Table 1). The sample included slightly more girls (244/428, 57.0%); the majority of participants were in ninth grade (247/428, 57.7%); 262/428 (61.2%) took the survey in English; and 176/428 (41.1%) lived with both parents and 160/428 (37.4%) with their mother only. Less than half (150/428, 35.1%) of the sample was born in the United States, with 47.1% (112/238) of those not born in the Unites States being recent immigrants (within 0–3 years). The parent's educational level was quite low for both mother and father, with 26.6% (114/428) of mothers not having had any school or not finishing high school, and 34.3% (147/428) of fathers not having had any school or not finishing high school. Interestingly, we noted that a large percentage of youth were unsure or did not know their parent's education level.

As Table 2 shows, 355/391 (90.8%) respondents had access to a mobile phone either through having their own or through borrowing or sharing one. Of those who had access to a mobile phone, 334/355 (94.1%) used SMS, with 113/276 (40.9%) sending and receiving more than 100 text messages per day. Youth reported using their mobile phones for a variety of reasons, including to take pictures, access the Internet, send or receive instant messages, play games, and send or receive email.

A total of 98.0% (385/393) of participants had access to the Internet (Table 2). Of those with access, 280/385 (72.7%) used the Internet daily and 330/385 youth use it for schoolwork (85.7%). Of 395 respondents, 384 (97.2%) had *at least one* social media account, and the mean number of accounts was 3.0 (range 0–8). In total, 75.8% (291/384) of adolescents logged in to their account daily. Of those with a social media account, 342/384 (89.1%) had a Facebook account with the next closest account being YouTube (205/384, 53.4%). Youth accessed social media to post photos or videos, to post comments, for instant messaging, and to update their status.

Table 3 presents SMS and social media behavior by US- versus foreign-born, survey language, and gender. As shown, girls were significantly more likely to use the Internet every day ($\chi^2_1 = 3.7$; 76.2% vs 67.3%; $P = .05$) and on average had significantly more social media accounts than their male counterparts ($t_{386} = 2.8$; mean 3.11 vs 2.61; $P < .01$). With respect to nativity, US-born participants on average had significantly more social media accounts than foreign-born participants ($t_{355} = –4.7$; 3.41 vs 2.54; $P < .001$). Survey language is often a proxy measure for acculturation and yielded several significant findings. Youth who took the survey in English were significantly more likely than those who took it in Spanish to have access to a mobile phone ($\chi^2_1 = 5.3$; 93.3% vs 86.3%; $P = .02$); to be high-volume texters ($\chi^2_2 = 16.8$; 49.4% vs 25.3%; $P < .001$); to use the Internet daily ($\chi^2_1 = 5.0$; 76.6% vs 66.0%; $P = .03$); to have a Facebook account ($\chi^2_1 = 9.9$; 90.9% vs 79.7%; $P = .002$); and to have a greater mean number of social media accounts ($t_{387} = 7.9$; 3.41 vs 2.07; $P < .001$).

Table 1. Characteristics of study participants (n = 428) and their parents.

Characteristic	Participant	Father	Mother
Age (years), mean	15.7		
Gender, n (%)			
Male	178 (41.6%)		
Female	244 (57.0%)		
No response	6 (1%)		
Grade, n (%)			
9th	247 (57.7%)		
10th	174 (57.7%)		
No response	7 (2%)		
Survey language, n (%)			
English	262 (61.2%)		
Spanish	166 (38.8%)		
Family structure, n (%)			
Both parents	176 (41.1%)		
Mother	160 (37.4%)		
Other	34 (37.3%)		
No response	59 (14%)		
US-born, n (%)			
Yes	150 (35.1%)		
No	238 (55.6%)		
No response	40 (9%)		
How long in United States (years), n (%)			
0–3	112 (47.1%)		
4–10	68 (29%)		
10+	27 (11%)		
No response	31 (13%)		
Parent’s education level, n (%)			
Did not have any schooling		41 (10%)	48 (11%)
Did not finish high school		73 (17%)	99 (23%)
Graduated from high school		53 (12%)	59 (14%)
Attended vocational/technical school		6 (1%)	15 (4%)
Some college		17 (4%)	26 (6%)
Graduated from college		35 (8%)	37 (9%)
Not sure/don’t know/no response		203 (47.4%)	145 (33.9%)

Table 2. Utilization of mobile phone, Internet, and social media.

Media usage	n	%
Mobile phone access		
Own a mobile phone	291	74.2%
Have access to a mobile phone	64	16%
Do not have access to a mobile phone	37	9%
Reasons for mobile phone use		
Take pictures	316	90.6%
Access the Internet	294	84.8%
Send or receive instant messages	280	80.9%
Play games	276	80.7%
Send or receive email	23	68%
SMS[a] use	334	94.4%
Frequency of SMS use (messages/day)		
Low (1–20)	84	30%
Medium (21–100)	79	29%
High (>101)	113	40.9%
Internet access	385	98.0%
Frequency of Internet use		
Daily	208	72.7%
Less than daily	105	27.3%
Reasons for Internet use		
School work	330	85.7%
Email	299	81.7%
Get health information	115	41.6%
Social networking account	384	97.2%
Mean number of social networking sites used	3.0 (range 0–8) (394)	
Type of social networking account		
Facebook	342	89.1%
YouTube	205	53.4%
Twitter	196	51.0%
Frequency of social networking use		
Daily	291	75.8%
Less often	93	24%
Reasons for use of social networking sites		
Post photos/videos	329	87.7%
Post comments	329	86.4%
Use instant messaging	318	86.0%
Update status	320	85.6%
Send private messages	273	73.0%
Play games	166	46.6%

[a] Short message service.

Table 3. Social media behaviors by gender, nativity, and survey language.

Behavior	Gender		Nativity		Survey language	
	Female (n = 244)	Male (n = 178)	US-born (n = 150)	Foreign-born (n = 238)	English (n = 262)	Spanish (n = 166)
Access to cell phone, n (%)	206 (90.8%)	143 (90.5%)	125 (91.2%)	195 (89.9%)	224 (93.3%)	126 (86.3%)*
Frequency of SMS [a] use (messages/day), n (%)						
Low (1–20)	45 (29%)	38 (33%)	27 (27%)	52 (35.9%)	42 (24%)	41 (43%)***
Medium (1–100)	40 (26%)	36 (32%)	32 (31%)	38 (26%)	47 (27%)	30 (32%)***
High (101+)	71 (46%)	40 (35%)	43 (42%)	55 (38%)	87 (49%)	24 (25%)***
Daily Internet use, n (%)	170 (76.2%)	105 (67.3%)*	107 (77.5%)	145 (68.7%)	183 (76.6%)	93 (66%)*
Social networking account, n (%)	224 (98.3%)	154 (95.7%)	137 (98.6%)	211 (96.4%)	239 (98.8%)	140 (94.6%)**
Facebook account, n (%)	196 (86.0%)	141 (87.6%)	125 (89.9%)	184 (84.0%)	220 (90.9%)	118 (79.7%)**
Mean number of social networking accounts	3.11	2.61**	3.41	2.54***	3.41	2.07***
Daily logging in to social networking account, n (%)	174 (77.7%)	113 (73.4%)	104 (75.9%)	154 (73.0%)	186 (77.8%)	102 (72.9%)

[a] Short message service.

*P < .05, **P < .01, ***P < .001.

Qualitative Phase

Staff and youth interviews yielded several themes regarding perceptions of an SMS and Facebook component to the ELYP program, including both positive and negative aspects of their use with Latino youth. The themes pulled from the qualitative data are divided into SMS- and Facebook-related themes separately.

Positive Aspects of an SMS Program Component

Staff and youth perceived SMS to have an overall positive impact because it would be an ideal way for staff and youth to become closer and to be more connected. One staff member stated that

> *some youth don't really have adults, like positive role models, so especially a lot of them would express that they don't do anything during the weekend, so I think it would be a nice comforting thing for them [to get a text message]. And then the reminder [message] would also obviously be a reminder and be helpful for them in that aspect, but I think overall, it's that you are thinking about them.*

When youth were asked what they had texted to staff in the past, their messages mainly concerned session times, checking whether a session had been cancelled, whether food was going to be available at a session, or other questions pertaining to group projects. While youth did not think it was necessary to have content-based reminder messages, staff thought motivational messages would be helpful reinforcements for youth.

Negative Aspects of an SMS Program Component

Staff expressed the two primary concerns for including an SMS component: (1) whether all youth participating in a program would have access to a mobile phone, and (2) that it could create the expectation that staff are available all hours of a day and every day. The first concern that not all youth have a mobile phone was noted by *both* staff and youth: "not all parents are on board with their students using cell phones, or using them all the time, or maybe even the idea of [staff] texting them." Additionally, if "not every student has a cell phone...you might have a very effective tool, but if it doesn't reach out to everyone it might alienate the other students." The second concern, expressed by staff, would create an expectation that staff would be available at all hours of the day. If this were the case, one staff member discussed a possible scenario in which a student text messages a staff member when the student is in a crisis. Even if youth were told at the beginning of the program that the text messages were not to be used after hours or in times of emergency, one staff member stated that if "that kid texts and no one is there to respond, in some way I would feel bad." One staff member commented that youth should not be dependent on staff and so SMS should clearly be about logistics of the program or reinforcement of the program content and should not be personal in any way.

Overall, staff thought that youth would want to receive between 2 and 4 text messages a week from staff pertaining either to logistics of the program, or to the curriculum and motivational messages. Youth said that, while they thought the logistics reminder messages would be helpful, receiving curriculum messages would be "a little" annoying and "would be helpful, but it isn't necessary." Additionally, both staff and youth thought that after school and during the evenings were the best times to send or receive text messages.

When consulted about 1-way versus 2-way text messaging, the overwhelming majority of youth declared that they preferred 2-way text messaging. Some youth stated that they would want to be able to ask a question back if they received a text from the program, or to simply ask a question of the program in general. One youth stated

> *it would be kind of hard because you have to tell them, like "oh, I can't make it" or "do I need to take something?" Then you don't see them until like the next session and then you ask them this and that, so it's easier if you can be like, "oh I got your text, its ok, whatever."*

Positive Aspects of a Facebook Component

There are four major concepts related to positive aspects of Facebook. First, according to *both* staff and youth, access is widespread and all youth have a Facebook page. One staff member commented that "even the ones who don't have a cell phone definitely check it at their friends house or maybe their own house, because usually parents are definitely willing to buy a computer but not so much give them a cell phone." Second, youth with a Facebook page reported checking their page daily, which was also noted by staff. The third concept is about ease of communication. Staff noted that "[Facebook] is the most consistent way that you could be able to communicate with [youth]," which in turn would keep them engaged in the program. Many staff mentioned that being able to post events and opportunities for youth to gain service learning hours, potential field trips, pictures from past activities, and videos of youth in the program would keep the youth engaged. Youth reported that they would Like or join a program-specific Facebook page, and they would also want to make comments on the page. The final positive aspect centered on the open access of Facebook, which provides youth the opportunity to be more honest and open about their thoughts and feelings, which according to staff not only has the potential of bringing a group of young people closer together, but also offers staff insight into the youth's context.

Negative Aspects of a Facebook Component

The major concern among staff about a health program-specific Facebook page was the possibility that it would enable cyberbullying of youth who joined the page, and that youth would post inappropriate pictures, videos, or comments to the Facebook page if it was not highly regulated and controlled by the overseeing organization. In addition, staff noted that not every student would be happy with the program all the time, "so if they get really frustrated or if they get kicked out of the program because of attendance, they could express some of that" on the Facebook page; in short, it is "a platform for [youth] to lash out." Additionally, while youth had responded that they would initially join a program-specific Facebook page, they stated that if they were unable to post comments or pictures, or have a personal interaction with the Facebook page, then they would not check it as often.

Discussion

Harnessing new technology to reach youth with positive and healthy messages could be extremely powerful, and therefore it is critical to fully understand how adolescents from different communities and cultures use technology. This study found results comparable with those of the Pew Hispanic Center teen survey on mobile and social media use [12]. The Pew report found that the use of mobile technology greatly differs among Latino teens depending on nativity and language, with 65% of native-born teens saying they communicated with friends using text messaging versus 26% of foreign-born teens, and 68% of English-dominant and 50% of bilingual young Latinos using text messaging daily to communicate with friends, compared with only 19% of Spanish-dominant Latinos. Similar to that report, the present study found significant differences by survey language and nativity, with 93% of youth who took the survey in English versus 86% of youth who took it in Spanish having access to a cell phone. Further, 42% of US-born youth versus 38% of foreign-born youth had a higher frequency of SMS use, and 49% of those who took the survey in English versus 25% who took it in Spanish had a higher SMS usage frequency.

The Pew report also found that 23% of young Latinos used social media sites such as Facebook or MySpace daily to communicate with their friends, and native-born Latinos were more likely to use social media sites. Similarly, this study found that 97% had a social media account and 89% had a Facebook account. Further, 91% of youth who took the survey in English versus 80% who took it in Spanish had a Facebook account, and 90% of US-born youth versus 84% of foreign-born youth had a Facebook account.

Overall, the qualitative findings were very positive with respect to use of SMS and social media for public health interventions. Although several negative aspects were mentioned, most could be addressed through appropriate system design and use by program staff.

The study also found that Facebook is slightly more feasible than SMS as a method of communicating health messages due to issues surrounding lack of access to mobile phones and the youth participants' overwhelming access to and usage of Facebook. Due to the majority of youth being non-US-born, even though they completed the survey in English, each component should be available in English and Spanish, and students should be able to opt in to either language.

The differences found in survey language indicate that Spanish-language respondents had significantly less access and less frequent use of social media. Furthermore, those who took the survey in Spanish were more likely to be recent immigrants, within the last 3 years. To address these significant and disadvantageous differences, programs must ensure that any social media component is available in both languages so as not to isolate the Spanish-speaking youth from this aspect of the intervention.

Since mobile phone access may be a financial issue, programs may need to expand efforts using free social media sites such as Facebook, although this does require access to a computer. Our findings indicate that use of Facebook is extremely common and may be particularly useful with youth who do not have access to SMS. If youth are simply unaware of these social media because of recent immigration, they can be encouraged to sign up for an account, or alternative strategies will need to be developed to reach these adolescents with the same information.

Although students expressed the desire to have 2-way messages, it may be more feasible for public health programs to have 1-way messages in an effort to not overwhelm staff. Youth need to be educated about the use of SMS and social media within the context of a program to increase their understanding that the program is sending them information and is not a 24-hour service, nor a means to contact the organization if there is a crisis. Rather, youth should be given phone numbers for local organizations that are equipped to handle crises such as 911, a local teen suicide hotline, and other local organizations as necessary to avoid liability issues.

Strengths and Limitations

This study focused on a new and growing youth population in the United States with significant public health needs. However, there are a few limitations of the current study. This was a cross-sectional sample of youth and therefore it is not possible to examine causal relationships between nativity or language use and SMS or social media behavior. Additionally, although nativity and language are frequently used as proxy measures for acculturation, future studies should use more salient and robust measures that capture the complex and dynamic process of acculturation. Finally, this study comprised a subsample of youth participating in a research and demonstration project (ELYP) and may not be fully representative of a national sample of Latino youth. However, this sample consisted of Latino youth from Central America, primarily El Salvador and Peru, which is representative of the growing Latino population in Maryland. Further, future studies ought to examine cultural and geographic differences around SMS and social media use, as there are certainly distinct characteristics and traditions between Latino subgroups.

Conclusions

SMS and social media are pervasive among Latino youth, and program staff and youth perceive these as credible and essential methods of communication in the context of public health programs. In fact, the frequency of use is so significant that public health interventions must continue to innovate and maximize new ways to reach young people to reinforce public health messages and education. Furthermore, given the differences found in this study by nativity and survey language, future research must delve deeper to understand differences in utilization and behavior by acculturation and other cultural factors such as parenting and family influences. Given the substantial role of parenting and family influences on adolescents living in immigrant families, and these differences by acculturation, future research must explore how these factors may mediate SMS and social media utilization and behavior.

Acknowledgments

This study was supported by a grant from the US Department of Health and Human Services, Office of Adolescent Health (grant # TP2AH000015). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US Department of Health and Human Services or the Office of Adolescent Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Quantitative survey administered to Latino youth.

[PDF File (Adobe PDF File), 70KB - jmir_v14i4e99_app1.pdf]

Multimedia Appendix 2

Key informant field guide for staff working with Latino youth.

[PDF File (Adobe PDF File), 38KB - jmir_v14i4e99_app2.pdf]

Multimedia Appendix 3

Triad field guide for Latino youth.

[PDF File (Adobe PDF File), 44KB - jmir_v14i4e99_app3.pdf]

References

1. The National Campaign to Prevent Teen Pregnancy. TheNationalCampaign.org, StayTeen.org, blog.TheNationalCampaing.org. 2008 May. A Look at Latinos: An Overview of Latina Teen Pregnancy & Birth Rates URL: http://www.thenationalcampaign.org/espanol/PDF/latino_overview.pdf [accessed 2012-06-27] [WebCite Cache ID 68jqQyQVC]
2. Trejos-Castillo E, Vazsonyi AT. Risky sexual behaviors in first and second generation Hispanic immigrant youth. J Youth Adolesc 2009 May;38(5):719-731. [doi: 10.1007/s10964-008-9369-5] [Medline: 19636766]
3. Anonymous. Young People at Risk: HIV/AIDS Among America's Youth. Atlanta, GA: Centers for Disease Control and Prevention; 2005.

4. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010 Apr;32(1):56-69. [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
5. Donner J. Research approaches to mobile use in the developing world: a review of the literature. Inf Soc 2008;24(3):140-159.
6. Lim MS, Hocking JS, Hellard ME, Aitken CK. SMS STI: a review of the uses of mobile phone text messaging in sexual health. Int J STD AIDS 2008 May;19(5):287-290. [doi: 10.1258/ijsa.2007.007264] [Medline: 18482956]
7. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
8. Atun RA, Sittampalam SR. A Review of the Characteristics and Benefits of SMS in Delivering Health Care: The Role of Mobile Phones in Increasing Accessibility and Efficiency in Healthcare. The Vodafone Policy Paper Series 4, no. 4. Berkshire, United Kingdom: Vodafone Group Plc; 2006:18-28.
9. Leach-Lemens C. Using mobile phones in HIV care and prevention. HIV AIDS Treat Pract 2009;137:2-8.
10. Kaplan WA. Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? Global Health 2006;2:9 [FREE Full text] [doi: 10.1186/1744-8603-2-9] [Medline: 16719925]
11. Gold J, Lim MS, Hellard ME, Hocking JS, Keogh L. What's in a message? Delivering sexual health promotion to young people in Australia via text messaging. BMC Public Health 2010;10:792 [FREE Full text] [doi: 10.1186/1471-2458-10-792] [Medline: 21190584]
12. Livingston G, Lopez MH. How young Latinos communicate with friends in the digital age. Washington, DC: Pew Hispanic Center; 2010 Jul 28. URL: http://www.pewhispanic.org/2010/07/28/how-young-latinos-communicate-with-friends-in-the-digital-age/ [accessed 2012-06-19] [WebCite Cache ID 68XpeoQWH]
13. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. BMC Public Health 2011;11:583 [FREE Full text] [doi: 10.1186/1471-2458-11-583] [Medline: 21777470]
14. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
15. Cornelius JB, St Lawrence JS. Receptivity of African American adolescents to an HIV-prevention curriculum enhanced by text messaging. J Spec Pediatr Nurs 2009 Apr;14(2):123-131. [doi: 10.1111/j.1744-6155.2009.00185.x] [Medline: 19356206]
16. Wu Y, Stanton BF, Galbraith J, Kaljee L, Cottrell L, Li X, et al. Sustaining and broadening intervention impact: a longitudinal randomized trial of 3 adolescent risk reduction approaches. Pediatrics 2003 Jan;111(1):e32-e38. [Medline: 12509592]
17. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
18. Harris KM. The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II-1996; Wave III-2002; Wave IV-2009. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill; 1994.
19. US Centers for Disease Control and Prevention. 2012 Jun 11. Youth Risk Behavior Surveillance System URL: http://www.cdc.gov/HealthyYouth/yrbs/index.htm [accessed 2012-06-27] [WebCite Cache ID 68jqbYM4I]
20. Lenhart A. Teens, smartphones & texting. Washington, DC: Pew Internet & American Life Project; 2012 Mar 19. URL: http://pewinternet.org/~/media//Files/Reports/2012/PIP_Teens_Smartphones_and_Texting.pdf [accessed 2012-06-19] [WebCite Cache ID 68XplDvI0]

Abbreviations

ELYP: Empowering Latino Youth Program
SMS: short message service

Edited by G Eysenbach; submitted 18.05.12; peer-reviewed by G Eysenbach; comments to author 11.06.12; revised version received 12.06.12; accepted 13.06.12; published 12.07.12.

Please cite as:
Vyas AN, Landry M, Schnider M, Rojas AM, Wood SF
Public Health Interventions: Reaching Latino Adolescents via Short Message Service and Social Media
J Med Internet Res 2012;14(4):e99
URL: http://www.jmir.org/2012/4/e99/
doi:10.2196/jmir.2178
PMID:22789678

Original Paper

The Effects of Combining Web-Based eHealth With Telephone Nurse Case Management for Pediatric Asthma Control: A Randomized Controlled Trial

David Gustafson[1], PhD; Meg Wise[1], MLS, PhD; Abhik Bhattacharya[2], CSSBB, PhD; Alice Pulvermacher[1], MS; Kathleen Shanovich[3], RN, MS; Brenda Phillips[4], MS; Erik Lehman[4], MS; Vernon Chinchilli[4], PhD; Robert Hawkins[5], PhD; Jee-Seon Kim[6], PhD

[1]Center for Health Enhancement Systems Studies, University of Wisconsin-Madison, Madison, WI, United States
[2]Blue Health Intelligence, Chicago, IL, United States
[3]Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, United States
[4]Department of Public Health Sciences, Penn State Hershey College of Medicine, Hershey, PA, United States
[5]Department of Journalism and Mass Communication, University of Wisconsin-Madison, Madison, WI, United States
[6]Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
David Gustafson, PhD
Center for Health Enhancement Systems Studies
University of Wisconsin-Madison
1513 University Avenue, Room 4109
Madison, WI, 53706
United States
Phone: 1 608 263 4882
Fax: 1 608 890 1438
Email: dhgustaf@wisc.edu

Abstract

Background: Asthma is the most common pediatric illness in the United States, burdening low-income and minority families disproportionately and contributing to high health care costs. Clinic-based asthma education and telephone case management have had mixed results on asthma control, as have eHealth programs and online games.

Objectives: To test the effects of (1) CHESS+CM, a system for parents and children ages 4–12 years with poorly controlled asthma, on asthma control and medication adherence, and (2) competence, self-efficacy, and social support as mediators. CHESS+CM included a fully automated eHealth component (Comprehensive Health Enhancement Support System [CHESS]) plus monthly nurse case management (CM) via phone. CHESS, based on self-determination theory, was designed to improve competence, social support, and intrinsic motivation of parents and children.

Methods: We identified eligible parent–child dyads from files of managed care organizations in Madison and Milwaukee, Wisconsin, USA, sent them recruitment letters, and randomly assigned them (unblinded) to a control group of treatment as usual plus asthma information or to CHESS+CM. Asthma control was measured by the Asthma Control Questionnaire (ACQ) and self-reported symptom-free days. Medication adherence was a composite of pharmacy refill data and medication taking. Social support, information competence, and self-efficacy were self-assessed in questionnaires. All data were collected at 0, 3, 6, 9, and 12 months. Asthma diaries kept during a 3-week run-in period before randomization provided baseline data.

Results: Of 305 parent–child dyads enrolled, 301 were randomly assigned, 153 to the control group and 148 to CHESS+CM. Most parents were female (283/301, 94%), African American (150/301, 49.8%), and had a low income as indicated by child's Medicaid status (154/301, 51.2%); 146 (48.5%) were single and 96 of 301 (31.9%) had a high school education or less. Completion rates were 127 of 153 control group dyads (83.0%) and 132 of 148 CHESS+CM group dyads (89.2%). CHESS+CM group children had significantly better asthma control on the ACQ (d = –0.31, 95% confidence limits [CL] –0.56, –0.06, P = .011), but not as measured by symptom-free days (d = 0.18, 95% CL –0.88, 1.60, P = 1.00). The composite adherence scores did not differ significantly between groups (d = 1.48%, 95% CL –8.15, 11.11, P = .76). Social support was a significant mediator for

CHESS+CM's effect on asthma control (alpha = .200, *P* = .01; beta = .210, *P* = .03). Self-efficacy was not significant (alpha = .080, *P* = .14; beta = .476, *P* = .01); neither was information competence (alpha = .079, *P* = .09; beta = .063, *P* = .64).

Conclusions: Integrating telephone case management with eHealth benefited pediatric asthma control, though not medication adherence. Improved methods of measuring medication adherence are needed. Social support appears to be more effective than information in improving pediatric asthma control.

Trial Registration: Clinicaltrials.gov NCT00214383; http://clinicaltrials.gov/ct2/show/NCT00214383 (Archived by WebCite at http://www.webcitation.org/68OVwqMPz)

(J Med Internet Res 2012;14(4):e101) doi:10.2196/jmir.1964

KEYWORDS

Asthma; asthma information; childhood disease; case management; patient education; eHealth; social support

Introduction

Asthma and other chronic diseases pose a great risk as the United States attempts to decrease its health care costs. Research suggests that the active ingredients of chronic disease management include long duration [1,2], assertive outreach [3], monitoring [4-6], prompts [7-9], action planning [10-12], case management [13-15], and peer [16-18] and family [19-21] support. All of these can potentially be provided by information and computer-based technologies and sensors.

Poor control of pediatric asthma affects low-income and minority children disproportionately and contributes to more than 14 million lost school days and 3 million lost parental workdays per year [22]. Daily controller medications can manage even severe asthma [23], but adherence is low, especially for the underserved [24]. Asthma education programs, with their low participation rates, have had mixed results [23]. Nurse case management can reduce asthma-related emergency care and hospitalization costs, but it is expensive [25,26]. Child-focused, Web-based asthma education and games (eHealth) have improved knowledge, asthma control, and medication adherence in the short run [27-29]. However, parents tend to overestimate their child's medication-taking skills and actual adherence [30]. This suggests that parental involvement might be beneficial for managing pediatric asthma. Integrating phone-based clinician care into asthma eHealth programs for adults has shown significant promise in behavioral and asthma outcomes [31]. Interventions such as CHESS+CM, based on self-determination theory [32] and self-efficacy [33], provide information, social support, and skill-building tools for self-managing the disease. These interventions, which aim to increase confidence, competence, relatedness, and autonomous motivation, have been used successfully in asthma education programs [28,29,32,34]. However, the factors associated with these theories have not been tested for their mediational effects on adherence to controller medications or asthma control. Understanding this is important for developing an asthma eHealth system that balances the various functions—information, social support, and skill building—to the best effect for children and their parents.

We, therefore, hypothesized that a parent-focused intervention that integrates monthly telephone nurse case management with a comprehensive, interactive asthma eHealth program could improve asthma control and medication adherence. We surmised that these effects would be mediated by social support, self-efficacy, and asthma information competence. This paper reports the results of a randomized controlled trial funded by the US National Institute of Nursing Research.

Methods

Intervention

The year-long intervention called CHESS+CM consisted of an eHealth program, Comprehensive Health Enhancement Support System (CHESS), and a monthly telephone call to the parent from an asthma nurse case manager (CM). CHESS is an umbrella name for several eHealth systems developed and tested for the past 25 years at the University of Wisconsin-Madison. CHESS modules provide information, adherence strategies, decision-making tools, and support services in attractive, easy-to-use formats. The most important strength of CHESS modules may be the closed, guided universe of tailored information and support in an integrated package with efficient navigation, eliminating the need for complicated search and discovery. In randomized efficacy trials, CHESS modules significantly improved quality of life and self-efficacy for women with breast cancer versus control and Internet groups, and quality of life and costs of care in people infected with the human immunodeficiency virus [35,36]. CHESS programs also have demonstrated the effectiveness of using self-determination and self-efficacy theories to improve information competence, health care participation, and social support among cancer patients [35,37]. The CHESS module used in this study was designed specifically for asthma.

The project was carried out with a University of Wisconsin-based team of educators, pharmacists, and nurse practitioners specializing in asthma rather than with staff from the five managed care organizations (MCOs) from which participants were recruited. University of Wisconsin-based nurse practitioners also monitored the progress of recruitment. The project director monitored the discussion group within CHESS to ensure that calls for help were rapidly addressed and that inaccurate information was not shared. The full trial protocol is available at http://www.webcitation.org/69E2cXZbo.

As Figure 1 shows, CHESS for asthma had three audiences: parents, children, and case managers. Parents received comprehensive information based on the National Asthma Education and Prevention Program guidelines [21,22,38], a peer discussion group, case manager email, and the Asthma Coach, which assesses the child's asthma and the parent's and child's well-being. CHESS provided tailored feedback and links to salient CHESS content and other interactive tools. Children received simplified information in game and audiovisual formats, as well as social support via a peer discussion group and personal stories. No major bug fixes or downtimes occurred. Asthma-related content was updated annually over the course of the study. Otherwise, no major modifications were made to the system. No significant secular events took place during the study period.

The case manager received tools to schedule monthly phone calls with the parent, view parents' Asthma Coach entries, enter phone call notes, and send and receive case management mail to and from the parent, as well as a "prescription pad" to place CHESS resources on the parent's home page [21]. Monthly case management calls to the parent assessed the child's asthma, medication adherence, and psychosocial challenges, and provided relevant education and support. On completing a call, the case manager entered notes in the case management toolbox and then sent the parent a summary via case manager email with links to salient CHESS resources, which appeared on the parents' CHESS home page, as shown in Figure 2 [21]. These features were designed with user input for content and usability. For a more complete description of the CHESS asthma module and its development, see Wise et al [21]. The module is available at https://chess.wisc.edu/asthmamobile/. The code name is *uwmadison* and the password is *testing*. Screenshots of the program are available on request.

Figure 1. Comprehensive Health Enhancement Support System (CHESS) services for parent, case manager, and child. CM = case management.

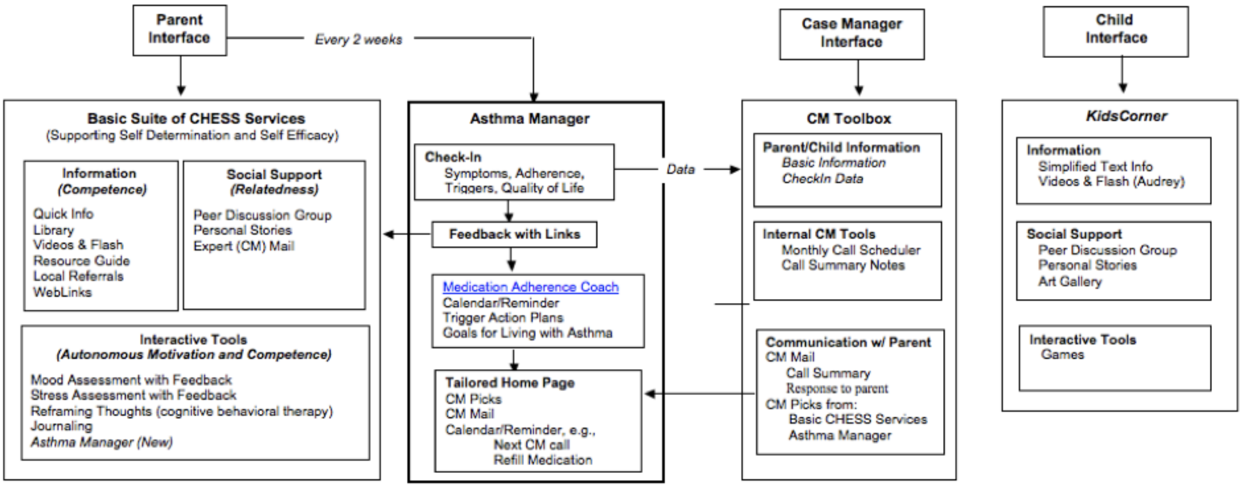
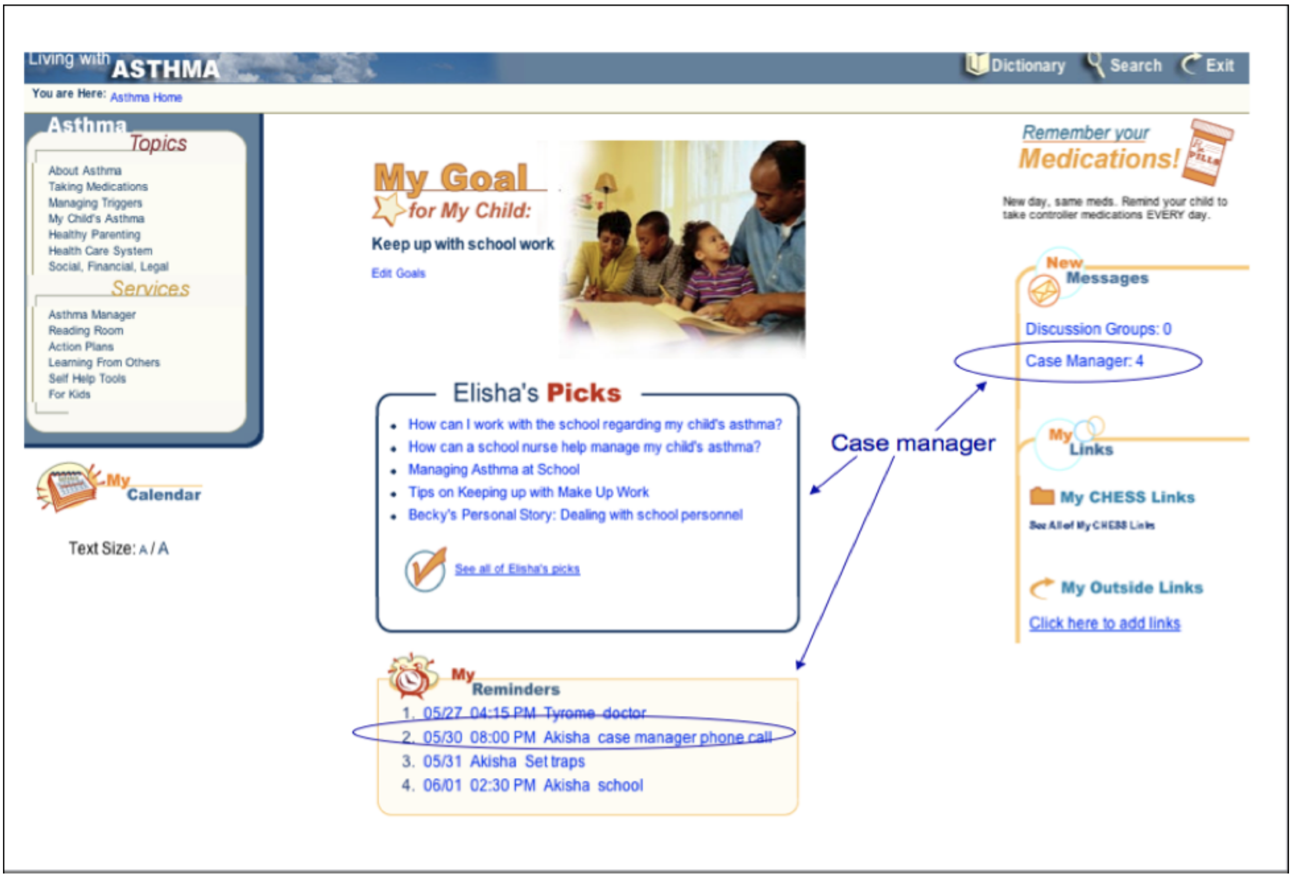


Figure 2. Parent's home page of the Comprehensive Health Enhancement Support System (CHESS) asthma module.



New day, same meds. Remind your child to take controller medications EVERY day.

- How can I work with the school regarding my child's asthma?
- How can a school nurse help manage my child's asthma?
- Managing Asthma at School
- Tips on Keeping up with Make Up Work
- Becky's Personal Story: Dealing with school personnel

1. 05/27 04:15 PM Tyrome doctor
2. 05/30 08:00 PM Akisha case manager phone call
3. 05/31 Akisha Set traps
4. 06/01 02:30 PM Akisha school

Invitation and Recruitment

Eligible participants were parents or other adults functioning as parents, such as grandmothers, who were able to read at a sixth-grade level and had children ages 4–12 years with a diagnosis of asthma (*International Classification of Diseases*, 9th revision code 493) or wheezing (code 786.07); a prescribed asthma controller medication; and poor medication adherence, defined as having missed more than one medication refill or having an emergency department visit or hospitalization because of poor asthma control. Originally children were identified through the health care utilization and pharmacy claim databases at four MCOs (MCOs 1–4) and the Wisconsin Medicaid Program from one urban–rural county (Dane County, which is also the home of the University of Wisconsin-Madison) and seven surrounding rural counties (Columbia, Dodge, Green, Iowa, Jefferson, Rock, and Sauk). MCO 5 in Milwaukee was added after it became clear that MCOs 1–4 could not produce enough participants with poorly controlled asthma. MCO 5 served an entirely Medicaid population in Milwaukee County and had the state's highest rates of asthma-related emergency department visits and overnight hospital stays [39].

All research materials and procedures were approved by the University of Wisconsin's Health Sciences Review Board, as well as by the ethics and review boards for each of the five MCOs. Recruitment was initiated by a letter from the MCO or Wisconsin Medicaid Program to parents of study-eligible patients with an opt-in or an opt-out card regarding a phone call from the study nurse, depending on the organization's institutional review board policies [40]. Recruiters screened for eligibility, described the study (as a comparison of two approaches to asthma control) and its risks and benefits, and scheduled an intake interview for people who agreed to participate.

The study had four key risks. (1) The key risk for participants was the possibility of their replacing effective medical care with unproven treatments. To reduce this risk, each survey asked whether parts of the child's medical treatment had been abandoned against medical advice. In addition, we scanned discussion group entries in CHESS. If we found indications of abandonment, we contacted parents to express our concern about potential risks. (2) Participants' misinterpreting information in CHESS posed another risk. The information was presented at a sixth-grade reading level and screened by asthma experts to reduce this risk. Disclaimers also cautioned that the computer is not a substitute for seeking medical attention and that comments in the social media may not be factual. (3) To reduce the risk of anonymity being breached, participants were assigned a blind code number. All data had names removed and code numbers attached. (4) Participants' divulging confidential information was another risk in the study. We frequently warned users about this, and we used digital signatures to warn users if CHESS was altered.

Enrollment

Parents and children participated in a prerandomization intake appointment with a study case manager at asthma clinics associated with MCOs 1–4 and, in MCO 5, at a community center. The study team traveled twice a month to the community center to conduct intakes after school and into the early evening

[40]. At all locations, childcare, snacks, and transportation were provided as needed. Parents were asked to bring the child's medications to the intake interviews. Intakes lasted 45–90 minutes. Potential participants were informed of (1) the nature and purpose of the study, (2) the financial compensation offered (US $35 for completing each of four surveys and providing 2 spirometry readings), (3) the types of data to be collected from claims records, surveys, and computer-use tracking, (4) the intervention to be given to the experimental group, (5) the nature and reasons for random assignment, (6) the measures used to insure the confidentiality of data collected, (7) the timeline of the study, (8) the duration of the study (12 months, to capture the seasonality of asthma), and (9) the University of Wisconsin-Madison being the research organization for the project. Consent was documented by obtaining signed, institutional review board-approved consent forms containing all of the above information. The consent forms were kept in a locked file at CHESS. Intake appointments assessed parental ability to read at a sixth-grade level by asking parents to read aloud the consent letter. The appointments also included recording medications and doses, a spirometry test, and the child's asthma history. A researcher administered a pretest survey with training on completing the asthma diary and provided individualized asthma education as needed. Regardless of study arm, case managers notified MCO staff about children with uncontrolled asthma for further evaluation.

Randomization

Researchers at the University of Wisconsin generated the random allocation sequence. Nurses conducting consent, assent, and pretests were given sequentially numbered envelopes containing the random assignment for each participants. It was not possible to blind participants or outcome assessors. We did blind those analyzing the data.

Participants were equally randomized according to their MCO and then blocked by severity and by Medicaid status. Randomization occurred upon receipt of the run in diaries for MCO 1-4 subject and after just the intake for MCO 5 participants [40]. The CHESS+CM group received a 45-minute training session. Laptop computers, land phone lines, and Internet service were provided, as needed. MCOs 1–4 and Wisconsin Medicaid Program participants received one-on-one training at home on an Internet-enabled computer with the live CHESS program. MCO 5 participants received group training at the community center where they had had their intake interview. Because that center lacked Internet access and most participants borrowed study laptops, training on using the laptop and CHESS was guided by an interactive compact disc. All users were instructed to use CHESS whenever they wished. No minimum expectations were placed on users.

All participants, regardless of study condition, received a call from the project manager 1 week after randomization to see how things were going. They also received with their mailed surveys at 3, 6, 9, and 12 months a packet of educational materials about asthma control, child development, parenting, and community resources. Parents and children returned to the clinic or community center for an exit interview that included taking the same measures used at the intake appointment. Exit interviews were conducted at home for 40 difficult-to-reach MCO 5 parents.

Measures

Asthma control was measured via two self-reports collected at baseline and at 3, 6, 9, and 12 months. Juniper and colleagues' Asthma Control Questionnaire (ACQ) was administered orally at the intake and exit interviews and self-administered via mailed surveys at the intervening time points [41]. This well-validated, 6-item, 7-point Likert scale measured daytime and nocturnal symptoms, missed school, and rescue medication use over the previous 7 days [41]. Lower scores indicated better asthma control. Scales were created as a mean of all items. Surveys with missing ACQ items were not calculated and counted as missing data. Symptom-free days were calculated from 2-week diaries as the absence of asthma-related nighttime awakenings, daytime symptoms, bronchodilator use, unscheduled physician visits, or school absences [42]. The study included a run-in period of 3 weeks before randomization during which participants kept asthma diaries that were used to collect baseline data on symptom-free days.

Adherence to asthma controller medications was measured as the mean composite score from self-reported 2-week asthma diaries at baseline and 3, 6, 9, and 12 months [42]. The ratio of medication to possession was calculated from MCO and Wisconsin Medicaid Program pharmacy claims data as the actual versus expected prescription refill rate for each participant's prescription.

Social support, self-efficacy, and information competence were adapted from 5-point Likert scales used successfully in prior CHESS studies [35]. The 6-item social-support scale assessed the availability and reliability of social, emotional, and instrumental support. The 6-item self-efficacy scale assessed asthma problem-solving skills and strategies, along with perceived competence, goal attainment, and comparison with others' skills. The 4-item information competence scale assessed parents' understanding of asthma information needs, difficulty in obtaining such information, satisfaction with their strategies, and level of control using the information. Scales were created as a mean of all items. If more than 1 item was missing, the scale score was not computed.

Statistical Analysis

Our study was powered based on an expected 320 dyads completing the trial; 259 completed the study. We anticipated a dropout rate of 20% but had a rate of 14% (42/301).

Descriptive statistics established baseline characteristics for the CHESS+CM and control groups and for participants with missing data. We used 1-way analysis of variance to compare differences between the CHESS+CM and control groups and between dropouts and those who completed the study. For missing 3-, 6-, 9-, and 12-month posttest data, we used a general linear model procedure to test for equality of mean scores meeting the baseline criteria followed by a 1-way analysis of variance to determine the *P* value for the general linear model. Scores for outcome variables with no significant mean differences between 3, 6, 9, and 12 months were averaged to create a score for the entire intervention time [43].

Main outcomes were an intent-to-treat analysis with a repeated-measures, mixed model to account for the correlation between the five time points within participants, and to analyze the differences between the time points and baseline within and between the control and CHESS+CM groups. A Bonferroni adjustment for multiple comparisons yielded adjusted *P* values and confidence limits for mean estimates within each set of comparisons at the four time periods. Finding no significant differences, we averaged participants' 3-, 6-, 9-, and 12-month data to create a score for the whole intervention time following the procedure described above. Comparisons of the change from baseline were made over the average of all posttest time points within the CHESS+CM group and the control group and between the two groups.

A multiple regression model, as described by MacKinnon [44], was used to analyze mediators for outcome variables that showed significant difference between the CHESS+CM and control groups.

Step 1 determined significant covariates by loading candidate variables for their premediated effects of the intervention tau on the outcome variable. Covariates included the outcome variable's pretest score, Medicaid status, age, time since diagnosis, and asthma severity, as well as the parent's age and education level and whether help with parenting was available from another adult. Covariates in step 2 included the pretest scores for the outcome and mediator variables and any significant covariate.

Step 2 used a multiple regression model for each mediator, whereby alpha = CHESS+CM's effect on the mediator, beta = mediator's effect on the dependent variable, and tau^1 = CHESS+CM's effect on the dependent variable after the mediational test.

Results

Recruitment and Enrollment

As Figure 3 shows, a total of 1988 invitation letters were mailed, 702 recruitment or eligibility screening calls were completed, and 305 parent–child dyads enrolled in the study. The enrollees accounted for 15.34% of the 1988 letters sent and 43.5% of the 702 completed phone calls. Data were collected from August 2, 2004 through August 16, 2007.

Reasons for nonenrollment were unable to be reached by phone, not eligible, did not have moderate to severe asthma, and too busy. Up to 3 did not come for their scheduled intake appointments. A total of 4 dyads dropped out after the intake but before randomization; thus, 301 were randomly assigned: 153 to the control group and 148 to the CHESS+CM group. Finally, 259 dyads (86.1%) completed the study. After randomization, 42 dropped out: 26 (17%) from the control group and 16 (11%) from the CHESS+CM group. The between-group dropout rate was not significant (*P* = .12). However, participants who dropped out were significantly less likely to be white or married, and more likely to be significantly younger, have lower asthma quality of life, and have less education. Children of dropouts had no significant differences in baseline ACQ scores, but had significantly lower pharmacy refill rates and more asthma-related school absences.

Response Rates and Data Availability

Available data rates, shown in Table 1, were highest for self-reported data at baseline and 12 months, which involved direct interaction with a researcher. For example, the ACQ response rate was 98.7% at baseline and 82.7% at 12 months, but 58.1%, 52.8%, and 49.5% at the intervening time points. Missing pharmacy refill data, however, were highest at 12 months.

Baseline Characteristics of Control and Intervention Groups

Table 2 shows no significant differences at baseline between the control and CHESS+CM groups for demographics, asthma status, Web use, or the mean outcome and mediator scores.

Intervention Effects on Main Outcomes

Table 3 shows the mean difference in scores for the outcome variables between baseline and the mean scores measured at 3, 6, 9, and 12 months for the control group, the CHESS+CM group, and the difference between the CHESS+CM and the control group.

Symptom-free days as measured from asthma diaries improved significantly for the CHESS+CM group (odds ratio 1.38, *P* = .01) and less so for the control group (odds ratio 1.20, *P* = .29), but there were no significant between-group differences (odds ratio 0.18, *P* = 1.00). Asthma control as measured on the ACQ improved significantly for CHESS+CM (–0.42 on a 7-point Likert scale with lower scores indicating better control, *P* = .001) and not significantly for the control group (–0.11, *P* = .22). The between-group difference (–0.31) was significant (*P* = .01). The composite medication adherence score did not change significantly either within group or between the groups, with a 0.58% increase (*P* = .87) for the control group and 2.06% increase (*P* = .55) for the CHESS+CM group, and a 1.48% between-group difference (*P* = .76). Both groups reported declining adherence from diaries and had significant improvements in medication refills.

Figure 3. CONSORT diagram of sample pool, recruitment, and participation.

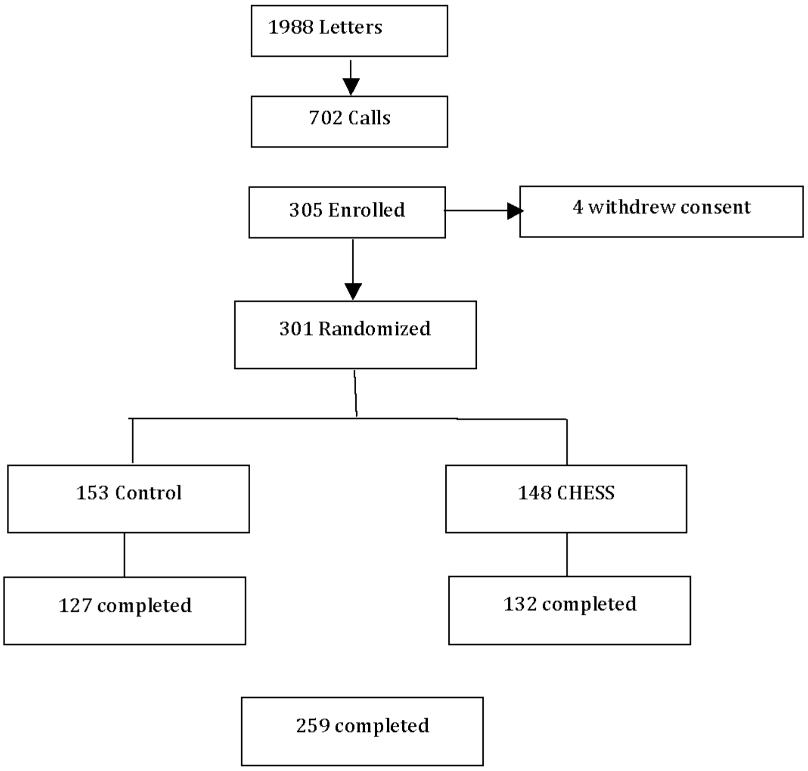


Table 1. Response rate and available data at study time points.

Measure	CHESS+CM[a] (n = 148)	Control (n = 153)	Total (n = 301)
Asthma Control Questionnaire			
Baseline	145 (98.0%)	152 (99.3%)	297 (98.7%)
3 months	94 (64%)	81 (53%)	175 (58.1%)
6 months	85 (57%)	74 (48%)	159 (52.8%)
9 months	86 (58%)	63 (41%)	149 (49.5%)
12 months	128 (86.5%)	121 (79.1%)	249 (82.7%)
Diary data [b]			
Baseline	128 (86.5%)	123 (80.4%)	251 (83.4%)
3 months	91 (61%)	79 (52%)	170 (56.5%)
6 months	84 (57%)	73 (48%)	157 (52.2%)
9 months	84 (57%)	65 (42%)	149 (49.5%)
12 months	108 (73.0%)	110 (72.9%)	218 (72.4%)
Pharmacy claims data			
Baseline	103 (69.6%)	102 (66.7%)	205 (68.1%)
3 months	64 (43%)	64 (42%)	128 (42.5%)
6 months	78 (53%)	75 (49%)	153 (50.8%)
9 months	70 (47%)	77 (50%)	155 (48.8%)
12 months	59 (40%)	60 (39%)	119 (39.5%)

[a] Comprehensive Health Enhancement Support System plus monthly nurse case management.

[b] Diary data measured symptom-free days and self-reported medication adherence.

Table 2. Mean baseline values for demographics and main outcome and mediator measures.

Characteristic	Control (n = 153)	CHESS+CM[a] (n = 148)	*P* value
Dropout, n (%)	26 (17%)	16 (11%)	.12
Children's characteristics			
Male gender, n (%)	87 (57%)	97 (66%)	.12
Mean age (years), mean (SD)	8.18 (2.45)	7.65 (2.61)	.11
Mean age at first asthma diagnosis (years), mean (SD)	3.16 (2.57)	2.79 (2.45)	.21
African American, n (%)	84 (55%)	87 (59%)	.39
Medicaid, n (%)	80 (52%)	74 (50%)	.69
Yes to: n (%)			
Prednisone at least once[b]	96 (63%)	101 (68%)	.33
Hospital stays for asthma[b]	28 (18%)	22 (15%)	.51
Emergency department for asthma[b]	78 (51%)	84 (57%)	.37
Unplanned clinic visit[b]	116 (75.8%)	110 (74.3%)	.71
Intensive care unit for asthma[b]	8 (5%)	4 (3%)	.39
Asthma specialist[b]	87 (57%)	84 (57%)	.89
Asthma action plan	78 (51%)	87 (59%)	.22
Parents' characteristics			
Mean age (years), mean (SD)	37.94 (8.06)	38.03 (9.81)	.92
Female gender, n (%)	145 (94.8%)	138 (93.2%)	.59
African American, n (%)	72 (47%)	78 (53%)	.25
Marital status: with partner, n (%)	73 (48%)	73 (49%)	.82
Highest level of education: high school or less	49 (32%)	47 (32%)	.67
Outcome variables			
ACQ[c], mean (SD) score	2.32 (1.11)	2.49 (1.26)	.21
Symptom-free days, odds ratio (SD)	0.45 (0.39)	0.47 (0.38)	.75
Composite adherence score, mean (SD) %	73.54 (47.81)	69.80 (26.96)	.43
Pharmacy refill possession ratio, mean (SD)	56.86 (27.14)	58.44 (26.68)	.67
Self-report 1, mean (SD) %	88.80 (51.4)	82.92 (27.09)	.30
Self-report 2, mean (SD) %	89.97 (32.11)	87.10 (26.99)	.48
Mediator variables			
Social support[d], mean (SD) score	3.42 (0.73)	3.54 (0.71)	.13
Self-efficacy[d], mean (SD) score	3.58 (0.67)	3.67 (0.62)	.25
Information competence[d], mean (SD) score	3.13 (0.55)	3.25 (0.63)	.59

[a] Comprehensive Health Enhancement Support System plus monthly nurse case management.

[b] In the past year.

[c] Asthma Control Questionnaire, response scale: 1 = excellent asthma control; 7 = very poor asthma control.

[d] Response scale: 1–5.

Table 3. Intervention effects: baseline compared with mean of 3-, 6-, 9-, and 12-month scores.

Outcome	Within-control	Within-CHESS+CM[a]	CHESS+CM – Control
Asthma control			
Symptom-free days odds ratio[b]			
Difference	1.20	1.38	0.18
95% CL[c]	0.98, 1.61	1.12, 1.71	-0.88, 1.60
P value	.29	.01	1.00
Asthma Control Questionnaire[d]			
Difference[d]	–0.11	–0.42	–0.31
95% CL	–0.29, 0.07	–0.60, –0.25	–0.56, –0.06
P value	.22	.001[b]	.01
Adherence			
Composite adherence scores[e]			
Difference	0.58%	2.06%	1.48%
95% CL	–6.24, 7.40	–4.74, 8.86	–8.15, 11.11
P value	.87	.55	.76
Pharmacy refill possession ratio			
Difference	17.7%	13.76%	–3.95%
95% CL	11.78, 23.62	7.83, 19.68	–12.33, 4.43
P value	.001[b]	.001[b]	.35
Self-report controller 1: inhaled corticosteroid[b]			
Difference	–13.42%	–1.78%	11.64%
95% CL	–21.49, –5.35	–18.67, –2.88	–8.65, 13.93
P value	.001[b]	.008	.65
Self-report controller 2: anticholingerics[b]			
Difference	–1.85%	0.95%	2.81%
95% CL	–12.05, 8.34	–8.73, 1.64	–11.26, 16.87
P value	.72	.85	.69

[a] Comprehensive Health Enhancement Support System plus monthly nurse case management.

[b] *P*< .01.

[c] Confidence limits.

[d] *P* ≤ .001.

[e] Sum of self-reported adherence data and pharmacy refill data.

Mechanisms of CHESS+CM Effect on Asthma Control

Figure 4 shows the prespecified secondary mediational analyses. Part a, which shows the premediated effect of CHESS+CM on the ACQ with all covariates entered into the model, was significant at tau = –.22, *P* = .03. Only Medicaid was a significant covariate. Therefore, we entered three covariates into each mediator model: Medicaid, pretest score from the ACQ, and mediator.

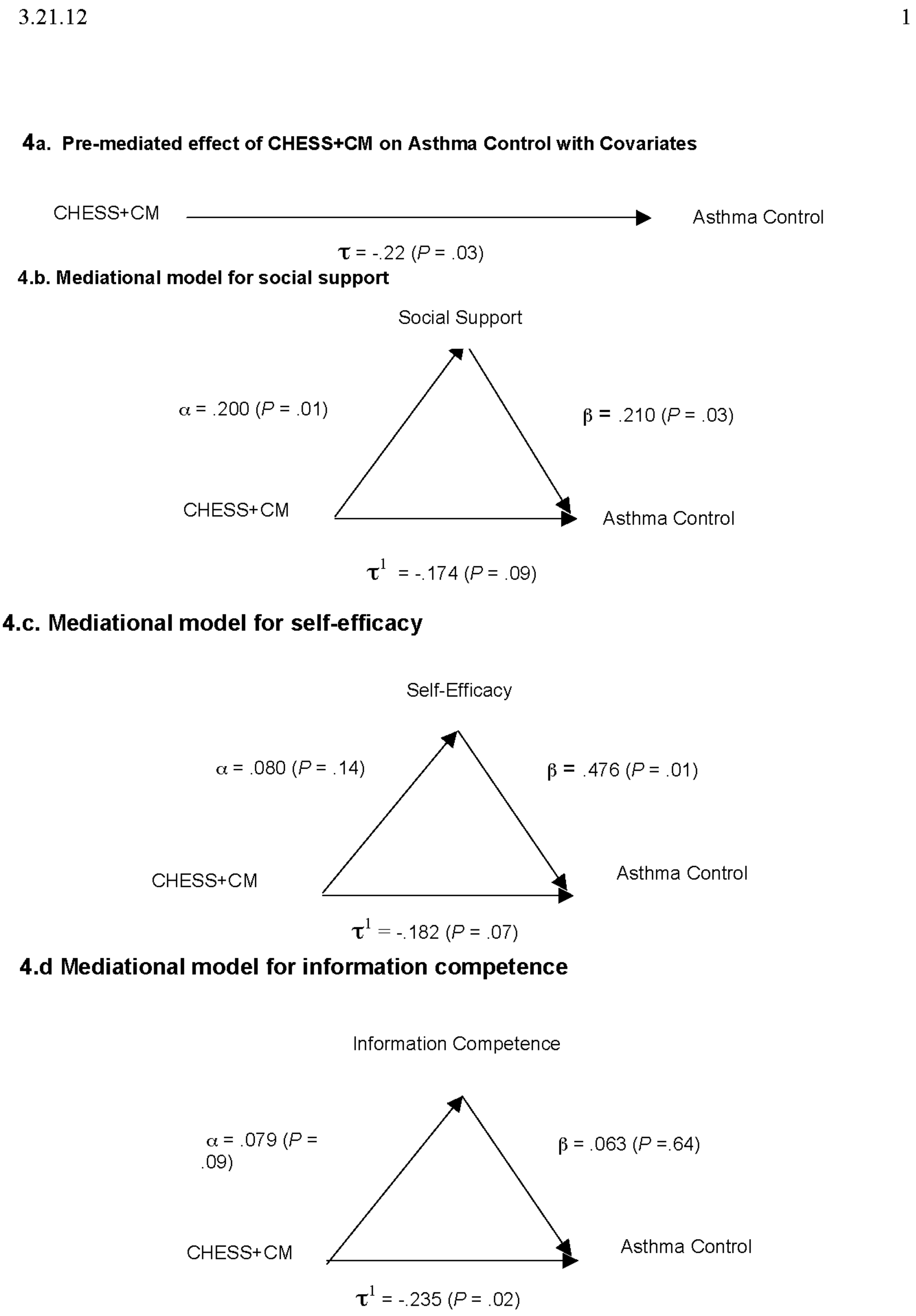
Figure 4 part b shows that CHESS+CM had a significant effect on social support at alpha = .200, *P* = .01, and social support had a significant effect on ACQ at beta = .210, *P* = .03. After mediation, the CHESS+CM effect on ACQ was no longer significant at tau^1 = –.174; *P* = .09, as hypothesized.

Figure 4 part c shows that CHESS+CM had a positive but nonsignificant effect on self-efficacy at alpha = .080, *P* = .14. Self-efficacy had a significant effect on ACQ at beta = .476, *P* = .01. After mediation, CHESS+CM no longer had a significant effect on ACQ at tau^1 = –.182, *P* = .07.

Figure 4 part d shows that CHESS+CM had a marginally significant positive effect on information competence at alpha = .079, *P* = .09; information competence had a positive but nonsignificant effect on ACQ at beta = .063, *P* = .64. After the mediational analysis, CHESS+CM's effect on ACQ remained significant at tau^1 = –.235; P = .02. Information competence, therefore, was not a significant mediator.

In sum, only social support was the only significant mediator for CHESS+CM's effect on asthma control.

Figure 4. Mediational analyses. CHESS+CM = Comprehensive Health Enhancement Support System plus monthly nurse case management.



CHESS Use

Table 4 and Table 5 present the number of logins, pages viewed, and time spent on the site. Table 4 is based on the total number of potential users (146); Table 5 is based on the number of actual users during each month of access. Figure 5 shows the percentage of potential users that actually used CHESS each month. Figure 5 shows a sharp drop-off from the first to the second month and then small declines in most other months. The same can be said for the extent of use. In the last month, usage rates moved up again, possibly because participants knew that the study was coming to an end. This may have led people to employ their last opportunity to use CHESS or it may reflect users wanting to prepare for the exit interview.

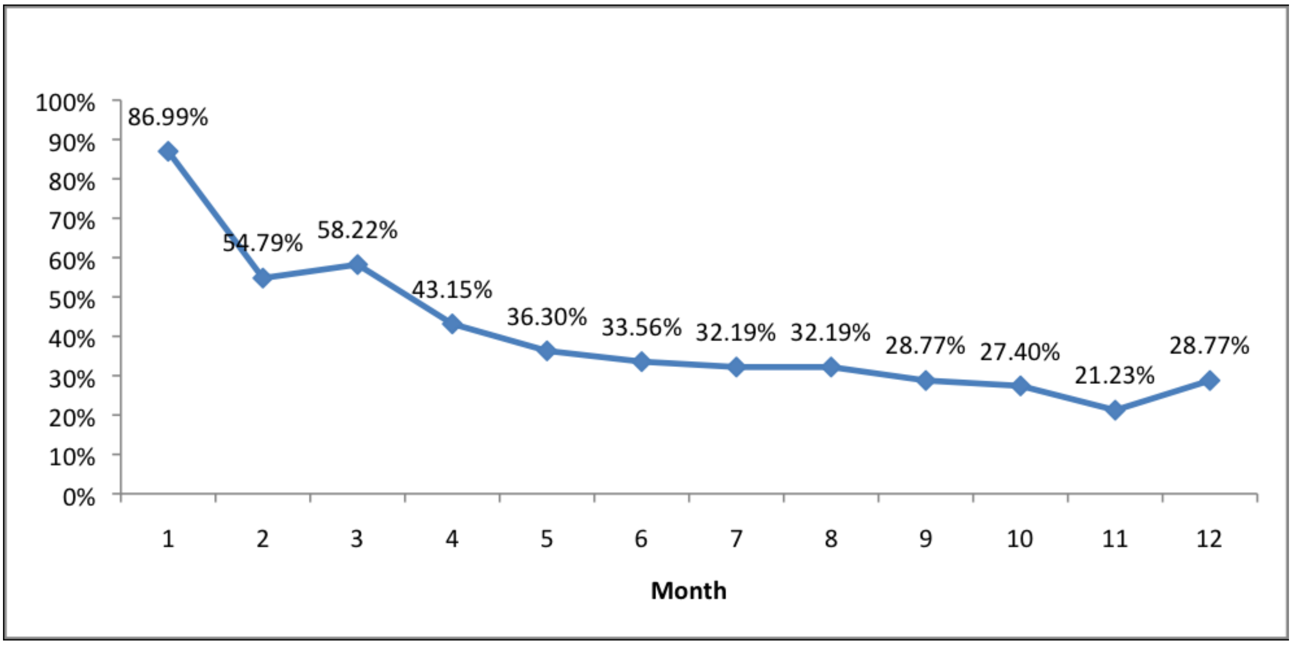
Table 4. Logins, pages viewed, and time spent for the 146 potential users.

Measure	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
Logins												
No.	146	146	146	146	146	146	146	146	146	146	146	146
Mean	10.64	4.15	3.57	2.66	3.12	2.52	2.45	2.60	3.05	2.23	1.40	2.88
SD	16.72	10.24	7.09	6.91	12.59	9.94	8.33	8.73	11.58	7.25	5.17	10.72
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	116	89	43	55	126	107	64	73	80	56	46	81
Pages viewed												
No.	146	146	146	146	146	146	146	146	146	146	146	146
Mean	87.64	29.22	24.12	15.15	13.65	12.08	9.51	8.75	9.56	7.57	6.99	16.47
SD	111.84	61.29	40.54	32.27	38.75	41.14	22.37	20.08	30.35	22.29	23.66	49.10
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	649	551	242	231	307	429	167	105	241	193	155	288
Time (minutes)												
No.	146	146	146	146	146	146	146	146	146	146	146	146
Mean	82.92	25.80	19.52	13.49	12.21	9.21	8.71	9.70	10.25	8.86	8.01	19.50
SD	131.34	68.01	47.36	48.15	51.45	34.63	34.73	33.03	47.87	46.94	40.23	90.78
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	760	700	450	533	524	365	354	265	471	517	403	777

Table 5. Logins, pages viewed, and time spent for actual users who logged into the Comprehensive Health Enhancement Support System (CHESS) each month.

Measure	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
Logins												
No.	127	80	85	63	53	49	47	47	42	40	31	42
Mean	12.24	7.58	6.13	6.17	8.60	7.51	7.60	8.09	10.62	8.13	6.61	10.02
SD	17.38	12.90	8.41	9.47	19.85	16.12	13.37	13.97	19.79	12.10	9.68	18.25
Minimum	1	1	1	1	1	1	1	1	1	1	1	1
Maximum	116	89	43	55	126	107	64	73	80	56	46	81
Pages viewed												
No.	127	80	85	63	53	49	47	47	42	40	31	42
Mean	100.76	53.33	41.44	35.11	37.60	36.00	29.53	27.19	33.24	27.63	32.90	57.24
SD	114.29	74.79	45.95	41.51	57.18	65.08	31.19	27.55	49.52	35.75	42.72	78.32
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	649	551	242	231	307	429	167	105	241	193	155	288
Time (minutes)												
No.	127	80	85	63	53	49	47	47	42	40	31	42
Mean	95.33	47.09	33.53	31.25	33.62	27.45	27.06	30.13	35.62	32.33	37.71	67.79
SD	136.60	86.45	58.28	69.69	81.53	55.78	57.39	53.02	84.73	86.10	81.64	160.61
Minimum	0	0	0	0	0	0	0	0	0	0	0	0
Maximum	760	700	450	533	524	365	354	265	471	517	403	777

Figure 5. Percentage of active Comprehensive Health Enhancement Support System (CHESS) users by month.



Discussion

Summary of Primary Outcomes

We report on a randomized controlled trial that integrated an asthma eHealth program called CHESS with case management from a monthly telephone call from an asthma nurse. The study enrolled 305 parent–child dyads and randomized 301 dyads. Half of our sample were African American and had a low income. Despite the digital divide, the intervention showed significant improvement on pediatric asthma control when measured by the ACQ, but not when measured as symptom-free days from asthma diaries. The intervention did not have a significant effect on adherence to asthma controller medications.

Pharmacy refill rates improved for both groups, yet self-reported adherence declined in both groups. Social support mediated CHESS+CM's effect on the ACQ. Despite positive trends, we found no significant effects for self-efficacy or information competence. Still, the study has important implications because of its focus on (1) low-income, high-risk populations (often African American), (2) an integrated system of nurse case management and eHealth, and (3) mediation analysis to identify the mechanisms of the effect of eHealth systems.

Mediational Effects of Social Support Versus Information

Of the three mediational analyses, social support was the only factor that significantly mediated the CHESS+CM effect on asthma control. CHESS+CM had a marginally significant impact on information competence ($P = .09$), but information competence had no impact on asthma control ($P = .64$). Conversely, CHESS+CM had a nonsignificant but promising trend ($P = .14$) toward improving self-efficacy. The trend was strongly associated with improvements in asthma control ($P = .01$).

These mediational analyses raise an intriguing question about the relative contribution of social support and information in pediatric asthma management and, as a result, the relative effort that should be spent in developing eHealth technologies, at least for asthma. CHESS+CM's modest improvements in information competence but lack of mediational effect for the ACQ may result from an insensitive measure or inadequate presentation of information about or the nature of asthma care, or may indicate that knowledge may not be as important as motivation. On the other hand, an exploratory study of a Web portal for pediatric diabetes patients and their parents reported that users put great value on obtaining information from the site [45]. It might be that, for this largely low-income and minority sample of parents, communicating specific asthma information by phone and case manager email with trusted, caring case managers is a more effective way to encourage asthma management than is reading information in the CHESS program. While these results may not hold for other chronic diseases, we encourage future researchers to consider the relative effects of social support and health information, and possible ways to deliver health information in a supportive manner. Interventions themselves may benefit from being simplified, leading to a more extensive adoption and use of eHealth systems.

Challenges of Measuring Medication Adherence

These results confirm other study results that show asthma diary data are unreliable because of lapses in daily record keeping and overreporting of adherence [46]. The self-reported decline in adherence over time in both groups, however, is puzzling and raises questions about whether participation in the study encouraged more candid responses. This might be interpreted as increasing levels of realistic self-evaluation. Missing data about pharmacy refills were high and may stem from administrative errors and a highly mobile Medicaid population. This research might have benefited from the use of electronic medication measurement devices [47], but these are costly, allow for dumping doses, and may augment adherence and thus reduce generalizability [48].

Limitations

Somewhat surprising was the control group's significant and sustained improvements in refilling asthma controller medications. Control group placebo effects are widely recognized, but primarily for double-blind medication trials rather than health education interventions. Notably, our control group received a welcoming and thorough nurse-led, hour-long intake that was, for many parents and children, their first asthma-focused clinical appointment. This intake included an assessment of the child's asthma and parent's well-being, asthma education as needed, and a warm handoff rather than a referral to follow-up care. Control group participants also received four quarterly mailings with seasonally tailored asthma information, as well as parenting and community resources [49].

In a routine application setting, the case managers would have been employed by the MCO (possibly making our results more optimistic) and the control group would not have received the extensive attention provided in this study (possibly making our results more conservative).

Participants in this eHealth trial were not blinded. This research examined several outcomes of interest (asthma control, symptom-free days, and medication adherence), thus increasing the risk for a type I error. While we did use a Bonferroni correction, we did not use such a correction in the mediation analysis. We do not report analyses comparing users with nonusers of CHESS. Doing so would have increased the length of an already complex paper, shifted attention away from the primary analyses, and introduced substantial biases in research reporting. Finally, eligibility requirements and informed consent limited generalizability of the results. Specifically, we required a level of literacy beyond that possessed by many low-income people.

When this study was initiated, smartphones were just becoming powerful vehicles for change. If we were to do this study over today, we would make several changes. We would use global positioning system tracking to identify when the child entered a prespecified high-risk location, such as a smoker's home. We would install more reminders to both parents and children. Our social media would have included a service in which parents could exchange tips on how best to promote adherence. We would have added a panic button and services to offer help if a child entered an asthma attack. We would have explored the addition of other sensors, such as a peak flowmeter attached to the smartphone. In a second study, we would compare CHESS alone versus CHESS+CM versus control.

Comparison with Prior Work

A 2011 Cochrane review of 21 randomized studies found that asthma telehealth care interventions did not show a clinically significant improvement in patients' quality of life or in the number of emergency department visits, but did show a significant reduction in the number of asthma-related hospitalizations [50]. Most interventions used telephones, and none combined telephone case management with self-guided Web-based education—and none measured the effects on asthma control. Like other pediatric asthma intervention studies with samples, the authors noted that an "active" control group may have reduced the effects on the primary outcomes [51]. Other interventions have improved pediatric asthma outcomes and cognitive learning processes [52,53]. An interactive game that improved children's asthma knowledge, self-management, and clinical outcomes did not significantly improve self-efficacy but found higher scores correlated with better self-management [53]. A school-based intervention improved asthma self-efficacy, knowledge, and asthma management activities as outcomes, but not asthma control [53]. An eHealth program improved asthma knowledge, which correlated with reduced use of rescue medicine and emergency department visits [54]. However, this is the first study to our knowledge that tested the mediational effects of cognitive learning factors on asthma control.

Similarly to the present study, others have found significant effects from brief asthma educational interventions analogous to our active control condition. A single, brief pediatric asthma educational intervention improved asthma outcomes—at least in the short term [55]. An evaluation of self-management support provided by in-home community health workers compared with an active control of three clinic-based pediatric asthma nurse education sessions found modest significant improvements in symptom-free days [56].

This is one of very few eHealth studies that have shown an effect on pediatric asthma caregivers. This is significant because it confirms the family's critical role in disease management [57].

Conclusion and Implications for Further Research

CHESS+CM provided information, social support, and interactive tools to help parents overcome barriers to managing their child's asthma, and secondarily to help the child participate in his or her asthma management [35]. The present study, however, could not identify whether the relative impact of the CHESS eHealth program or monthly phone conversations with the case manager affected these outcomes. Further analyses are needed to identify the specific effects for the separate components of CHESS and case management for different participant profiles and to provide important clues about how asthma education can be tailored better to meet the complex needs of managing pediatric asthma within the family context.

From an eHealth development perspective, more research is needed into the conditions under which it makes sense to invest heavily in various aspects of disease management [58], such as information versus social support. In the present study, the case manager provided asthma information in a supportive and encouraging manner during the monthly phone call—perhaps conflating the relative contribution of information and social support in improving asthma control. In sum, continuous condition-specific and population-specific research and refinement are needed to develop and implement effective eHealth programs.

Finally, information and communication technologies like the one used here might be cost beneficial in disease management [59]. Efficacy studies of information and communication technologies in chronic disease self-management are promising [60,61]. People with addictions tend to view information and communication technologies favorably [62]. They acknowledge more drug use and psychiatric symptoms online than in face-to-face interviews [63]. Computerized screening and brief interventions have been shown to reduce problem drinking [64-66]. A recent review [67] found positive outcomes in 29 of 32 randomized trials of personal computer interventions offering a single service, such as texting and giving reminders, for various chronic diseases. Randomized controlled trials of smartphone systems are just beginning to appear.

Acknowledgments

The authors wish to acknowledge the following: National Institute of Nursing Research (Grant 5R01 NR007889-03) for funding; The University of Wisconsin-Madison GCRC (Grant 3-1145-01) for clinical space to conduct intake interviews; Dean Health Plan, Physicians Plus Insurance Corporation, Group Health Cooperative, Unity Health Insurance, Managed Health Services, and the Wisconsin Medicaid Program for identifying and screening eligible participants. We thank Robert Lemanske, MD, Marcus Cohen, MD, and Donald Bukstein, MD, for reviewing the CHESS asthma program and enlisting the case managers, and Christine Sorkness, PhD, for reviewing the CHESS asthma module content. We thank Pam Richardson, RN, Sally Hollman, BS, RN, Rhonda Christenson, RN, Amy McCullough, DOM PC, and Elisha Terre, BA, for conducting the recruitment, intakes, and monthly telephone case management. Thanks to Eunkyung Kim, MS, for assistance with data analysis, and Andrea Gianopolis, BA, Maureen Fitzgerald, BA, Roberta Johnson, MA, MEd, and Amanda Dinauer, BA, for editorial and bibliographical assistance.

Study protocols were approved by the University of Wisconsin Health Sciences Institutional Review Board on March 9, 2004. The clinical trial registry number for Internet Telehealth for Pediatric Asthma Case Management (CHESS) is NCT00214383.

Funding: National Institute for Nursing Research 5R01 NR007889-03.

Authors' Contributions

David H Gustafson, PhD, as the Principal Investigator, framed the study hypotheses and design and guided the overall development and implementation of the CHESS+CM intervention. For this paper, Dr Gustafson guided the preparation, framed the interpretation of the findings, played a significant role in writing, and approved the final version.

Meg Wise, MLS, PhD, as the project director in years 1 and 2, led the development of the intervention and study tools and as co-investigator in subsequent years coordinated the data analysis team and conducted the descriptive and 1-way analyses of variance. For this paper, Dr Wise played a significant role in writing, reviewing the literature, and coordinating the analysis.

Abhik Bhattacharya, PhD, as the biostatistician, developed the blocking procedures for randomization and managed and analyzed the data. For this paper, Dr Bhattacharya wrote sections on missing data and contributed to sections on the main outcomes. He carefully reviewed all the tables and text for this paper and approved the final version.

Alice D Pulvermacher, MS, as project director in years 3–5, managed the acquisition of data and the implementation of the project, and contributed to refining the implementation design. For this paper, Ms Pulvermacher reviewed drafts and approved the final version.

Kathleen K Shanovich, BSN, MS, as the lead asthma nurse case manager, contributed significantly to the CHESS content development and supervised the intake and exit interview processes and telephone case management during the study implementation. For this paper, Ms Shanovich reviewed drafts and approved the final version.

Brenda R Phillips, MS, conducted the analyses for symptom-free days from the asthma diary data. For this paper, Ms Phillips contributed to and reviewed the statistical methods and results sections, and reviewed and approved the final version.

Erik B Lehman, MS, conducted the statistical analyses for the Asthma Control Questionnaire and medication adherence. For this paper, Mr Lehman contributed to and reviewed the statistical methods and results sections, and reviewed and approved the final version.

Vernon M Chinchilli, PhD, developed, designed, and supervised the analyses for all primary outcome variables. For this paper, Dr Chinchilli contributed to and reviewed the statistical methods and results sections and reviewed and approved the final version.

Robert*P* Hawkins, PhD, as a co-investigator, contributed to the design of the research protocols and CHESS program content and to data analysis and interpretation. For this paper, Dr Hawkins provided substantive comments and approved the final version.

Jee-Seon Kim, PhD, designed and conducted the mediational analyses. For this paper, Dr Kim wrote the methods and results sections and developed the figures related to the mediational analyses.

Conflicts of Interest

The authors have no financial interests in the eHealth system evaluated here, although Gustafson, Wise, and Hawkins were lead members of the development team.

Multimedia Appendix 1

CONSORT eHealth checklist V1.6 [68].

[PDF File (Adobe PDF File), 467KB - jmir_v14i4e101_app1.pdf]

References

1. McKay JR, Van Horn DH, Oslin DW, Lynch KG, Ivey M, Ward K, et al. A randomized trial of extended telephone-based continuing care for alcohol dependence: within-treatment substance use outcomes. J Consult Clin Psychol 2010 Dec;78(6):912-923. [doi: 10.1037/a0020700] [Medline: 20873894]
2. Maisto SA, Zywiak WH, Connors GJ. Course of functioning 1 year following admission for treatment of alcohol use disorders. Addict Behav 2006 Jan;31(1):69-79. [doi: 10.1016/j.addbeh.2005.04.008] [Medline: 15919159]
3. Godley MD, Godley SH, Dennis ML, Funk R, Passetti LL. Preliminary outcomes from the assertive continuing care experiment for adolescents discharged from residential treatment. J Subst Abuse Treat 2002 Jul;23(1):21-32. [Medline: 12127465]
4. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, MOBITEL Investigators. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009 Aug;11(3):e34 [FREE Full text] [doi: 10.2196/jmir.1252] [Medline: 19687005]
5. Noh JH, Cho YJ, Nam HW, Kim JH, Kim DJ, Yoo HS, et al. Web-based comprehensive information system for self-management of diabetes mellitus. Diabetes Technol Ther 2010 May;12(5):333-337. [doi: 10.1089/dia.2009.0122] [Medline: 20388042]
6. Helzer JE, Badger GJ, Rose GL, Mongeon JA, Searles JS. Decline in alcohol consumption during two years of daily reporting. J Stud Alcohol 2002 Sep;63(5):551-558. [Medline: 12380851]
7. Davis JR, Glaros AG. Relapse prevention and smoking cessation. Addict Behav 1986;11(2):105-114. [Medline: 3739795]
8. Larimer ME, Palmer RS, Marlatt GA. Relapse prevention. An overview of Marlatt's cognitive-behavioral model. Alcohol Res Health 1999;23(2):151-160. [Medline: 10890810]
9. Seale JP, Shellenberger S, Velasquez MM, Boltri JM, Okosun I, Guyinn M, et al. Impact of vital signs screening & clinician prompting on alcohol and tobacco screening and intervention rates: a pre-post intervention comparison. BMC Fam Pract 2010 Mar;11:18 [FREE Full text] [doi: 10.1186/1471-2296-11-18] [Medline: 20205740]
10. Witkiewitz KA, Marlatt GA. High-risk situations: relapse as a dynamic process. In: Witkiewitz K, Marlatt G, editors. Therapist’s Guide to Evidence-Based Relapse Prevention. Amsterdam: Elsevier Academic Press; 2007:19-33.

11. van Osch L, Lechner L, Reubsaet A, Wigger S, de Vries H. Relapse prevention in a national smoking cessation contest: effects of coping planning. Br J Health Psychol 2008 Sep;13(Pt 3):525-535. [doi: 10.1348/135910707X224504] [Medline: 17631695]
12. Washton AM. Cocaine Addiction: Treatment, Recovery, and Relapse Prevention. New York, NY: WW Norton & Company; 1991.
13. Sullivan, WP, Wolk JL, Hartmann DJ. Case management in alcohol and drug treatment: improving client outcomes. Fam Soc 1992;73(4):195-204. [doi: 10.1606/1044-3894.1677]
14. Dennis M, Scott CK, Funk R. An experimental evaluation of recovery management check-ups (RMC) for people with chronic substance use disorders. Eval Program Plann 2003 Aug;26(3):339-352.
15. McLellan AT, Hagan TA, Levine M, Meyers K, Gould F, Bencivengo M, et al. Does clinical case management improve outpatient addiction treatment. Drug Alcohol Depend 1999 Jun 1;55(1-2):91-103. [Medline: 10402154]
16. Lorig KR, Ritter PL, Laurent DD, Fries JF. Long-term randomized controlled trials of tailored-print and small-group arthritis self-management interventions. Med Care 2004 Apr;42(4):346-354. [Medline: 15076811]
17. Rapp RC, Siegal HA, Fisher JH. A strengths-based model of case management/advocacy: adapting a mental health model to practice work with persons who have substance abuse problems. NIDA Res Monogr 1992;127:79-91. [Medline: 1436007]
18. Shaw BR, Jeong Yeob H, Hawkins RP, McTavish FM, Gustafson DH. Communicating about self and others within an online support group for women with breast cancer and subsequent outcomes. J Health Psychol 2008 Oct;13(7):930-939. [doi: 10.1177/1359105308095067] [Medline: 18809644]
19. Palermo TM, Wilson AC, Peters M, Lewandowski A, Somhegyi H. Randomized controlled trial of an Internet-delivered family cognitive-behavioral therapy intervention for children and adolescents with chronic pain. Pain 2009 Nov;146(1-2):205-213. [doi: 10.1016/j.pain.2009.07.034] [Medline: 19695776]
20. DuBenske LL, Gustafson DH, Shaw BR, Cleary JF. Web-based cancer communication and decision making systems: connecting patients, caregivers, and clinicians for improved health outcomes. Med Decis Making 2010 Dec;30(6):732-744. [doi: 10.1177/0272989X10386382] [Medline: 21041539]
21. Wise M, Gustafson DH, Sorkness CA, Molfenter T, Staresinic A, Meis T, et al. Internet telehealth for pediatric asthma case management: integrating computerized and case manager features for tailoring a Web-based asthma education program. Health Promot Pract 2007 Jul;8(3):282-291. [doi: 10.1177/1524839906289983] [Medline: 16928987]
22. Bloom B, Cohen RA, Freeman G. Summary health statistics for U.S. children: National Health Interview Survey, 2009. Vital Health Stat 10 2010 Dec(247):1-82. [Medline: 21563639]
23. Wolf FM, Guevara JP, Grum CM, Clark NM, Cates CJ. Educational interventions for asthma in children. Cochrane Database Syst Rev 2002(4):CD000326.
24. Laster N, Holsey CN, Shendell DG, Mccarty FA, Celano M. Barriers to asthma management among urban families: caregiver and child perspectives. J Asthma 2009 Sep;46(7):731-739. [doi: 10.1080/02770900903082571] [Medline: 19728215]
25. Fox P, Porter PG, Lob SH, Boer JH, Rocha DA, Adelson JW. Improving asthma-related health outcomes among low-income, multiethnic, school-aged children: results of a demonstration project that combined continuous quality improvement and community health worker strategies. Pediatrics 2007 Oct;120(4):e902-e911 [FREE Full text] [doi: 10.1542/peds.2006-1805] [Medline: 17908746]
26. Karnick P, Margellos-Anast H, Seals G, Whitman S, Aljadeff G, Johnson D. The pediatric asthma intervention: a comprehensive cost-effective approach to asthma management in a disadvantaged inner-city community. J Asthma 2007 Feb;44(1):39-44. [doi: 10.1080/02770900601125391] [Medline: 17365203]
27. Coffman JM, Cabana MD, Halpin HA, Yelin EH. Effects of asthma education on children's use of acute care services: a meta-analysis. Pediatrics 2008 Mar;121(3):575-586 [FREE Full text] [doi: 10.1542/peds.2007-0113] [Medline: 18310208]
28. Lieberman DA. Management of chronic pediatric diseases with interactive health games: theory and research findings. J Ambul Care Manage 2001 Jan;24(1):26-38. [Medline: 11189794]
29. Shegog R, Bartholomew LK, Gold RS, Pierrel E, Parcel GS, Sockrider MM, et al. Asthma management simulation for children: translating theory, methods, and strategies to effect behavior change. Simul Healthc 2006;1(3):151-159. [doi: 10.1097/01.SIH.0000244456.22457.e8] [Medline: 19088584]
30. Burgess SW, Sly PD, Morawska A, Devadason SG. Assessing adherence and factors associated with adherence in young children with asthma. Respirology 2008 Jun;13(4):559-563. [doi: 10.1111/j.1440-1843.2008.01292.x] [Medline: 18422868]
31. Finkelstein JA, Lozano P, Fuhlbrigge AL, Carey VJ, Inui TS, Soumerai SB, Pediatric Asthma Care Patient Outcomes Research Team. Practice-level effects of interventions to improve asthma care in primary care settings: the Pediatric Asthma Care Patient Outcomes Research Team. Health Serv Res 2005 Dec;40(6 Pt 1):1737-1757. [doi: 10.1111/j.1475-6773.2005.00451.x] [Medline: 16336546]
32. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
33. Bandura A. Self-efficacy mechanism in human agency. Am Psychol 1982 Feb;37(2):122-147. [doi: 10.1037/0003-066X.37.2.122]
34. Smith S, Mitchell C, Bowler S. Patient-centered education: applying learner-centered concepts to asthma education. J Asthma 2007 Dec;44(10):799-804. [doi: 10.1080/02770900701645256] [Medline: 18097853]

35. Gustafson DH, Hawkins R, Pingree S, McTavish F, Arora NK, Mendenhall J, et al. Effect of computer support on younger women with breast cancer. J Gen Intern Med 2001 Jul;16(7):435-445. [Medline: 11520380]
36. Gustafson DH, Hawkins R, Boberg E, Pingree S, Serlin RE, Graziano F, et al. Impact of a patient-centered, computer-based health information/support system. Am J Prev Med 1999 Jan;16(1):1-9. [Medline: 9894548]
37. Wise M, Han JY, Shaw B, McTavish F, Gustafson DH. Effects of using online narrative and didactic information on healthcare participation for breast cancer patients. Patient Educ Couns 2008 Mar;70(3):348-356. [doi: 10.1016/j.pec.2007.11.009] [Medline: 18201859]
38. National Heart, Lung, and Blood Institute, National Asthma Education and Prevention Program. US Department of Health and Human Services, National Institutes of Health. 2007. Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma: Full Report URL: http://www.nhlbi.nih.gov/guidelines/asthma/asthgdln.pdf [accessed 2012-03-23] [WebCite Cache ID 66Nfx9tVg]
39. Wisconsin Asthma Program. Wisconsin Department of Health Services, Division of Public Health, Bureau of Environmental and Occupational Health. 2010. Burden of Asthma in Wisconsin URL: http://www.dhs.wisconsin.gov/eh/asthma/pdf/BurdenofAsthma2010Web.pdf [accessed 2012-03-22] [WebCite Cache ID 66MYf4wS9]
40. Wise M, Pulvermacher A, Shanovich KK, Gustafson DH, Sorkness C, Bhattacharya A. Using action research to implement an integrated pediatric asthma case management and eHealth intervention for low-income families. Health Promot Pract 2010 Nov;11(6):798-806. [doi: 10.1177/1524839909334621] [Medline: 19515862]
41. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999 Oct;14(4):902-907 [FREE Full text] [Medline: 10573240]
42. Santanello NC. Pediatric asthma assessment: validation of 2 symptom diaries. J Allergy Clin Immunol 2001 May;107(5 Suppl):S465-S472. [Medline: 11344376]
43. Moeinedin F, Moineddin R, Jadad AR, Hamid JS, To T, Beyene J. Application of biomedical informatics to chronic pediatric diseases: a systematic review. BMC Med Inform Decis Mak 2009;9:22 [FREE Full text] [doi: 10.1186/1472-6947-9-22] [Medline: 19416540]
44. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002 Mar;7(1):83-104. [Medline: 11928892]
45. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 Diabetes Portal--the management tool, the generator, and the gatekeeper: qualitative study. J Med Internet Res 2010;12(2):e17 [FREE Full text] [doi: 10.2196/jmir.1267] [Medline: 20511179]
46. Milgrom H, Bender B, Ackerson L, Bowry P, Smith B, Rand C. Noncompliance and treatment failure in children with asthma. J Allergy Clin Immunol 1996 Dec;98(6 Pt 1):1051-1057. [Medline: 8977504]
47. Bender B, Milgrom H, Apter A. Adherence intervention research: what have we learned and what do we do next? J Allergy Clin Immunol 2003 Sep;112(3):489-494. [Medline: 13679805]
48. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]
49. Tinkelman D, Schwartz A. School-based asthma disease management. J Asthma 2004 Jun;41(4):455-462. [Medline: 15281331]
50. McLean S, Chandler D, Nurmatov U, Liu J, Pagliari C, Car J, et al. Telehealthcare for asthma: a Cochrane review. CMAJ 2011 Aug 9;183(11):E733-E742. [doi: 10.1503/cmaj.101146] [Medline: 21746825]
51. Schaffer SD, Tian L. Promoting adherence: effects of theory-based asthma education. Clin Nurs Res 2004 Feb;13(1):69-89. [doi: 10.1177/1054773803259300] [Medline: 14768768]
52. Wolf FM, Guevara JP, Grum CM, Clark NM, Cates CJ. Educational interventions for asthma in children. Cochrane Database Syst Rev 2003(1):CD000326. [doi: 10.1002/14651858.CD000326] [Medline: 12535395]
53. Bartholomew LK, Sockrider M, Abramson SL, Swank PR, Czyzewski DI, Tortolero SR, et al. Partners in school asthma management: evaluation of a self-management program for children with asthma. J Sch Health 2006 Aug;76(6):283-290. [doi: 10.1111/j.1746-1561.2006.00113.x] [Medline: 16918856]
54. Krishna S, Francisco BD, Balas EA, König P, Graff GR, Madsen RW. Internet-enabled interactive multimedia asthma education program: a randomized trial. Pediatrics 2003 Mar;111(3):503-510. [Medline: 12612228]
55. Lukacs SL, France EK, Barón AE, Crane LA. Effectiveness of an asthma management program for pediatric members of a large health maintenance organization. Arch Pediatr Adolesc Med 2002 Sep;156(9):872-876. [Medline: 12197793]
56. Krieger J, Takaro TK, Song L, Beaudet N, Edwards K. A randomized controlled trial of asthma self-management support comparing clinic-based nurses and in-home community health workers: the Seattle-King County Healthy Homes II Project. Arch Pediatr Adolesc Med 2009 Feb;163(2):141-149. [doi: 10.1001/archpediatrics.2008.532] [Medline: 19188646]
57. Brown N, Gallagher R, Fowler C, Wales S. The role of parents in managing asthma in middle childhood: an important consideration in chronic care. Collegian 2010;17(2):71-76. [Medline: 20738059]
58. Gentles SJ, Lokker C, McKibbon KA. Health information technology to facilitate communication involving health care providers, caregivers, and pediatric patients: a scoping review. J Med Internet Res 2010;12(2):e22 [FREE Full text] [doi: 10.2196/jmir.1390] [Medline: 20562092]

59. Popovici I, French MT, McKay JR. Economic evaluation of continuing care interventions in the treatment of substance abuse: recommendations for future research. Eval Rev 2008 Dec;32(6):547-568. [doi: 10.1177/0193841X08316311] [Medline: 18334678]
60. 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]
61. Solomon M, Wagner SL, Goes J. Effects of a Web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012;14(1):e32 [FREE Full text] [doi: 10.2196/jmir.1924] [Medline: 22353433]
62. Cunningham J, Wild TC, Walsh GW. Interest in self-help materials in a general population sample of drinkers. Drugs Educ Prev Policy 1999 Jul;6(2):209-213. [doi: 10.1080/09687639997179]
63. Rosen CS, Henson BR, Finney JW, Moos RH. Consistency of self-administered and interview-based Addiction Severity Index composite scores. Addiction 2000 Mar;95(3):419-425. [Medline: 10795362]
64. 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]
65. Murray E, McCambridge J, Khadjesari Z, White IR, Thompson SG, Godfrey C, et al. The DYD-RCT protocol: an on-line randomised controlled trial of an interactive computer-based intervention compared with a standard information website to reduce alcohol consumption among hazardous drinkers. BMC Public Health 2007;7:306 [FREE Full text] [doi: 10.1186/1471-2458-7-306] [Medline: 17963483]
66. Saitz R, Palfai TP, Freedner N, Winter MR, Macdonald A, Lu J, et al. Screening and brief intervention online for college students: the ihealth study. Alcohol Alcohol 2007 Feb;42(1):28-36 [FREE Full text] [doi: 10.1093/alcalc/agl092] [Medline: 17130139]
67. Gustafson DH, Boyle MG, Shaw BR, Isham A, McTavish F, Richards S, et al. An e-health solution for people with alcohol problems. Alcohol Res Health 2011;33(4):327-337 [FREE Full text]
68. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ACQ: Asthma Control Questionnaire
CHESS: Comprehensive Health Enhancement Support System
CHESS+CM: Comprehensive Health Enhancement Support System plus monthly nurse case management
CL: confidence limits
MCO: managed care organization

Edited by G Eysenbach; submitted 12.10.11; peer-reviewed by W Slack, G Kreps, D Dalan, S Patolia; comments to author 11.01.12; revised version received 31.01.12; accepted 25.04.12; published 26.07.12.

Please cite as:
Gustafson D, Wise M, Bhattacharya A, Pulvermacher A, Shanovich K, Phillips B, Lehman E, Chinchilli V, Hawkins R, Kim JS
The Effects of Combining Web-Based eHealth With Telephone Nurse Case Management for Pediatric Asthma Control: A Randomized Controlled Trial
J Med Internet Res 2012;14(4):e101
URL: http://www.jmir.org/2012/4/e101/
doi:10.2196/jmir.1964
PMID:22835804

Original Paper

Effects of Functional Interactivity on Patients' Knowledge, Empowerment, and Health Outcomes: An Experimental Model-Driven Evaluation of a Web-Based Intervention

Luca Camerini[1], PhD; Peter Johannes Schulz[1], PhD

Institute of Communication and Health, Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland

Corresponding Author:
Luca Camerini, PhD
Institute of Communication and Health
Faculty of Communication Sciences
Università della Svizzera italiana
via G. Buffi 13
Lugano, 6900
Switzerland
Phone: 41 0586664821
Fax: 41 0586664647
Email: lukame@gmail.com

Abstract

Background: The effectiveness of eHealth interventions in terms of reach and outcomes is now well documented. However, there is a need to understand not only whether eHealth interventions work, but also what kind of functions and mechanisms enhance their effectiveness. The present investigation contributes to tackling these challenges by investigating the role played by functional interactivity on patients' knowledge, empowerment, and health outcomes.

Objectives: To test whether health knowledge and empowerment mediate a possible relationship between the availability of interactive features on an eHealth application and individuals' health outcomes. We present an empirical, model-driven evaluation of the effects of functional interactivity implemented in an eHealth application, based on a brief theoretical review of the constructs of interactivity, health knowledge, empowerment, and health outcomes. We merged these constructs into a theoretical model of interactivity effects that we tested on an eHealth application for patients with fibromyalgia syndrome (FMS).

Methods: This study used a pretest–posttest experimental design. We recruited 165 patients and randomly assigned them to three study groups, corresponding to different levels of functional interactivity. Eligibility to participate in the study required that patients (1) be fluent in Italian, (2) have access to the Internet, (3) report confidence in how to use a computer, and (4) have received a diagnosis of FMS from a doctor. We used structural equation modeling techniques to analyze changes between the pretest and the posttest results.

Results: The main finding was that functional interactivity had no impact on empowerment dimensions, nor direct observable effects on knowledge. However, knowledge positively affected health outcomes ($b = –.12$, $P = .02$), as did the empowerment dimensions of meaning ($b = –.49$, $P < .001$) and impact ($b = –.25$, $P < .001$).

Conclusion: The theoretical model was partially confirmed, but only as far as the effects of knowledge and empowerment were concerned. The differential effect of interactive functions was by far weaker than expected. The strong impact of knowledge and empowerment on health outcomes suggests that these constructs should be targeted and enhanced by eHealth applications.

(J Med Internet Res 2012;14(4):e105) doi:10.2196/jmir.1953

KEYWORDS

Interactivity; health literacy; patient empowerment; fibromyalgia; Internet intervention

Introduction

The effectiveness of eHealth interventions in terms of reach and outcomes is now well documented. Several reviews and meta-analyses showed the benefits of designing, implementing, delivering, and maintaining health programs on the Internet [1-7]. Still, Bennett and Glasgow [8] pointed out that one of the major challenges for eHealth research is to evaluate the differential effects of the enabling functions implemented in the applications, rather than consider them as black boxes [8]. In other words, there is a need not only to know *whether* eHealth interventions work, but also to understand *how* they achieve their effectiveness.

The question of how eHealth interventions work has at least two aspects. The first is to identify the effective elements in these interventions; this comes down to achieving a more precise understanding of the independent variable. As it is among the prime features that distinguish Internet interventions from other mediated communication interventions, an obvious subject to turn to is the role played by interactivity for the outcome of eHealth interventions. Indeed, in the health care domain, interactive applications are favored, and optimistic claims on their effectiveness abound [9-16]. However, the body of evidence on the effects of interactivity on behavioral and health outcomes is not very broad. This is perhaps linked to the fact that the definition and operationalization of interactivity have always been difficult [15]. According to Rafaeli and Ariel [15] and Rafaeli [17], a rough distinction can be made between the conceptualizations focused on functions of features and those focused on users.

Interactivity as an attribute of technology (eg, [18]) can be defined as “the extent to which users can participate in modifying the form and content of a mediated environment in real time” [19,20]. In this sense, traditional media are low in terms of interactivity, while new technologies, such as the Internet, share a high level of interactivity. This perspective is known as functional interactivity and offers a theoretically robust way to operationalize an eHealth intervention in terms of its enabling functions.

The second aspect is to better understand the process that creates the effect. This means to find intervening variables that mediate or moderate the effect of eHealth interventions. As eHealth interventions mostly aim at improving self-management and self-help capabilities, approaches from the larger field of patient autonomy come into focus: health literacy (or rather one of its elements, patient knowledge) and patient or health empowerment. Both are assumed or have been shown to affect health outcomes.

In its original definition, health literacy was conceived as “the ability to read and comprehend prescription bottles, appointment slips, and other essential health-related material” [21]. Reading and numeracy are considered essential skills an individual should possess in order to navigate the health care system [22,23]. This functional perspective, however, has been criticized for being too narrow and missing much of the richness of the deeper meaning and purpose of literacy for people (eg, [24]). As a consequence, broader definitions have been proposed [24,25]. For this study, we adopted the critical perspective on health literacy and health knowledge proposed by Schulz and Nakamoto [26]. These authors conceived health literacy and knowledge as a function of basic reading and writing skills, declarative knowledge, procedural knowledge, and judgmental skills.

Patient or health empowerment, in turn, was theorized as the process by which people, organizations, and communities gain mastery over issues of concern to them [27-39]. Thomas and Velthouse [40], building on the psychological and organizational literature [41-43], proposed a model of empowerment as a multidimensional construct, composed of four cognitive variables (or task assessments): impact (or the degree to which behavior is seen as “making a difference”), competence (or the degree to which a person can perform task activities skillfully), meaningfulness (or the individual’s intrinsic caring about a given task), and choice (or weather a person’s behavior is perceived as self-determined). In the health care domain, powerlessness has been associated with ill health [44], whereas empowerment is considered a determinant of improved health status [45-47] and eHealth interventions, especially online support groups, are considered an effective empowering tool [48-50]. A recent review and meta-analysis of the effectiveness of Web-based interventions on patient empowerment showed that eHealth significantly contributes to increased empowerment, even when compared with face-to-face consultations or usual care [51].

Links between eHealth interactivity on the one hand, and health literacy and empowerment on the other, are plausible to assume, for it takes knowledge and literacy to benefit from eHealth applications, and bringing them up in medical consultation gives a more active role to patients, who thus claim autonomy. Knowledge and empowerment can also be assumed to be affected by eHealth interventions and their qualities. In a sense, such links are obvious to assume, but they are largely unexplored by research and therefore not established at all [52]. Our research aimed at contributing to this area, resting on the assumption that eHealth fosters patient autonomy. In particular, we hypothesized that the availability of interactive functions in eHealth interventions would positively affect users’ knowledge and their empowerment, and further that knowledge and empowerment would positively affect health outcomes.

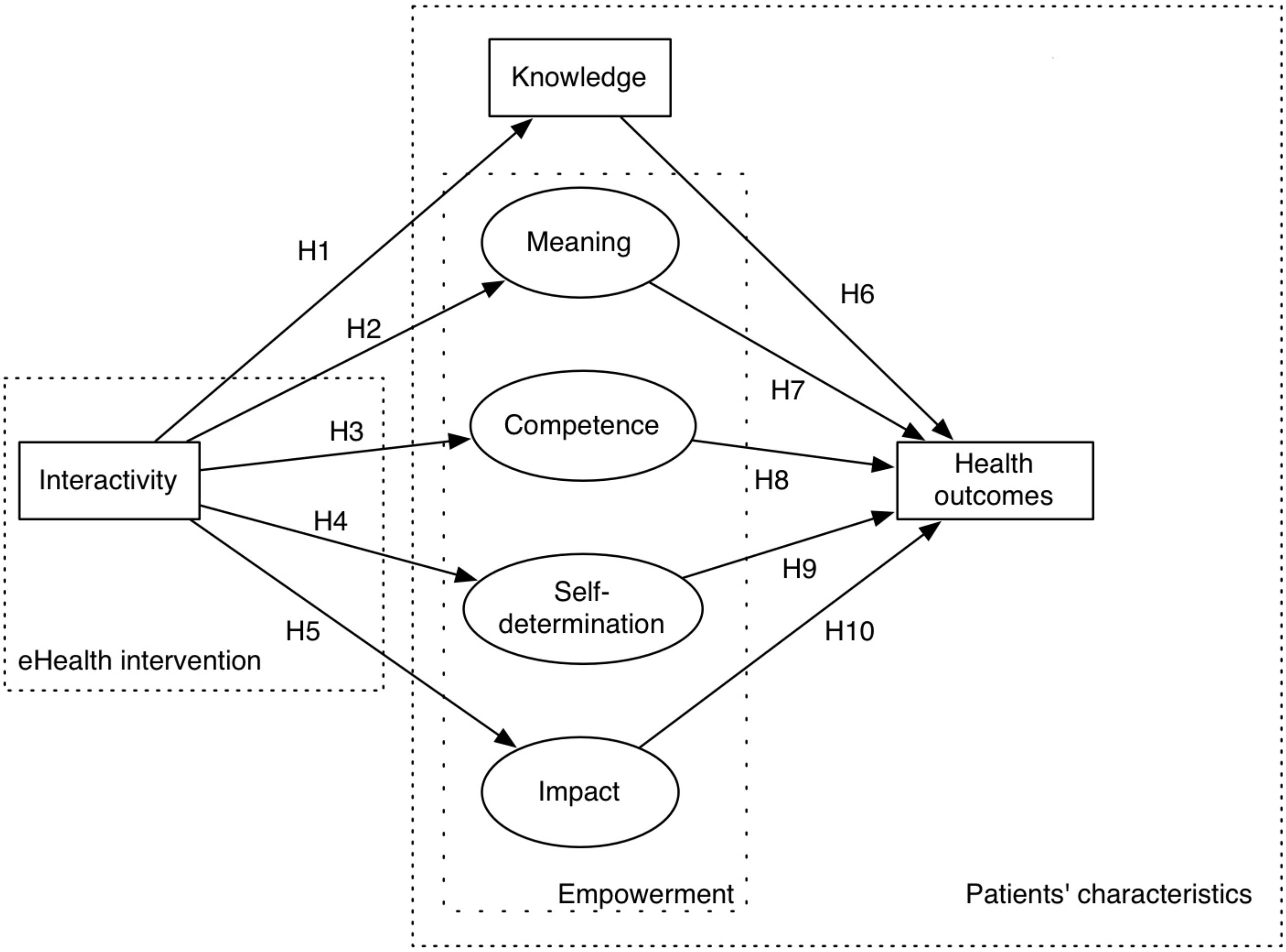
The main aim of this study was to test whether health knowledge and empowerment mediate a possible relationship between the availability of interactive features on an eHealth application and individuals’ health outcomes. We expected that interactivity increases knowledge and empowerment, which in turn improve patient assessment of fibromyalgia impact, our health outcome measure. Similar assumptions for other interventions and conditions were already tested—and partially confirmed—in other studies relating interactive eHealth applications with knowledge or empowerment, or both [10,53-56]. Therefore, this study tested a specific model that conceives aspects of patient autonomy as mediating a possible influence of interactivity on assessment of impact. To keep a confirmatory rather than an exploratory approach to these relationships, we did not directly test alternative models (eg, that knowledge and empowerment could also moderate the effect of interactivity). Figure 1 illustrates our model of interactivity effects on patients’ assessment of fibromyalgia impact.

Each arrow in Figure 1 represents a hypothesis, all of which are a specification of the general mediator hypothesis mentioned.

Specifically, patients who use an application that offers interactive functions become more knowledgeable (H1) and achieve a higher score on the empowerment dimensions of meaning (H2), competence (H3), self-determination (H4), and impact (H5) than patients not offered these functions. In turn, the higher the level of knowledge (H6), meaning (H7), competence (H8), self-determination (H9), and impact (H10), the better (ie, lower) the patients assess their fibromyalgia impact.

The mediator analysis rests on the existence of an influence of interactivity condition on outcome, in the sense that patients who are offered interactive function will assess fibromyalgia impact better. We also tested this.

Figure 1. A model of functional interactivity effects on knowledge, empowerment, and health outcomes. Arrows indicate directional effects. H1-H10 = main hypotheses.



Methods

Reference Population

The reference population in this study consisted of patients with fibromyalgia syndrome (FMS). According to the American College of Rheumatology [57], FMS is a condition characterized by chronic widespread pain and tenderness in 11 or more of the 18 specific tender point sites. Although the medical evidence is still lacking for precise diagnostic criteria for FMS, there are three major symptoms that are usually associated with the disease: pain, sleep disorders, and fatigue [58-60]. Alongside these somatic factors are many other psychological dimensions observed in FMS patients, such as anxiety, stress, and depression [61-63]. People with FMS usually have other concurrent conditions, including diabetes, high blood pressure, and back pain. FMS outcomes are generally measured with the Fibromyalgia Impact Questionnaire (FIQ) [64-66] or the Fibromyalgia Assessment Status (FAS) [67]. The FIQ includes items covering all the disease-specific domains, accounting for functional disability, job ability, pain intensity, sleep function, stiffness, anxiety, depression, and the overall sense of wellbeing. The FAS focuses on the three main characteristics of FMS: fatigue, sleep disturbances, and pain. Both of these measures have good psychometric properties, in terms of reliability and validity. For this study, we used the FIQ, as the FAS was still under development and testing. In other words, our health outcome variable was patients' assessment of fibromyalgia impact.

The eHealth Intervention

To test our hypotheses, we developed a Web-based eHealth intervention for patients with FMS. This application, called ONESELF, was developed in collaboration with health professionals (rheumatologists, physiotherapists, and general practitioners), and it is fully compliant with the Health On the Net Foundation guidelines (HONcode). The HONcode prescribes guidelines on the quality of the contents and the overall usability of an application. Indeed, these intrinsic factors can play a decisive role in the ultimate effects of an online intervention, and compliance with these guidelines helps to temper—and to some extent rule out—potential biases caused by usability issues.

The application enabled asynchronous and synchronous interactions with health professionals and laypeople. The system also included static informative sections in a virtual library that provided users with relevant information on the disease. First aid and frequently asked questions sections provided brief and practical information on syndrome management. A virtual gymnasium provided patients with tailored multimedia contents on several physical exercises that constitute the wider part of the nonpharmacological treatment of FMS. Later, a section with testimonies, where patients could post their stories and read stories of other people with the same health condition, enhanced the dimension of social support. We also designed and implemented synchronous interaction via a chat room and asynchronous interaction in an online forum. Patients used these tools to communicate with the physicians and among themselves. A more detailed description of the design of the application is presented elsewhere [68,69], together with qualitative insights on the user experience with the system, which was generally considered useful, usable, and comprehensible. Since the first release of ONESELF in June 2008, more than 600 FMS patients, mostly from Switzerland and Italy (the site language being Italian), have registered.

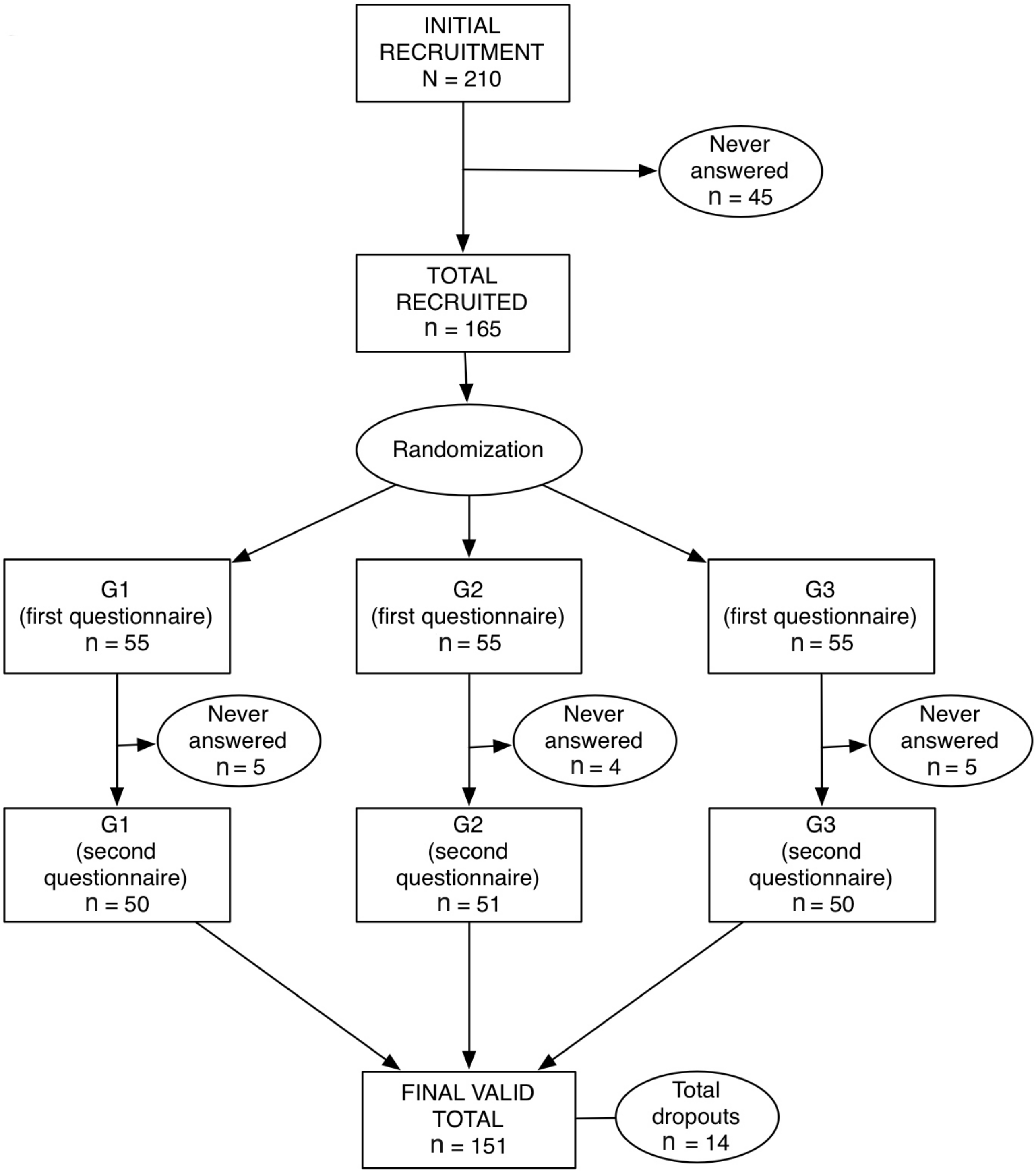
Study Design and Procedure

This study used a pretest–posttest experimental design. Patients were contacted by two means: a list of patients who were members of the *Associazione Fibromialgici Svizzeri Sezione Ticino* (Ticino Fibromyalgia Patients Association) and patients visiting health professionals (rheumatologists and physiotherapists). Health professionals were involved in the recruitment to assure that patients received a diagnosis of FMS from a doctor. To be eligible for the study, patients had to meet a set of inclusion criteria: (1) be fluent in Italian, (2) have access to the Internet, (3) be confident in using a computer, and (4) have received a diagnosis of FMS from a doctor. Every patient who matched these criteria was given a letter briefly describing the aims of the study, together with a contact form. If interested in the study, patients had to send the form back to the research team, filled in with contact details. The research team contacted the patients by phone and email. Patients were introduced to the study and asked to register on the ONESELF website. After registration, they had to accept an informed consent statement and, finally, complete the first questionnaire. After completing the questionnaire, patients could access the website and start the navigation. After 5 months of access to the site, a second questionnaire was presented and completed. A maximum of three reminders were used to maximize response rate. By the end of the recruitment process, a total of 165 patients had agreed to participate in the study (Figure 2).

To investigate the effect of functional interactivity, we created three different versions of the ONESELF application, each implementing different enabling functions. Patients were randomly assigned to one of the three versions and blinded to the others, using a computer utility that assigned them to a randomly selected experimental condition until the conditions were equally filled. Patients in group 1 (n = 55) were given a static version of ONESELF, including only the library, the virtual gymnasium, the testimonials, and the generic sections such as the first aid, the frequently asked questions, and other common contents (eg, contacts, legal notices, and ownership disclosure). No interactive enabling tools such as the Web forum or the chat room were present in this version. This group was considered as the reference or control group. Patients in group 2 (n = 55) were given an interactive-only version of ONESELF, including the Web forum, the chat room, and the generic sections. Static sections were not implemented in this version. Patients in group 3 (n = 55) were given the full version of ONESELF, including both static and interactive components.

The rationale behind this choice goes back to the functional approach to interactivity. While the library, the virtual gymnasium, and the testimonials did not allow any input from the user other than traditional hypertextual navigation, the Web forum and the chat room enabled synchronous and asynchronous interactions. Although this kind of approach has been criticized for being too generic [15], it is useful to capture the contribution of different enabling functions [70], and it was adopted in other studies on interactivity as well (eg, [53,71-73]).

Figure 2. Attrition flow from the initial recruitment phase. G1 = static-only study group; G2 = interactive-only study group; G3 = full application study group.



Measurements

Each construct of the eHealth intervention effect model was translated into operational measures. The assessments were conducted using standardized online questionnaires. The questionnaires were pretested for face and content validity with two focus groups, with 4 health professionals and 4 patients respectively. The measures used in the study were as follows.

Demographic Characteristics

Demographics in the questionnaires included age, gender, and level of education. We used age as an exogenous covariate in the model analysis, since it can discriminate other focal constructs (eg, level of pain varies with age). Gender was measured but was not controlled for, since the large majority of FMS patients are female and our sample reflected this [74].

Knowledge

Knowledge was measured following the approach of critical or integrative models of health literacy. Thus, we assessed knowledge with 10 multiple choice questions adapted from the Mayo Clinic website relating to FMS symptoms, etiology, treatments, and management. Each answer was coded 1 when correct and 0 when incorrect. The final measure of knowledge was obtained by a mean score calculation of the 10 items, with a theoretical range from 0 (no correct responses) to 1 (all correct responses).

Empowerment

Empowerment was measured according to the scale proposed by Spreitzer [75] but adapted to the FMS domain. This measure reflects the multidimensionality of the construct of empowerment, which is a combination of meaning, competence, self-determination, and impact. Each one of the subdimensions is treated as a latent construct with three observed indicators. Each indicator was measured on a 7-point Likert scale.

Health Outcomes

Health outcomes were measured with the FIQ [64,65] in its Italian version [76]. The FIQ is a validated questionnaire that consists of 20 indicators to assess patients' disability to carry out everyday activities, patients' intensity of pain, and the interference of FMS with patients' sleep and emotional state. The FIQ provides a single score ranging from 0 to 100, where a higher score indicates a greater impact of FMS on the patient. For this reason, it should be considered a measure of negative health outcomes. According to Bennett [65], the average FMS patient scores about 50. Because of the high theoretical variance of this measure compared with the others', we transformed the FIQ raw score on a 0–10 scale and used the result as a single manifest indicator throughout the analyses.

Additional Covariates

The number of years since the first diagnosis was measured as a single additional covariate. By doing so, it is possible to control for potential differences that are due to the illness experience accumulated over time.

In addition to this set of self-report measures, we operationalized functional interactivity by delivering different versions of ONESELF to the three groups of patients, as explained above.

Analyses

The main analyses were conducted using structural equation modeling techniques [77]. SPSS AMOS 18 (IBM Corporation, Somers, NY, USA) was used for the analyses. Specifically, we examined the effect of interactivity on knowledge, meaning, competence, self-determination, and impact. Because of the relatively small sample size, it was not advisable to run a test on a model including the five focal dependent variables at the same time. Thus, a structural equation model reflecting an analysis of covariance for the analysis of change was implemented and tested five times, varying the five focal dependent variables. Additionally, to determine whether a mediating effect existed, we ran an analysis of variance with the fibromyalgia impact as dependent variable and study condition as factor.

The model accounts for autoregressive effects of the focal dependent variables (knowledge and the four empowerment dimensions) at time 1 on the same variable at time 2, as well as outcomes at time 1 on outcomes at time 2. This strategy allows controlling for pretest scores and adjustment for measurement errors. To reduce sampling error, we included age and years since first diagnosis in the model as additional covariates, as well as the initial score on health outcomes. The main endogenous variable was the patients' assessment of fibromyalgia impact.

Results

Sample Characteristics and Application Usage

The sample included 165 patients with FMS. Age ranged from 27 to 72 (mean 49.93, SD 9.93) years. Most patients were female (157, 95.2% women; 8, 4.8% men). The modal educational level was high school (63.6%); 43 (26.1%) had a lower educational level and 17 (10.3%), a higher. The mean time since first diagnosis was 5.7 (SD 4.8) years. At baseline, the average FIQ score was 5.98 (SD 1.82). Table 1 presents descriptive sample characteristics.

Table 2 reports the mean values of the relevant constructs of the model, as measured at pretest (T1) and posttest (T2).

Table 1. Descriptive sample characteristics.

Categorical or dichotomous variable	Range	n	%
Gender, n (%)			
Female		157	95.2%
Male		8	4.8%
Level of education, n (%)			
None		4	2.4%
Elementary school		9	5.5%
Middle school		30	18.2%
High school		105	63.6%
University		17	10.3%
Continuous variables, mean (SD)		**Mean**	**SD**
Age (years)	27–72	49.93	9.93
Years since first diagnosis	0–40	5.7	4.8
Health outcomes (FIQ[a]) at baseline	0.73–9.07	5.98	1.82

[a] Fibromyalgia Impact Questionnaire, score transformed on a scale of 0 to 10.

Table 2. Descriptive statistics (mean, SD) for experimental measures for study groups 1–3 at time 1 (pretest) and time 2 (posttest)[a].

Variable name	Time 1			Time 2		
	Group 1 (n = 55)	Group 2 (n = 55)	Group 3 (n = 55)	Group 1 (n = 50)	Group 2 (n = 51)	Group 3 (n = 50)
Health outcomes	5.9 (1.7)	6.1 (1.7)	5.9 (1.8)	5.9 (1.7)	6.0 (1.5)	5.8 (1.8)
Knowledge	6.1 (1.8)	5.6 (1.9)	5.9 (1.6)	6.4 (1.9)	6.1 (1.7)	6.4 (1.6)
Meaning-1	6.3 (1.4)	6.4 (1.3)	6.5 (0.9)	6.6 (0.8)	6.3 (1.3)	6.4 (1.0)
Meaning-2	6.0 (1.6)	6.1 (1.2)	5.9 (1.5)	6.2 (1.3)	6.1 (1.3)	5.8 (1.5)
Meaning-3	6.3 (1.2)	6.2 (1.2)	6.2 (1.2)	6.4 (1.0)	6.3 (1.0)	5.9 (1.4)
Competence-1	5.3 (1.6)	5.1 (1.8)	5.2 (1.7)	5.6 (1.4)	5.5 (1.4)	5.3 (1.7)
Competence-2	4.5 (1.8)	4.4 (2.0)	4.6 (1.9)	5.0 (1.6)	4.8 (1.8)	4.8 (1.9)
Competence-3	5.3 (1.8)	5.0 (1.9)	5.7 (1.3)	5.3 (1.4)	4.9 (1.5)	5.5 (1.4)
Self-determination-1	4.3 (1.9)	4.2 (1.9)	4.6 (2.0)	4.7 (1.7)	4.9 (1.6)	5.1 (1.8)
Self-determination-2	3.5 (2.3)	3.5 (1.9)	4.2 (2.2)	3.7 (1.9)	4.1 (2.0)	4.5 (2.0)
Self-determination-3	4.0 (2.3)	4.0 (1.9)	4.1 (2.0)	4.1 (2.0)	4.5 (1.9)	4.5 (2.0)
Impact-1	4.4 (2.0)	4.4 (1.9)	4.2 (2.2)	5.3 (1.4)	5.1 (1.7)	4.8 (1.8)
Impact-2	3.5 (1.8)	3.6 (2.0)	3.8 (2.1)	4.2 (1.7)	4.5 (1.7)	4.2 (1.8)
Impact-3	3.7 (2.0)	4.3 (2.0)	4.0 (1.9)	4.5 (1.8)	5.0 (1.5)	4.1 (1.9)

[a] Three items compose each empowerment construct.

Regarding usage of the application, the mean number of visits as recorded by a log file analyzer was 13.27 (SD 6.68). The mean time spend on the website was 4.8 (SD 3.2) minutes per visit. Table 3 reports some usage data related to the main features implemented in the experimental conditions. Given that these figures suggest that the application was not used very often, in a preliminary phase of the analysis, we introduced the number of visits and the time spent on the application as potential moderators of the relationships implied by the theoretical model. For the sake of brevity, we cannot report the whole preliminary evaluation, but the main result was that, at the observed level of usage, we found no significant moderating effect. Although this result does not exclude the presence of a trend (ie, that running the study for a longer period would generate significant interaction effects), it suggests that usage of the application had no strong influence on the relationships investigated in the present analysis. It must be acknowledged, however, that the scant usage of the system may understate the results of the present analysis. Under the assumption that more frequent usage should lead to larger effects, the differential impact of some features that were seldom used (eg, the chat room) might stay covered. We accounted for this in the power and effect size considerations that are detailed in Multimedia Appendix 1 [78-82].

Table 3. Usage data (mean, SD) across experimental conditions.

Variable	Group 1	Group 2	Group 3
Time per visit (minutes)	4.2 (10.2)	4.8 (16.4)	5.4 (13.2)
Number of visits to library	8.13 (3.5)	NA[a]	7.6 (3.5)
Number of visits to virtual gymnasium	9.0 (3.1)	NA	10.2 (0.6)
Number of visits to testimonials	12.9 (2.5)	NA	11.4 (1.2)
Number of visits to Web forum	NA	10.5 (22.1)	8.5 (8.1)
Number of visits to chat room	NA	2.1 (1.3)	1.8 (4.3)

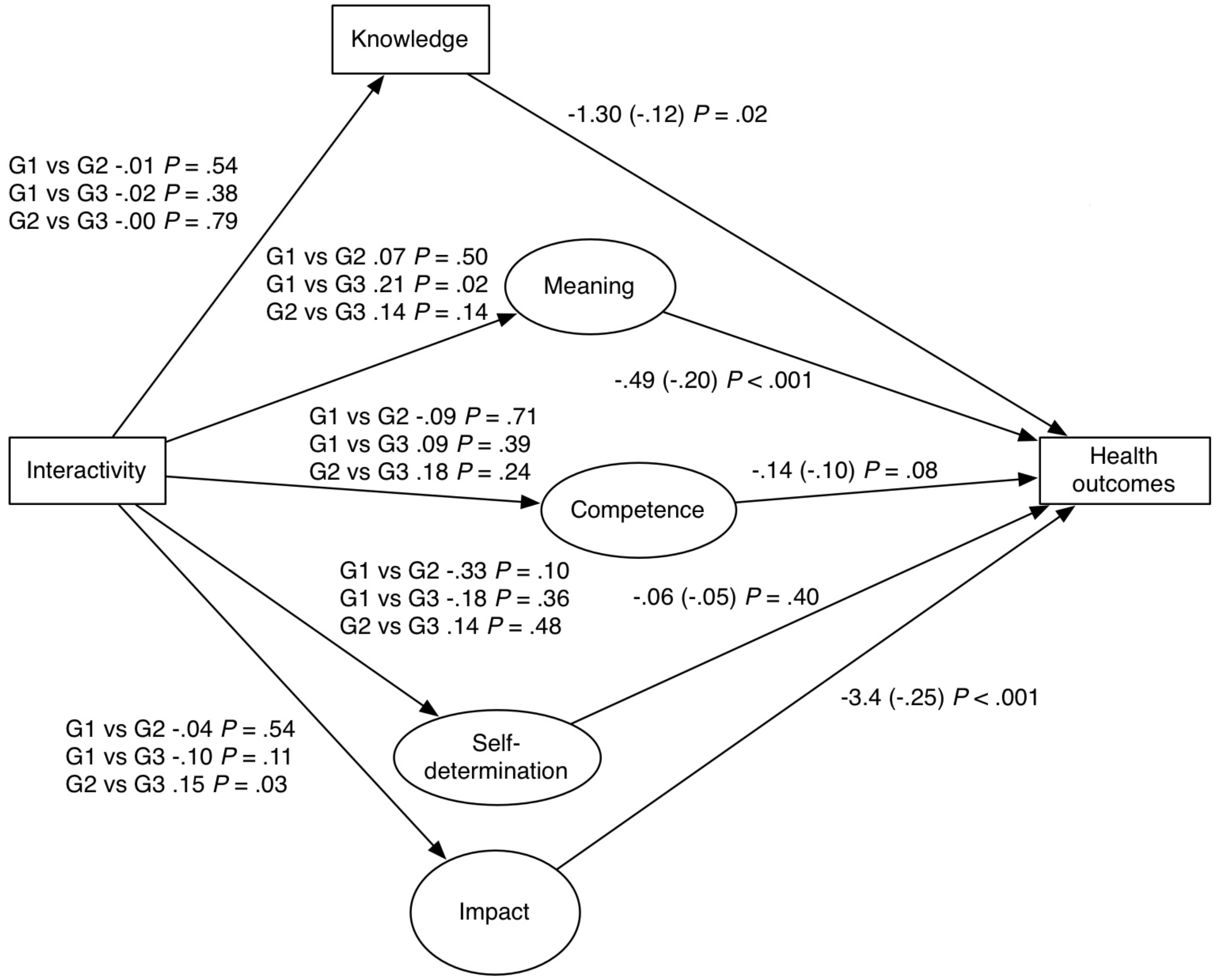
[a] Not applicable.

Testing Hypotheses

Our hypotheses detailed the role of knowledge and patient empowerment as mediating variables of a possible effect of the availability of interactive elements in an eHealth intervention. Results on the differences between the three study groups indicate no significant difference in fibromyalgia impact ($F_{2,148} = 0.824$, $P = .44$). This means that the two groups who were offered interactive elements did not benefit more from the website than the group offered only static elements. The absence of a difference renders the question of mediating variables obsolete. However, we did compute the model (and report it here) to determine at what stage the hypothesized relationships failed to emerge.

The following five steps address one of the five focal dependent variables (knowledge, and the four dimensions of empowerment). Each section first reports the results of the effects of one of the focal dependent variables on the endogenous variable, health outcomes (operationalized as patient assessment of fibromyalgia impact), and then the effects of the main independent variable on the focal dependent variable. This means that hypotheses 1 and 6 are treated in one step, as are hypotheses 2 and 7, 3 and 8, an so on. To make the presentation clearer, we summarize the results in Figure 3. All five models are available in Multimedia Appendix 2, Multimedia Appendix 3, Multimedia Appendix 4, Multimedia Appendix 5, and Multimedia Appendix 6. The five steps do not mention the autoregressive effects of the focal dependent variables at time 1 on time 2 and the endogenous variable at time 1 on time 2. All these effects were highly significant.

Figure 3. Overall model results. G1 = static-only study group; G2 = interactive-only study group; G3 = full application study group. Standardized coefficients are in parentheses (see Multimedia Appendix 2-7 for details).



The Role of Knowledge

As predicted by the hypothesized model (hypothesis 6), *knowledge* at posttest significantly predicted health outcomes at posttest. On average, for every 1 unit that knowledge increased, negative health outcomes were predicted to decrease by 1.3 units, everything else being equal. We further hypothesized that interactivity enabled in the application increases patients' knowledge (hypothesis 1). This result should reflect a significant difference between groups 2 and 3 (higher score) and group 1 (lower score). Despite the significant effect of knowledge on patient assessment of fibromyalgia impact, interactivity did not have any impact on knowledge, as the absence of significant differences between the study groups testifies. Given this result, hypothesis 1 was not confirmed.

The Role of Meaning

As predicted by the hypothesized model (hypothesis 7), *meaning* at posttest significantly predicted assessment of impact at posttest. On average, for every 1 unit that meaning increased, negative assessments were predicted to decrease by 0.49 units, everything else being equal. We further hypothesized that interactivity enabled in the application should increase the patients' meaning score (hypothesis 2). There should be a significant difference between the groups with (higher score) and without (lower score) interactivity. The only direct effect of functional interactivity on meaning can be observed in the significant difference between group 1 (ie, patients provided with the static version of the tool) and group 3 (ie, patients provided with the full-fledged application). However, the direction of the effect is not in as expected. Patients who had access to both static and interactive elements on the site scored significantly lower in meaning than did people with the static version only. Therefore, hypothesis 2 was not supported.

The Role of Competence

Contrary to the eHealth effect model prediction (hypothesis 8), *competence* at posttest did not significantly predict patient assessment of fibromyalgia impact at posttest. We also hypothesized that the interactivity enabled in the application should increase the patients' competence score (hypothesis 3). However, interactivity did not have any such impact on competence, as testified by the absence of significant differences between the study groups. Given this result, hypothesis 3 was not confirmed.

The Role of Self-determination

Contrary to the eHealth effect model prediction, *self-determination* at posttest did not significantly predict patient assessment of fibromyalgia impact at posttest. Thus, hypothesis 9 was not supported. We also hypothesized that the interactivity enabled in the application should increase the patients' self-determination score (hypothesis 4). No significant direct effect of functional interactivity was observed, as indicated by the absence of significant differences between the study groups. Therefore, hypothesis 4 was not confirmed.

The Role of Impact

In accordance with the eHealth effect model prediction (hypothesis 10), *impact* at posttest significantly predicted health outcomes at posttest. The impact of the syndrome as perceived by the patient appears to make the largest difference in health outcomes. From this experimental data, we observed that for every 1 unit impact score increased, negative health outcomes were predicted to decrease on average by 3.4 units (or a quarter of a standard deviation), which is indeed a relevant finding. Given this result, hypothesis 10 was confirmed. We also hypothesized that the interactivity enabled in the application should increase the patients' impact score (hypothesis 5). This result should reflect a significant difference between groups 2 and 3 (higher scores) and group 1 (lower score). The only significant difference, though, was found between groups 2 and 3. Patients provided with the interactive-only application (group 2) scored significantly higher on impact than did patients who had access to both the interactive and the static elements (group 3). Hypothesis 5, consequently, was not confirmed.

Although not of primary relevance in this analysis, it is also interesting that the number of years that patients had FMS influenced their impact score. On average, for every 1 additional year of illness experience, the patients' impact score was predicted to decrease by 0.04 units. The total effect of years of illness experience on health outcomes was 0.07, meaning that an additional year of FMS on average increased negative health outcomes by 0.07 units, and did so by decreasing the impact that patients perceived the syndrome to have on them. In other words, the lower a patient's experienced control over the syndrome (impact score), the worse the health outcomes score will be.

Discussion

In summary, the presence of interactive elements in our eHealth intervention did not affect knowledge, did not affect patient empowerment in the expected direction (but reduced the empowerment dimension of meaning), and did not improve the health outcome of perceived fibromyalgia impact. In contrast to other studies, ours did not find beneficial effects of functional interactivity. However, knowledge and two dimensions of empowerment (meaning and self-determination) did affect health outcomes. Overall, the experimental findings suggest that only some of the hypothesized relationships held true (see Figure 3 and Multimedia Appendix 7).

The strongest relationships concern the effects of knowledge, meaning, and impact on health outcomes. These results show that cognition and empowerment, at least in some of its dimensions, are strong predictors of health outcomes, and eHealth applications should try to target and enhance these individual characteristics.

The relationships between functional interactivity and knowledge, empowerment, and health outcomes was not supported. Indeed, functional interactivity had an impact on two empowerment dimensions (meaning and impact), but in both cases the effect was not in line with the original predictions. In all cases, patients provided with the less complete version of the eHealth application reported higher scores on these dimensions of empowerment than did those exposed to the full application. This unexpected result may be due to the possibility that viewing static information helped individuals in group 1 make more sense of their condition than did a combination of static and interactive features. An alternative explanation is that patients provided with the full-featured application were overloaded by the functions and this badly affected their ability to understand their health condition.

Both unexpected results may also be unintended consequences of item wording. The empowerment dimension of impact was measured with items that varied the notion of control and influence over a person's quality of life and health status in general, while all items used for the other three dimensions specifically referred to FMS. Adding interactive to static features may have increased (or, given the other results, more likely not affected) a sense of impact related specifically to FMS (which was hypothesized), but at the same time reduced a sense of general impact (which is closer to what was actually measured). This highlights the general possibility that condition-specific eHealth applications, along with their intended effects on condition-specific empowerment, may have adverse effects on a general sense of empowerment. Or, in plain words: patients who are effectively being told their condition is manageable and controllable may get the impression that other conditions are not. This has implications for the conceptualization and measurement of empowerment as well as for the design of eHealth applications.

The empowerment dimension of meaning was measured with items varying the notion of the importance of coping with FMS. The finding that the scores decreased when interactive features were added to static ones can be interpreted in two ways, depending on how respondents may have understood the term *importance*. If importance was understood as relevance, the unexpected finding could be interpreted as an undesirable distraction from actually important matters, caused by the interactive elements. If, however, importance was understood as urgency, the result could indicate a desirable effect: as patients improved their ways of coping (not least by the eHealth application), the subject of coping became less urgent to think about. Methodologically, this would imply that the measure used for this dimension needs to be revised.

The relationships between interactivity and knowledge, self-determination, or competence were not confirmed. Regarding knowledge, a plausible explanation is that the patients were generally knowledgeable about the condition even before the pretest (indeed, the mean score at pretest was 0.60, SD 0.2).

Because of this, the application did not make a significant difference with regard to this construct. The same holds true for competence, for which initial score was relatively high across the study groups (5.25, SD 1.7). Regarding self-determination, at least two interpretations seem reasonable. First, while self-determination is a construct referring to individuals' autonomy, the application was not meant to increase independence in facing the condition, but rather to support embeddedness within a community of patients. Second, interactivity is only one of the factors that may contribute to increase the overall sense of autonomy of a patient with a chronic condition. Although it is theoretically reasonable to include this dimension as part of empowerment, the specific context of investigation may not fully reflect its relevance. Further theoretical work is needed to expand this issue.

It is difficult to relate these results to the empirical literature, which is scarce. Empirical studies that are concerned with interactivity in health applications, if they exist at all, address the effects on quite different dependent variables, such as attitudes to and satisfaction with the application [53], efficiency of professional learning [54], risk assessments [56], and, for political campaign websites, amount of time spent on the site and content recall [73]. Warnick et al [73] pointed out, quite in line with our results, that too much interactivity can decrease content recall. Neither empowerment nor health outcomes have, to our knowledge, been empirically related to interactivity, and of course the application we used has not been studied, nor has our measure of empowerment been used, as it was developed for this study. Hopeful expectations for a beneficial impact of eHealth interactivity (eg, [10]) should be tempered unless other empirical research supports them.

Generally, support for the mediator perspective on interactivity effects (with knowledge and empowerment functioning as mediators) was weak because the hypothesized effects of interactivity on knowledge and empowerment could not be established.

Study Limitations

Some limitations of this study should be acknowledged. In respect to interactivity, the preferred approach mostly relied on the functional perspective [18]. This choice allowed testing for the different enabling functions implemented in the eHealth intervention. However, the dimension of perceived interactivity [15,17] may play a major role as well. In this study, we assumed that the chosen functions had a different degree of interactivity without considering whether the patients perceived this difference. Ideally, both functional and perceived interactivity must be considered in an integrated and holistic model, as shown in a recent contribution to the study of this construct [83].

The construct of knowledge was largely based on the critical and integrative perspective on health literacy. However, the aggregate average score provided by the knowledge test may not fully capture its theoretical complexity. For example, in Schulz and Nakamoto's conception, individual judgment skills should ideally be included in a full measure of integrative health literacy [26].

Regarding our methods, an objection can be advanced for the lack of a pure control group in the study phase. Patients provided with the static version of the application were used as a baseline, but we did not include patients who were not exposed to the application. This limitation was mainly linked to feasibility issues in reaching a sufficiently large number of FMS patients. In addition, it is not uncommon to provide people in the control group with some kind of neutral intervention, and this principle was applied in the study.

Implications and Future Directions

Despite these limitations, this study provided interesting insights into the differential effect of enabling functions implemented in eHealth interventions, providing some evidence of the impact of interactivity in the health context.

First, the combination of interactive and static features is not necessarily a turnkey solution to enhance patients' knowledge and empowerment. For example, the presence of interactive features did not significantly improve the individual level of knowledge. When tested as stand-alone feature, however [84], the library of information as implemented in ONESELF had an impact on patients' knowledge. Given that we measured knowledge in relation to FMS, however, this result is hardly generalizable to other health conditions. More holistic and comprehensive measures of individual knowledge would help in comparing results on this construct.

A second consideration relates to other studies [56,70,73], which found that interactivity can enhance individual efficacy and self-determination. The present study rejected this claim by testing the effects of this construct on a more holistic conceptualization of empowerment.

A third, more methodological observation is that testing individual features separately or grouped into clusters decreases the effect size of the overall application. The present analysis is not suggesting that one should limit an eHealth intervention to some functions, but rather that the choice of these features should be theory or goal driven. This is even more important considering that our results showed that the effect of the full-fledged application was generally weaker than a more focused version. It is possible that patients provided with the more complex application needed more time before mastering the system and the impact was somehow delayed. A longitudinal evaluation may provide further insights into this hypothesis. Indeed, the stability of a certain effect is rarely tested in eHealth studies [85]. This is partly due to the costs associated with longitudinal evaluations and partly linked to the high level of attrition [82] between pretest and posttest.

At the time of writing this paper, research on health communication and the Internet is rapidly evolving. While the debate on the theories and constructs that must be considered in the design and evaluation of eHealth applications is lively, several efforts are being made to define standard procedures to conduct and report empirical studies on eHealth interventions [86]. In line with the scientific debate around eHealth, an underlying theme reappeared throughout this work: the need to integrate a theory-driven approach into eHealth research. The model of eHealth effects defined and tested in this study is a first step on this route, but is far from complete, especially as the model did not hold. Future directions of this endeavor should focus on adapting, complementing, and refining the model and its evaluation. The constructs included in the model of eHealth effects posited some limitations at the theoretical level. For example, the construct of interactivity should be refined to include both functional and perceived perspectives. Although Rafaeli and Ariel [15] argued that interactivity is primarily a subjective construct, exclusion of interactivity at the functional level would not provide any insight into the mechanisms that occur within eHealth applications. A more integrative perspective such as the one recently proposed by Yoo [83], who combined medium and audience interactivity, is preferable and should be pursued to refine the model.

Along the same lines, while the general validity of the model of empowerment as proposed by Thomas and Velthouse [40] and Spreitzer [75] was assessed in the health care setting, other models of empowerment (eg, [87]) might be considered and integrated to develop an improved measure of this construct.

In general terms, a model should be as predictive and explanatory as it is parsimonious. Without overexpanding the eHealth effects model, there are some constructs that should be considered for its improvement.

The first two of these constructs relate to the literacy and knowledge dimensions. According to Schulz and Nakamoto [26], the construct of judgment skills as the ability to integrate procedural and declarative knowledge into one's own experience should be included in the model to complement the construct of knowledge. Before integrating the judgment dimension, however, further effort should be put into translating it into operational measures. Additionally, the construct of media literacy [88] can further improve the cognitive component of the model. Rosenbaum and colleagues [89] defined media literacy as the understanding of the influences that occur between the media, producers, and users. Thus, media literacy can be considered a mediator between exposure to a certain medium and its effects on opinions and behavior. As such, together with health literacy, it can make a significant difference in the effects of eHealth interventions.

Eventually, an important construct that should be included in the model is computer self-efficacy, or one's confidence in using a computer and gaining benefits from it [90,91]. Studies on eHealth often have generalizability issues because the reference population is composed of people who are implicitly considered skilled in using a computer. The CONSORT-EHEALTH guidelines for conducting research on eHealth stress the need to avoid this assumption and urge systematically testing individuals' ability and confidence in using technology [86]. The construct of computer self-efficacy and its measurement can serve to control for this aspect, while providing further insights into the mechanisms that determine the effectiveness of eHealth interventions.

Acknowledgments

The authors wish to thank the *Lega Ticinese per la Lotta Contro il Reumatismo* (LTCR) for their invaluable support of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH Checklist V 1.6.

[PDF File (Adobe PDF File), 544KB - jmir_v14i4e105_app1.pdf]

Multimedia Appendix 2

Model of the effect of interactivity on knowledge.

[PDF File (Adobe PDF File), 66KB - jmir_v14i4e105_app2.pdf]

Multimedia Appendix 3

Model of the effect of interactivity on meaning.

[PDF File (Adobe PDF File), 66KB - jmir_v14i4e105_app3.pdf]

Multimedia Appendix 4

Model of the effect of interactivity on competence.

[PDF File (Adobe PDF File), 66KB - jmir_v14i4e105_app4.pdf]

Multimedia Appendix 5

Model of the effect of interactivity on self-determination.

[PDF File (Adobe PDF File), 66KB - jmir_v14i4e105_app5.pdf]

Multimedia Appendix 6

Model of the effect of interactivity on impact.

[PDF File (Adobe PDF File), 66KB - jmir_v14i4e105_app6.pdf]

Multimedia Appendix 7

Summary of hypotheses tested and subsequent results.

[PDF File (Adobe PDF File), 48KB - jmir_v14i4e105_app7.pdf]

Multimedia Appendix 8

CONSORT-EHEALTH Checklist V 1.6 [92].

[PDF File (Adobe PDF File), 544KB - jmir_v14i4e105_app8.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 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
2. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
3. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
4. Ybarra ML, Eaton WW. Internet-based mental health interventions. Ment Health Serv Res 2005 Jun;7(2):75-87. [Medline: 15974154]
5. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]
6. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
7. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
8. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
9. Abbate J. Inventing the Internet. Cambridge, MA: MIT Press; 1999.
10. Fotheringham MJ, Owies D, Leslie E, Owen N. Interactive health communication in preventive medicine: internet-based strategies in teaching and research. Am J Prev Med 2000 Aug;19(2):113-120. [Medline: 10913902]
11. 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]
12. Neuhauser L, Kreps GL. Rethinking Communication in the E-health Era. J Health Psychol 2003 Jan;8(1):7-23. [doi: 10.1177/1359105303008001426] [Medline: 22113897]
13. Kalyanaraman S, Sundar SS. The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? J Commun 2006;56(1):110-132. [doi: 10.1111/j.1460-2466.2006.00006.x]
14. Bock BC, Graham AL, Whiteley JA, Stoddard JL. A review of web-assisted tobacco interventions (WATIs). J Med Internet Res 2008;10(5):e39 [FREE Full text] [doi: 10.2196/jmir.989] [Medline: 19000979]
15. Rafaeli S, Ariel Y. Assessing interactivity in computer-mediated research. In: Joinson A, McKenna K, Postmes T, Reips U, editors. The Oxford Handbook of Internet Psychology. Oxford: Oxford University Press; 2009:71-88.
16. Downes EJ, Mcmillan SJ. Defining interactivity: a qualitative identification of key dimensions. New Media Society 2000;2(2):157-179. [doi: 10.1177/14614440022225751]
17. Rafaeli S. Interactivity: from new media to communication. In: Hawkins RP, Wieman JM, Pingree S, editors. Advancing Communication Science: Merging Mass and Interpersonal Processes. Newbury Park, CA: Sage Publications; 1988:110-134.
18. Sundar SS. Theorizing Interactivity?s effects. Inf Soc 2004;20(5):385-389. [doi: 10.1080/01972240490508072]

19. Steuer J. Defining virtual reality: dimensions determining telepresence. J Commun 1992;42(4):73-93. [doi: 10.1111/j.1460-2466.1992.tb00812.x] [Medline: 22299406]
20. Rogers EM. Diffusion of Innovations. 4th edition. New York, NY: Free Press; 1995.
21. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. JAMA 1999 Feb 10;281(6):552-557. [Medline: 10022112]
22. Williams MV, Baker DW, Honig EG, Lee TM, Nowlan A. Inadequate literacy is a barrier to asthma knowledge and self-care. Chest 1998 Oct;114(4):1008-1015 [FREE Full text] [Medline: 9792569]
23. Institute of Medicine. Health Literacy: A Prescription to End Confusion. Washington, DC: National Academies Press; 2004.
24. Nutbeam D. Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int 2000;15(3):259-267. [doi: 10.1093/heapro/15.3.259]
25. Nutbeam D. Health promotion glossary. Health Promot Int 1998;13(4):349-364. [doi: 10.1093/heapro/13.4.349]
26. Schulz PJ, Nakamoto K. Emerging themes in health literacy. Stud Commun Sci 2005;5(2):1-10.
27. Rappaport J. In praise of paradox: a social policy of empowerment over prevention. Am J Community Psychol 1981 Feb;9(1):1-25. [Medline: 7223726]
28. Rappaport J. Studies in empowerment: introduction to the issue. Prev Hum Serv 1984;3(2):1-7.
29. Zimmerman MA, Rappaport J. Citizen participation, perceived control, and psychological empowerment. Am J Community Psychol 1988 Oct;16(5):725-750. [Medline: 3218639]
30. Perkins DD, Zimmerman MA. Empowerment theory, research, and application. Am J Community Psychol 1995 Oct;23(5):569-579. [Medline: 8851340]
31. Rappaport J. Terms of empowerment/exemplars of prevention: toward a theory for community psychology. Am J Community Psychol 1987 Apr;15(2):121-148. [Medline: 3604997]
32. Wallerstein N, Bernstein E. Empowerment education: Freire's ideas adapted to health education. Health Educ Q 1988;15(4):379-394. [Medline: 3230016]
33. Schulz AJ, Israel BA, Zimmerman MA, Checkoway BN. Empowerment as a multi-level construct: perceived control at the individual, organizational and community levels. Health Educ Res 1995;10(3):309-327.
34. Israel BA, Checkoway B, Schulz A, Zimmerman M. Health education and community empowerment: conceptualizing and measuring perceptions of individual, organizational, and community control. Health Educ Q 1994;21(2):149-170. [Medline: 8021145]
35. Zimmerman MA. Psychological empowerment: issues and illustrations. Am J Community Psychol 1995 Oct;23(5):581-599. [Medline: 8851341]
36. Deci EL. Intrinsic Motivation. New York, NY: Plenum Press; 1975.
37. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
38. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
39. Conger JA, Kanungo RN. The empowerment process: integrating theory and practice. Acad Manag Rev 1988;13(3):471-482. [doi: 10.2307/258093]
40. Thomas KW, Velthouse BA. Cognitive elements of empowerment: an interpretive model of intrinsic task motivation. Acad Manag Rev 1990;15(4):666-681. [doi: 10.2307/258687]
41. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [Medline: 2782727]
42. Motivation reconsidered: the concept of competence. Psychol Rev 1959 Sep;66:297-333. [Medline: 13844397]
43. Deci EL, Ryan RM. Intrinsic Motivation and Self-determination in Human Behavior. New York, NY: Plenum; 1985.
44. Wallerstein N. Powerlessness, empowerment, and health: implications for health promotion programs. Am J Health Promot 1992;6(3):197-205. [Medline: 10146784]
45. Bergsma LJ. Empowerment education: The link between media literacy and health promotion. Am Behav Sci 2004;48(2):152-164.
46. Wallerstein N. What is the Evidence on Effectiveness of Empowerment to Improve Health?. Copenhagen: WHO Regional Office for Europe; 2006.
47. Bergsma LJ, Carney ME. Effectiveness of health-promoting media literacy education: a systematic review. Health Educ Res 2008 Jun;23(3):522-542 [FREE Full text] [doi: 10.1093/her/cym084] [Medline: 18203680]
48. Fox NJ, Ward KJ, O'Rourke AJ. The 'expert patient': empowerment or medical dominance? The case of weight loss, pharmaceutical drugs and the Internet. Soc Sci Med 2005 Mar;60(6):1299-1309. [doi: 10.1016/j.socscimed.2004.07.005] [Medline: 15626525]
49. Amichai-Hamburger Y, McKenna KYA, Tal SA. E-empowerment: empowerment by the internet. Comput Human Behav 2008;24(5):1776-1789. [doi: 10.1016/j.chb.2008.02.002]
50. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]

51. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
52. Schulz PJ, Nakamoto K. “Bad” literacy, the Internet, and the limits of patient empowerment. In: AAAI Spring Symposium Series. 2011 Presented at: Artificial Intelligence & Health Communication; Mar 21-23, 2011; Stanford, CA, USA URL: http://www.aaai.org/ocs/index.php/SSS/SSS11/paper/view/2430/2853
53. Teo HH, Oh LB, Liu C. Wei K-K. An empirical study of the effects of interactivity on web user attitude. Int J Hum Comput Stud 2003;58:281-305. [doi: 10.1016/S1071-5819(03)00008-9]
54. Cook DA, Levinson AJ, Garside S. Time and learning efficiency in Internet-based learning: a systematic review and meta-analysis. Adv Health Sci Educ Theory Pract 2010 Dec;15(5):755-770. [doi: 10.1007/s10459-010-9231-x] [Medline: 20467807]
55. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2:183-196. [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]
56. Zikmund-Fisher BJ, Dickson M, Witteman HO. Cool but counterproductive: interactive, Web-based risk communications can backfire. J Med Internet Res 2011;13(3):e60 [FREE Full text] [doi: 10.2196/jmir.1665] [Medline: 21868349]
57. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum 1990 Feb;33(2):160-172. [Medline: 2306288]
58. Belt NK, Kronholm E, Kauppi MJ. Sleep problems in fibromyalgia and rheumatoid arthritis compared with the general population. Clin Exp Rheumatol 2009;27(1):35-41. [Medline: 19327227]
59. Clauw DJ. Assessing and diagnosing fibromyalgia in the clinical setting. J Clin Psychiatry 2008 Nov;69(11):e33. [Medline: 19200425]
60. Arnold LM, Bradley LA, Clauw DJ, Glass JM, Goldenberg DL. Evaluating and diagnosing fibromyalgia and comorbid psychiatric disorders. J Clin Psychiatry 2008 Oct;69(10):e28. [Medline: 19192431]
61. Wilson HD, Robinson JP, Turk DC. Toward the identification of symptom patterns in people with fibromyalgia. Arthritis Rheum 2009 Apr 15;61(4):527-534 [FREE Full text] [doi: 10.1002/art.24163] [Medline: 19333980]
62. Buskila D, Cohen H. Comorbidity of fibromyalgia and psychiatric disorders. Curr Pain Headache Rep 2007 Oct;11(5):333-338. [Medline: 17894922]
63. Goldenberg DL. Psychological symptoms and psychiatric diagnosis in patients with fibromyalgia. J Rheumatol Suppl 1989 Nov;19:127-130. [Medline: 2607508]
64. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991 May;18(5):728-733. [Medline: 1865419]
65. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. Clin Exp Rheumatol 2005;23(5 Suppl 39):S154-S162. [Medline: 16273800]
66. Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. Arthritis Res Ther 2009;11(4):R120 [FREE Full text] [doi: 10.1186/ar2783] [Medline: 19664287]
67. Salaffi F, Sarzi-Puttini P, Girolimetti R, Gasparini S, Atzeni F, Grassi W. Development and validation of the self-administered Fibromyalgia Assessment Status: a disease-specific composite measure for evaluating treatment effect. Arthritis Res Ther 2009;11(4):R125 [FREE Full text] [doi: 10.1186/ar2792] [Medline: 19686606]
68. Caiata Zufferey M, Schulz PJ. Self-management of chronic low back pain: an exploration of the impact of a patient-centered website. Patient Educ Couns 2009 Oct;77(1):27-32. [doi: 10.1016/j.pec.2009.01.016] [Medline: 19321286]
69. Schulz PJ, Rubinelli S, Caiata-Zufferey M, Hartung U. Coping with chronic lower back pain: designing and testing the online tool ONESELF. J Comput Mediat Commun 2010;15(4):625-645. [doi: 10.1111/j.1083-6101.2009.01509.x]
70. Sundar SS. Social psychology of interactivity in human-website interaction. In: Joinson A, McKenna A, Postmes T, Reips U, editors. The Oxford Handbook of Internet Psychology. Oxford: Oxford University Press; 2007:89-102.
71. McMillan SJ, Hwang JS. Measures of perceived interactivity: an exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. J Advert 2002;31(3):29-42 [FREE Full text]
72. Amichai-Hamburger Y, Fine A, Goldstein A. The impact of Internet interactivity and need for closure on consumer preference. Comput Human Behav 2004;20(1):103-117. [doi: 10.1016/S0747-5632(03)00041-4]
73. Warnick B, Xenos M, Endres D, Gastil J. Effects of campaign-to-user and text-based interactivity in political candidate campaign web sites. J Comput Mediat Commun 2006;10(3):00. [doi: 10.1111/j.1083-6101.2005.tb00253.x]
74. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995 Jan;38(1):19-28. [Medline: 7818567]
75. Spreitzer GM. Psychological empowerment in the workplace: dimensions, measurement, and validation. Acad Manag J 1995;38(5):1442-1465. [doi: 10.2307/256865]
76. Sarzi-Puttini P, Atzeni F, Fiorini T, Panni B, Randisi G, Turiel M, et al. Validation of an Italian version of the Fibromyalgia Impact Questionnaire (FIQ-I). Clin Exp Rheumatol 2003;21(4):459-464. [Medline: 12942697]

77. Kline RB. Principles and Practice of Structural Equation Modeling. 2nd edition. New York, NY: Guilford Press; 2005.
78. Jaccard J, Wan CK. LISREL Approaches to Interaction Effects in Multiple Regression. Thousand Oaks, CA: Sage Publications; 1996.
79. Jackson DL. Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. Struct Equ Modeling 2003;10:128-141. [doi: 10.1207/S15328007SEM1001_6]
80. Jaccard J, Turrisi R, Wan CK. Interaction Effects in Multiple Regression. Thousand Oaks, CA: Sage Publications; 2003.
81. Bollen KA, Long JS. Testing Structural Equation Models. Newbury Park, CA: Sage Publications; 1993.
82. 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]
83. Yoo CY. Modeling audience interactivity as the gratification-seeking process in online newspaper. Commun Theory 2011;21(1):67-89. [doi: 10.1111/j.1468-2885.2010.01376.x]
84. Camerini L, Frisch AL, Schulz PJ. Internet-based patient education intervention for fibromyalgia: a model-driven evaluation. In: AAAI Spring Symposium Series. 2011 Presented at: Artificial Intelligence & Health Communication; Mar 21-23, 2011; Stanford, CA, USA URL: http://www.aaai.org/ocs/index.php/SSS/SSS11/paper/view/2376
85. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700. [Medline: 12020305]
86. Baker TB, Gustafson DH, Shaw B, Hawkins R, Pingree S, Roberts L, et al. Relevance of CONSORT reporting criteria for research on eHealth interventions. Patient Educ Couns 2010 Dec;81 Suppl:S77-S86. [doi: 10.1016/j.pec.2010.07.040] [Medline: 20843621]
87. Menon ST. Psychological empowerment: Definition, measurement, and validation. Can J Behav Sci 1999;31(3):161-164. [doi: 10.1037/h0087084]
88. Potter WJ. Theory of Media Literacy: A Cognitive Approach. Thousand Oaks, CA: Sage Publications; 2004.
89. Rosenbaum JE, Beentjes JWJ, Konig RP. Mapping media literacy: key conceptsfuture directions. In: Beck CS, editor. Communication Yearbook 32. New York, NY: Routledge; 2008:313-353.
90. Compeau DR, Higgins CA. Computer self-efficacy: development of a measure and initial test. Manag Inf Syst Q 1995;19(2):189-211. [doi: 10.2307/249688]
91. Poynton, TA. Computer literacy across the lifespan: a review with implications for educators. Comput Human Behav 2005;21(6):861-872. [doi: 10.1016/j.chb.2004.03.004]
92. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

FAS: Fibromyalgia Assessment Status
FIQ: Fibromyalgia Impact Questionnaire
FMS: fibromyalgia syndrome

Edited by R Crutzen; submitted 28.09.11; peer-reviewed by G Kreps, H Hochheiser, L Morrison; comments to author 01.11.11; revised version received 30.01.12; accepted 24.05.12; published 18.07.12.

Please cite as:
Camerini L, Schulz PJ
Effects of Functional Interactivity on Patients' Knowledge, Empowerment, and Health Outcomes: An Experimental Model-Driven Evaluation of a Web-Based Intervention
J Med Internet Res 2012;14(4):e105
URL: http://www.jmir.org/2012/4/e105/
doi:10.2196/jmir.1953
PMID:22810046

Original Paper

Use of Email and Telephone Prompts to Increase Self-Monitoring in a Web-Based Intervention: Randomized Controlled Trial

Mary L Greaney[1], PhD; Kim Sprunck-Harrild[1], MSW, MPH; Gary G Bennett[2,3], PhD; Elaine Puleo[4], PhD; Jess Haines[5], PhD; K Vish Viswanath[1,6], PhD; Karen M Emmons[1,7], PhD

[1]Center for Community-Based Research, Dana-Farber Cancer Institute, Boston, MA, United States
[2]Department of Psychology and and Neuroscience, Duke University, Durham, NC, United States
[3]Duke Global Health Institute, Duke University, Durham, NC, United States
[4]Department of Public Health, University of Massachusetts Amherst, Amherst, MA, United States
[5]Department of Family Relations and Applied Nutrition, University of Guelph, Guelph, ON, Canada
[6]Department of Society, Human Development, and Health, Harvard School of Public Health, Boston, MA, United States
[7]Department of Society, Human Development & Health, Harvard School of Public Health, Boston, MA, United States

Corresponding Author:
Mary L Greaney, PhD
Center for Community-Based Research
Dana-Farber Cancer Institute
450 Brookline Ave, LW664
Boston, MA, 02215
United States
Phone: 1 617 582 7940
Fax: 1 617 582 5390
Email: mary_greaney@dfci.harvard.edu

Abstract

Background: Self-monitoring is a key behavior change mechanism associated with sustained health behavior change. Although Web-based interventions can offer user-friendly approaches for self-monitoring, engagement with these tools is suboptimal. Increased use could encourage, promote, and sustain behavior change.

Objective: To determine whether email prompts or email plus telephone prompts increase self-monitoring of behaviors on a website created for a multiple cancer risk reduction program.

Methods: We recruited and enrolled participants (N = 100) in a Web-based intervention during a primary care well visit at an urban primary care health center. The frequency of daily self-monitoring was tracked on the study website. Participants who tracked at least one behavior 3 or more times during week 1 were classified as meeting the tracking threshold and were assigned to the observation-only group (OO, n = 14). This group was followed but did not receive prompts. Participants who did not meet the threshold during week 1 were randomly assigned to one of 2 prompting conditions: automated assistance (AA, n = 36) or automated assistance + calls (AAC, n = 50). During prompting periods (weeks 2–3), participants in the AA and AAC conditions received daily automated emails that encouraged tracking and two tailored self-monitoring reports (end of week 2, end of week 3) that provided feedback on tracking frequency. Individuals in the AAC condition also received two technical assistance calls from trained study staff. Frequency of self-monitoring was tracked from week 2 through week 17.

Results: Self-monitoring rates increased in both intervention conditions during prompting and declined when prompting ceased. Over the 16 weeks of observation, there was a significant between-group difference in the percentage who met the self-monitoring threshold each week, with better maintenance in the AAC than in the AA condition ($P < .001$). Self-monitoring rates were greater in the OO group than in either the AA or AAC condition ($P < .001$).

Conclusions: Prompting can increase self-monitoring rates. The decrease in self-monitoring after the promoting period suggests that additional reminder prompts would be useful. The use of technical assistance calls appeared to have a greater effect in promoting self-monitoring at a therapeutic threshold than email reminders and the tailored self-monitoring reports alone.

Trial Registration: ClinicalTrials.gov NCT01415492; http://clinicaltrials.gov/ct2/show/NCT01415492 (Archived by WebCite at http://www.webcitation.org/68LOXOMe2)

(J Med Internet Res 2012;14(4):e96) doi:10.2196/jmir.1981

KEYWORDS

Web-based health promotion intervention; self-monitoring; prompts

Introduction

Self-monitoring of behaviors and health measures such as diet, smoking, physical activity, and weight is a key behavior change mechanism. Adherence to self-monitoring regimens is associated with greater behavior change [1-6], weight loss [5,7-10], and long-term maintenance of weight loss [11,12]. However, a major challenge has been initiating and maintaining use of self-monitoring tools. Prior to the widespread availability of personal computers and mobile devices, most self-monitoring was done with paper and pencil. Technology offers excellent potential for increasing the ease and engagement of self-monitoring.

Participants in Web-based interventions who self-monitored their weight or physical activity, or both, have been shown to have greater success than those who do not [5,13,14]. However, the use of self-monitoring tools has been suboptimal. For example, in a Web-based worksite behavior change intervention, only 11% of participants (41/378 participants) used the self-monitoring tools [15]. Limited use and attrition are concerns in Web-based intervention studies requiring participant enrollment [16,17], as many of these interventions are designed for multiple visits and consistent use. Previous research suggests that email or telephone prompts may promote website use [15,18-20].

Despite its importance, there is limited information about the frequency of self-monitoring in Web-based interventions. This may be because process data are seldom reported [14,21]. Nonetheless, the current literature indicates that self-monitoring rates in Web-based interventions are low [14,22]. An evaluation of the HealthPartners 10,000 Steps online program, which included six promotional activities, found that although 74% of participants tracked their behavior at least once, only 9% tracked their steps weekly throughout the 21-week intervention [23]. Notably higher rates of self-monitoring have been reported in interventions that included in-person meetings with study staff or group sessions, or both [5,18,24,25], but the costs associated with in-person meetings substantially reduce the possibility of the intervention being sustained. Thus, the purpose of this study was to examine the feasibility of implementing a prompting intervention in a Web-based health promotion intervention.

Methods

Healthy Directions 2

The prompting study (hereafter referred to as substudy) described in this paper was a substudy of Healthy Directions 2, a randomized controlled trial of a multiple risk factor cancer prevention intervention conducted in two urban primary care health centers located in metropolitan Boston, USA. In Healthy Directions 2, patients from 33 participating providers were recruited and randomly assigned to study arms at the provider level. Enrollment eligibility included being a health center patient, being 18+ years of age, having a scheduled well visit or chronic disease management appointment, and being able to read English. Patients were ineligible if they had undergone cancer treatment in the previous year or had a diagnosis of dementia, blindness, neurodegenerative disease, or psychiatric illness (including substance abuse, psychosis, or schizophrenia in the previous 5 years).

The Healthy Directions 2 intervention targeted multiple cancer risk factors and was designed to (1) promote physical activity, (2) reduce red meat intake, (3) increase fruit and vegetable consumption, (4) promote daily multivitamin use, and (5) promote smoking cessation, as applicable. Intervention components included an endorsement of behavior change by the participant’s health care provider; materials delivered through the study website or in print, based on participant preference; intervention materials for participants’ friends and family members; and links to community-based resources. The materials emphasized the importance of consistent and continued tracking of health behaviors. The website included a user-friendly section where patients could self-monitor all targeted behaviors at once. Although daily self-monitoring was encouraged, the website allowed participants to enter data for the day they logged into the website and for the 2 days prior. After entering data, participants received immediate feedback in the form of graphs and descriptive text. Participants could also view their data over time, to assess overall progress. Intervention materials were available via the Web or as a print packet. When joining the study, each participant randomly assigned to the intervention received a bottle of multivitamins, a pedometer, intervention materials or logon information for the study website, and a US $5 gift card. The Healthy Directions 2 study was approved by the institutional review board at Harvard Vanguard Medical Associates.

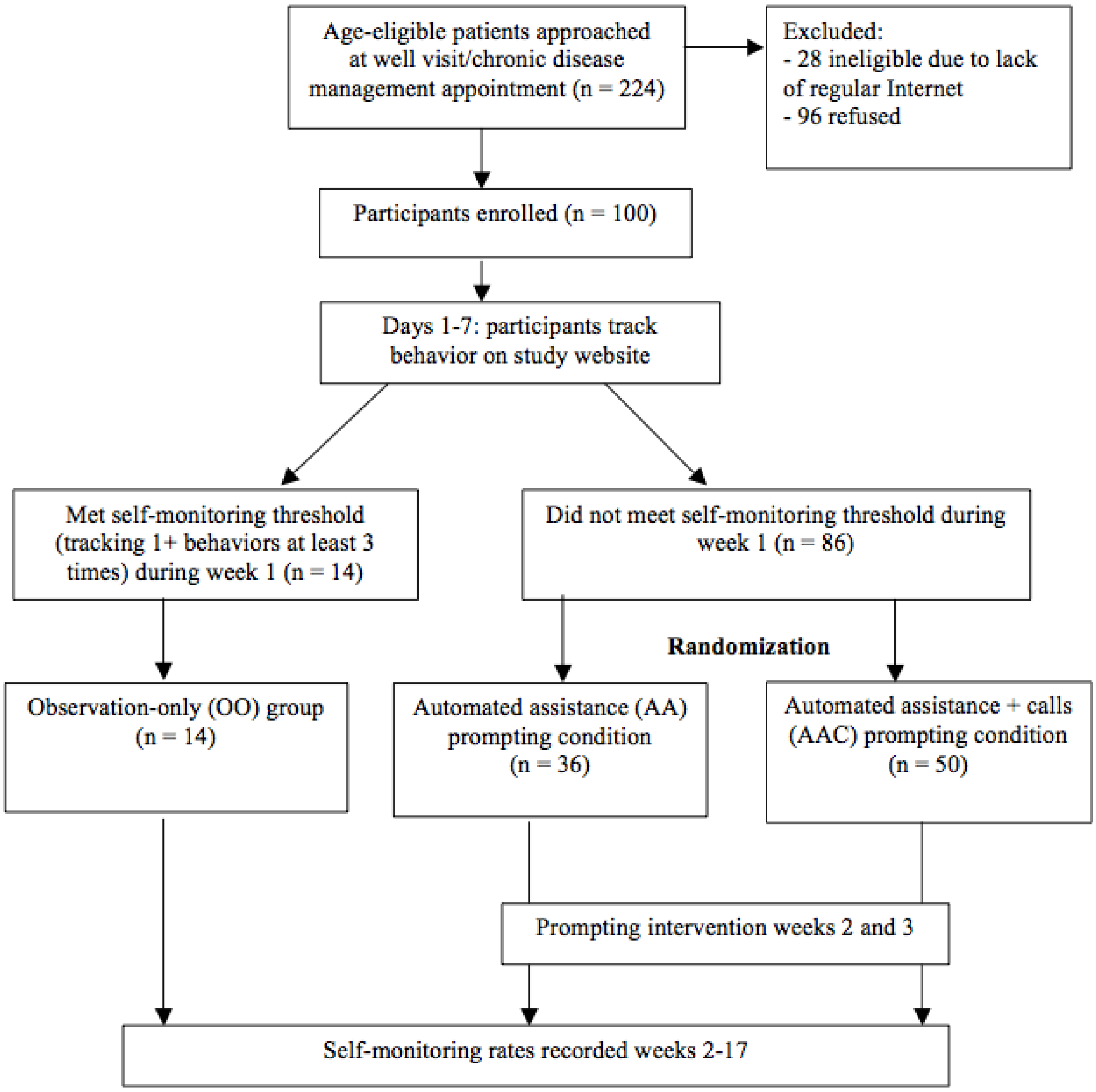
Recruitment

After completing recruitment for the parent study, we recruited an additional 100 participants for the substudy. This substudy was separate from the parent study. Eligibility requirements were the same as those for the parent study, plus the following additional criteria: (1) having an email address, (2) having the ability to access the Internet daily, and (3) being willing to receive the Healthy Directions 2 intervention via the Web only.

Recruitment for both the parent study and the substudy was the same. Eligible patients were sent an introductory letter that outlined the study and let them know that they may be approached and invited to join the study at their upcoming appointment. At check-in, study staff met the patient and verbally introduced the study, and interested individuals provided written informed consent and completed a self-administered baseline survey.

During recruitment for the substudy, eligible participants were made aware that they may receive an additional intervention that would include emails and possibly two technical assistance calls. After completing the survey, each substudy participant received a bottle of multivitamins, a pedometer, login information for the study website, and a US $5 gift card for completing the survey. Recruitment for the substudy was limited to one site (8 providers) and took place in March 2010. As Figure 1 shows, 224 people were approached to join the substudy, 28 were ineligible, and 100 enrolled (96 declined; response rate: 51%). The substudy was approved by the institutional review board at Harvard Vanguard Medical Associates.

Figure 1. Flow of participants through the substudy.



Prompting Conditions

We set a minimum threshold of self-monitoring at least one behavior 3 or more times per week [13]. Participants who met this threshold during week 1 did not receive prompts (observation-only group [OO]; n = 14) but were followed throughout the study as a comparison group with the two prompting conditions. Participants who did not meet the self-monitoring threshold during week 1 were randomly assigned, based on primary care physician, following the randomization scheme of the parent study to receive one of 2 prompting interventions: automated assistance (AA) or automated assistance + calls (AAC). Participants were actively prompted during weeks 2 and 3, and frequency of self-monitoring was tracked from week 2 through week 17.

Automated Assistance

Participants assigned to the AA condition (n = 36) received 2 weeks of daily emails during the prompting period (weeks 2 and 3) that encouraged them to track their behaviors via the study website. Email messages changed daily and included a brief message about the benefits of self-monitoring and a hyperlink to the study website. Participants could choose to respond directly to the email with their tracking information instead of logging into the website; study staff uploaded emailed self-monitoring data to the study website. Participants also received two tailored self-monitoring reports: the first at the end of week 2 and the second at the end of week 3. Reports provided feedback to the individual about his or her frequency of tracking for each of the behaviors during the previous week. The reports mirrored the self-monitoring graphs available on the website; we hoped that seeing the graphs would encourage individuals who had not self-monitored to visit the website and track their behaviors. Reports were viewed as part of the prompting intervention. If participants did not self-monitor, their reports reiterated the information on the benefits of self-monitoring that was included on the daily emails and encouraged self-monitoring via the study website.

Automated Assistance + Calls

Participants randomly assigned to the AAC condition (n = 50) were sent the emails and tailored self-monitoring reports, detailed above, and received two technical assistance calls. The first call was made at the end of the first week of prompting (week 2) and the second call took place at the end of the second week of prompting (week 3). The calls, conducted by a trained health coach, were designed to be brief (less than 5 minutes) and focused on troubleshooting technical questions (eg, trouble logging in or how to self-monitor on the website).

Measures

Tracking Measures

We created two categories of tracking measures. The first, self-monitoring measures, focused on the frequency of monitoring. We created the second, threshold measures, to examine the minimum weekly therapeutic threshold of self-monitoring 3 or more times per week. We used multiple measures to attempt to fully capture participants' interaction with monitoring.

Self-monitoring Measures

We determined the total number of self-monitoring events (range 0–112). We then determined the total number of weeks during which participants self-monitored at least once (range 0–16), the greatest number of continuous weeks in which participants self-monitored (range 0–16), and the frequency with which participants self-monitored each week (range 0–7).

Threshold Measures

Using the frequency of self-monitoring each week, we determined whether participants met the weekly threshold each week (yes or no), total number of weeks during which participants met the threshold (range 0–16), and the greatest number of continuous weeks in which participants met the threshold (range 0–16).

Demographic Measures

The baseline survey included items to assess race, Latino ethnicity, marital status, frequency of Internet use, comfort level in using computers, and participants' financial situation by asking participants to rate the "money situation" in their household (comfortable with extras, enough but no extras, have to cut back, and can't make ends meet). Age, sex, and primary care provider was determined by data obtained from participants' electronic medical records.

Behavioral Variables

All behaviors were assessed using validated measures [26-28]. The outcomes were dichotomized as to whether the participant met the recommendation for that specific behavior (75+ minutes of vigorous or 150+ minutes of moderate physical activity/week; 3 or fewer servings of red meat/week; 5+ servings of fruits and vegetables/day; a multivitamin 6–7 times/week; and not smoking).

Analysis

We obtained descriptive statistics for key variables (using SAS version 9.1; SAS Institute, Cary, NC, USA). We used chi-square statistics and analysis of variance to assess differences in demographics and in meeting the behavioral recommendations between those who met and those who did not meet the weekly self-monitoring threshold of tracking at least one behavior 3 or more times during week 1 (OO vs AA and AAC). We used analyses of variance to assess differences between the three groups (OO, AA, and AAC) and the self-monitoring variables and the threshold variables.

To examine the impact of the prompting conditions from baseline through the prompting period, we tested a series of 2 (group: prompting conditions) × 3 (time: weeks 1–3) repeated-measures models. Separate models were used for each dependent variable (eg, self-monitored each week, met weekly threshold of tracking at least one behavior 3 or more times/week, or frequency of self-monitoring each week). We used binomial repeated-measures models for the dichotomous outcomes and general liner models for the continuous outcomes. The models were specified with a within-group factor of time and a between-group factor of prompting condition. We conducted similar analyses to examine the impact of the prompting conditions over the 16 observational weeks. We then included the OO group in the analyses and conducted repeated-measures models with 3 (group: OO, AA, and AAC) × 16 (time: weeks 2–17) and post hoc tests to explore any differences between the groups.

Results

Participants

The sample (N = 100) was 53% (53/100) male, with a mean age of 45.6 years, and was racially and ethnically diverse, with 37% (37/100) of the sample being black and 8% (8/100) being Latino or Latina. Most participants were college graduates (70/100, 70%) and reported frequent Internet use (see Table 1). Demographic or behavioral variables did not differ between the OO group and the combined AA and AAC conditions. Among participants in the AAC condition, 80% (40/50 participants) completed call 1 and 70% (35/50 participants) completed call 2. In total, 6 participants required technical assistance (needed a user name or password, had emails going to their spam filter, or were not sure what to do).

Table 1. Demographic characteristics of the study sample (N = 100).

Characteristic	Data
Age (years)	
Mean (SD)	45.6 (14.9)
Range	21–84
Race/ethnicity, n (%) [a]	
White	52 (52%)
Hispanic	8 (8%)
Black	37 (37%)
Other	3 (3%)
Sex, n (%) [a]	
Male	53 (53%)
Female	47 (47%)
Education, n (%) [a]	
High school diploma or less	5 (5%)
Some college	24 (24%)
College graduate or more	70 (70%)
Married, n (%)	63 (63%)
Self-rated health status, n (%) [a]	
Fair/poor	7 (7%)
Good	44 (44%)
Excellent/very good	47 (47%)
Household financial situation, n (%) [a]	
Comfortable, with some extras	49 (49%)
Enough, but no extras	32 (32%)
Have to cut back	15 (15%)
Cannot make ends meet	4 (4%)
Frequency of Internet use (/week), n (%) [a]	
once	5 (5%)
2–4 times	14 (14%)
5+ times	80 (80%)
Comfort level using a computer, n (%) [a]	
Very uncomfortable/uncomfortable	7 (7%)
Comfortable/very comfortable	91 (91%)

[a] Due to missing values, percentages may not total 100%.

Self-Monitoring Rates

Overall, 99.92% of participants who self-monitored, regardless of prompting condition and frequency of self-monitoring, tracked all four (nonsmokers) or five (smokers) of the targeted behaviors each time that they self-monitored (1219 of 1220 times tracked). In addition, during 76.8% (205/267) of the weeks when participants self-monitored at least once, they met the weekly threshold for self-monitoring during that week (tracking 1 or more behaviors 3 or more times). In all conditions, self-monitoring rates decreased over time.

Self-Monitoring in the Observation-Only Group

The OO group consisted of study participants who met the weekly self-monitoring threshold (tracking 1 or more behaviors 3 or more times) during week 1 and therefore did not receive prompts. Within this group, some self-monitoring occurred in 100% of the observational weeks.

Figure 2 shows the percentage of participants within each group (OO, AA, and AAC) who self-monitored at least once during a week. As seen, the OO group had a greater percentage of people self-monitoring at least once per week than either the AA or AAC group. However, self-monitoring rates for all groups declined over time.

Figure 3 shows the percentage of participants within each group (OO, AA, and AAC) who met the self-monitoring threshold during a week. As shown, between 21% (3/14) and 86% (12/14) of participants in the OO group met the threshold each week. The OO group had a greater percentage of participants who met the self-monitoring threshold each week than either the AA or AAC group. Self-monitoring in the OO group dropped most precipitously during the first 6 weeks; rates then stabilized, with around 30%–40% of this group continuing to track in the remaining weeks and a slightly smaller percentage meeting the weekly threshold.

Figure 2. Percentage of participants, by group, self-monitoring 1 or more times per week (N = 100). The prompting intervention occurred during weeks 2 and 3. AA = automated assistance group, AAC = automated assistance + calls group, OO = observation-only group.

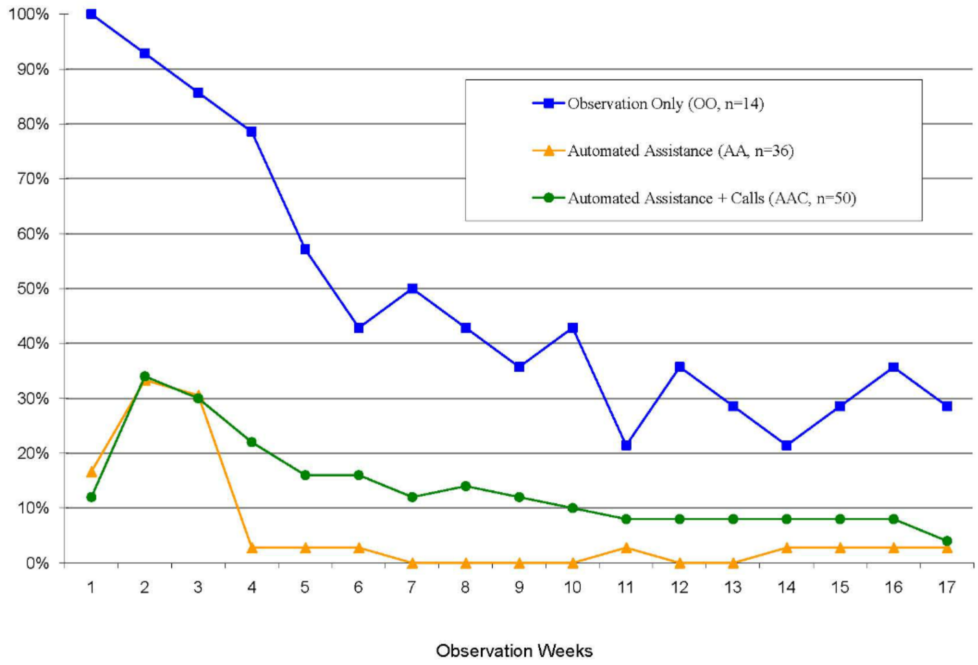
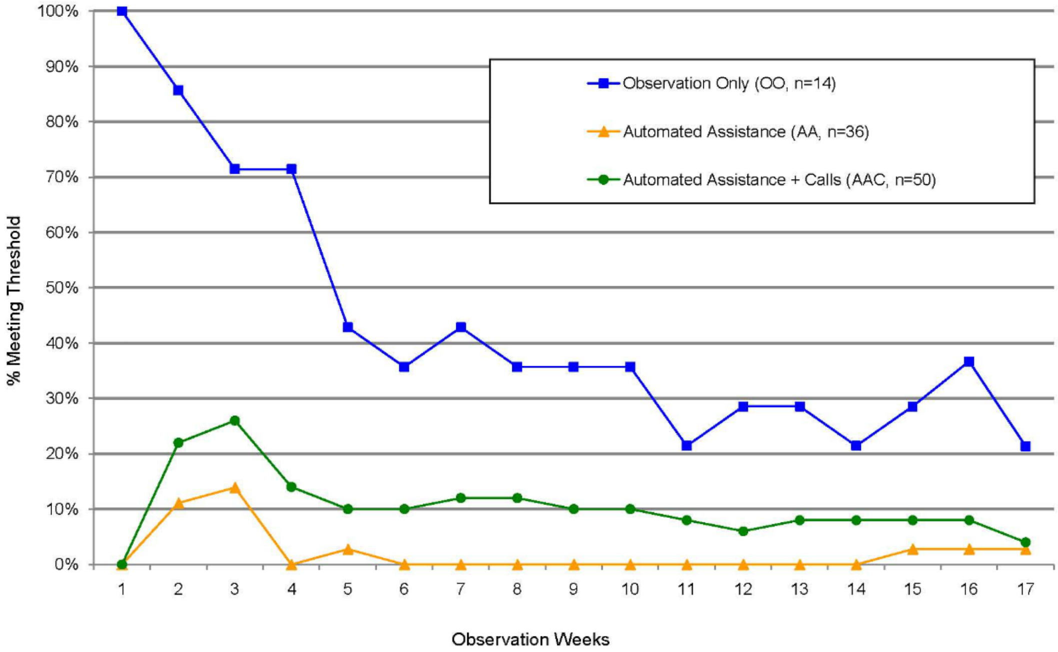


Figure 3. Percentage of participants, by group, meeting the self-monitoring threshold by week (N = 100). The prompting intervention occurred during weeks 2 and 3. The weekly threshold for self-monitoring was defined as tracking 1 or more behaviors 3 or more times during a week. AA = automated assistance group, AAC = automated assistance + calls group, OO = observation-only group.



Impact of the Prompting Interventions

During the Prompting Period

Time had a within-group effect on whether self-monitoring occurred each week (*P* < .001), on whether the threshold was met each week (*P* = .006), and on frequency of self-monitoring each week (*P* < .001). The between-group differences and group-by-time interactions were not significant.

Over the Course of the Observation Period

The repeated-measures models examining whether self-monitoring occurred each week revealed a within-group effect for time (*P* < .001) but not a between-group effect. The analyses examining whether the threshold was met each week revealed a within-group effect (*P* < .001) and a between-group effect (*P* = .009). The AA condition was less likely than the AAC condition to meet the threshold each week (odds ratio 0.27, 95% confidence interval 0.09–0.72). The group-by-time interaction was not significant. The analyses examining frequency of self-monitoring each week determined that there was a within-group effect (*P* < .001) but not a between-group effect.

In the AAC condition, there were more weeks when people self-monitored their behaviors at least once than in the AA condition (16/16, 100% vs 10/16, 63% weeks, Figure 2). Similarly, in the AAC condition at least one participant met the weekly threshold during a greater number of weeks than in the AA condition (16/16, 100% vs 6/16, 31% weeks, Figure 3). In addition, there was a significant difference between the AA and AAC conditions in the mean number of weeks during which participants met the threshold (AAC: mean 1.8, SD 3.9 weeks; AA: mean 0.4, SD 0.8 weeks; *P* = .04). There were, however, no significant differences between the two prompting conditions in the self-monitoring variables (frequency of self-monitoring, total number of weeks self-monitored, and greatest number of continuous weeks self-monitored) or in the greatest number of continuous weeks when the participants met the threshold.

Differences Between the OO and the Prompting Conditions

The results of the repeated-measures models and post hoc analyses examining the three self-monitoring outcomes determined that the OO group was significantly different from both of the prompting conditions on all outcomes (*P* < .001), with results favoring the OO group.

Response to Email Prompts

Only a small percentage of participants (9/86, 10%) replied to the tracking reminder emails with their tracking information, but of those who responded, 67% (6/9) replied multiple times.

Discussion

Consistent and continued self-monitoring is an instrumental strategy for initiating and maintaining behavior change [1-5,7-12], and technology provides a platform that enables participants to receive immediate, tailored feedback. Data show that 78.3% of the US population have access to the Internet [29], and that the use of Web-based interventions is increasing [17]. The Internet and Web-based interventions provide the opportunity to develop and implement behavior change interventions that actively promote self-monitoring. Nonetheless, self-monitoring rates remain low, and the potential for Web-based self-monitoring has not yet been realized, particularly in interventions that do not include in-person interaction with study staff (eg, health coaches). Providing assistance or prompts to encourage self-monitoring could be a key strategy leading to consistent and sustained self-monitoring

[4,18]. Testing and developing sustainable strategies to initiate self-monitoring and to support consistent tracking is warranted.

The purpose of this study was to examine the effect of two limited prompting interventions on increasing participants' use of self-monitoring tools available on the study website. In this study, participants who did not meet the self-monitoring threshold of tracking at least one behavior 3 or more times during the first week of the study were enrolled in one of two prompting conditions. Both prompting conditions yielded a modest but significant increase in self-monitoring each week (yes vs no), meeting the weekly threshold (yes vs no), and frequency of self-monitoring/week during the prompting period. After prompting ceased, self-monitoring rates decreased in both prompting conditions. Although the significant within-group effect remained throughout the observation period, the increased self-monitoring during the prompting period coupled with the subsequent decline suggest that reminder emails can increase prompting, but that additional subsequent email reminders may be useful to help sustain this increase. Future research is needed to examine the effectiveness for different prompting intervals [30].

Over the course of the study, a greater percentage of participants in AAC, the group that received technical assistance calls, met the weekly self-monitoring threshold, suggesting that brief contact, even in the context of technical assistance calls, may be beneficial for promoting and sustaining tracking. This limited contact with study staff may have increased motivation to track by removing technical barriers, or simply served as a gentle push to self-monitor. In total, 80% of participants in the AAC condition completed call 1, and 70% completed call 2; higher completion rates may have resulted in the AAC prompting intervention having a greater impact on self-monitoring. Future research is needed to explore the nature of interactions with staff that prompt higher levels of engagement (modality, frequency, content, etc), and whether a more explicit focus on enhancing motivation to track may yield greater self-monitoring.

Over the course of the study, the OO group had the highest rates of self-monitoring each week. However, even with this group's strong start, their self-monitoring during weeks 3–5 declined precipitously, suggesting that this may be a critical time to intervene. Beginning with week 4, when prompting ended, self-monitoring rates decreased in both the AA and AAC groups. This general across-the-board attrition is similar to those seen in other studies [31,32] and is of concern.

It is noteworthy that participants who self-monitored chose to track multiple behaviors simultaneously 99.9% of the time. This is encouraging and suggests that tracking multiple behaviors is not burdensome. The willingness to track multiple behaviors in this study may have been due, in part, to the beta testing of the website that was conducted to inform a user-friendly design. The website was specifically structured so that all tracking could be completed on one page, thus making it easy to track all behaviors at the same time. In addition, it is striking that in 76.8% (205/267) of the weeks when participants self-monitored they met the threshold for that week, suggesting that specific strategies could be helpful for individuals who are willing to self-monitor but do not reach a weekly therapeutic threshold.

The response rate to the tracking emails was low. It is possible that participants who were motivated to track simply logged in to the website and recorded their self-monitoring data themselves without assistance. It also could be that participants found it difficult to respond to the emails because of the timing of the sent messages (ie, unable to respond to the email during the workday). One Web-based study found that participants who actively responded to emails had a greater increase in fruit and vegetable intake than did participants who did not reply to emails [20].

Most often tracking has been limited to pencil and paper. Burke and colleagues reported that participants found self-monitoring via a personal digital assistant to be more socially acceptable than monitoring in paper logs [30]. The Healthy Directions 2 website was designed to be accessed via a computer, and if accessed via a smart phone, the user was required to login and would need to use the zoom capacity to clearly read text and enter self-monitoring data. Future Web-based interventions may want to consider ways that smart phones and other electronic devices can be used to allow participants to easily enter self-monitoring data that are uploaded to the study website.

Study limitations include a modest sample size; a largely well-educated sample, which may limit generalizability; and the lack of a nonprompting control group that included only participants who did not self-monitor during the first week of the study. Study strengths include having a diverse racial and ethnic sample and a design that allowed comparison with a group already motivated to record their behavior.

It would be helpful if interventions that include self-monitoring provided information about these rates or used consistent definitions of self-monitoring. But, to date, the literature has not provided either, making it difficult to compare adherence across studies [33]. In addition, research is needed to understand what external factors or personal attributes and characteristics motivate initial self-monitoring without prompting. Better understanding and use of these elements in interventions could be used to improve participation by less self-motivated populations. A 2009 literature review on the use of prompts in behavior change interventions concluded that tailoring of periodic prompts through regular contact with a counselor was associated with positive behavior change [29]. It may be that prompting interventions, including the prompting strategy, the frequency of contact, and the content of the message, should be individually tailored based on individual characteristics and behaviors that are assessed at baseline and throughout the intervention period. This will be an important area for future research.

The 2010 US Affordable Care Act emphasizes prevention and engaging patients as active participants in their health care [34,35]. Encouraging self-monitoring could play an important role in this effort, and comparing different prompting intervention to determine the most effective intervention is important. The use of technology-based tools that include feedback on progress and goals could make self-monitoring more accessible. However, little information is available about factors associated with self-monitoring, particularly self-monitoring conducted via personal computers and mobile devises, and less is known about best practices of using prompts to increase self-monitoring. Research is needed to understand what mechanisms can be used to increase use of self-monitoring tools to therapeutic levels. Different strategies may be needed to help individuals initiate and develop a strong foundation of self-monitoring and to maintain their motivation to continue tracking over long periods of time.

Acknowledgments

We thank Harvard Vanguard Medical Associates for the opportunity to conduct our research in their clinical sites. This work was supported by NIH grants R01 CA123228 and 1K05 CA124415. We also would like to thank the study participants and the Healthy Directions 2 staff: Louisa Bloomstein, Joanne Chin, Lindsay ElShazly, Kristie Usher, and Val Varner.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Consort EHealth Checklist V1.6 [36].

[PDF File (Adobe PDF File), 2MB - jmir_v14i4e96_app1.pdf]

References

1. Bowen D, Raczynski J, George V, Feng Z, Fouad M. The role of participation in the women's health trial: feasibility study in minority populations. Prev Med 2000 Nov;31(5):474-480. [doi: 10.1006/pmed.2000.0731] [Medline: 11071827]
2. Tinker LF, Rosal MC, Young AF, Perri MG, Patterson RE, Van Horn L, et al. Predictors of dietary change and maintenance in the Women's Health Initiative Dietary Modification Trial. J Am Diet Assoc 2007 Jul;107(7):1155-1166. [doi: 10.1016/j.jada.2007.04.010] [Medline: 17604744]
3. Aittasalo M, Miilunpalo S, Kukkonen-Harjula K, Pasanen M. A randomized intervention of physical activity promotion and patient self-monitoring in primary health care. Prev Med 2006 Jan;42(1):40-46. [doi: 10.1016/j.ypmed.2005.10.003] [Medline: 16297442]
4. Ferrier S, Blanchard CM, Vallis M, Giacomantonio N. Behavioural interventions to increase the physical activity of cardiac patients: a review. Eur J Cardiovasc Prev Rehabil 2011 Feb;18(1):15-32. [doi: 10.1097/HJR.0b013e32833ace0e] [Medline: 20502341]
5. Conroy MB, Yang K, Elci OU, Gabriel KP, Styn MA, Wang J, et al. Physical activity self-monitoring and weight loss: 6-month results of the SMART trial. Med Sci Sports Exerc 2011 Aug;43(8):1568-1574. [doi: 10.1249/MSS.0b013e31820b9395] [Medline: 21200337]
6. Carels RA, Darby LA, Rydin S, Douglass OM, Cacciapaglia HM, O'Brien WH. The relationship between self-monitoring, outcome expectancies, difficulties with eating and exercise, and physical activity and weight loss treatment outcomes. Ann Behav Med 2005 Dec;30(3):182-190. [doi: 10.1207/s15324796abm3003_2] [Medline: 16336069]
7. Boutelle KN, Kirschenbaum DS. Further support for consistent self-monitoring as a vital component of successful weight control. Obes Res 1998 May;6(3):219-224. [Medline: 9618126]
8. Burke LE, Conroy MB, Sereika SM, Elci OU, Styn MA, Acharya SD, et al. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. Obesity (Silver Spring) 2011 Feb;19(2):338-344 [FREE Full text] [doi: 10.1038/oby.2010.208] [Medline: 20847736]
9. Wing RR, Hamman RF, Bray GA, Delahanty L, Edelstein SL, Hill JO, Diabetes Prevention Program Research Group. Achieving weight and activity goals among diabetes prevention program lifestyle participants. Obes Res 2004 Sep;12(9):1426-1434. [doi: 10.1038/oby.2004.179] [Medline: 15483207]
10. VanWormer JJ, Martinez AM, Martinson BC, Crain AL, Benson GA, Cosentino DL, et al. Self-weighing promotes weight loss for obese adults. Am J Prev Med 2009 Jan;36(1):70-73. [doi: 10.1016/j.amepre.2008.09.022] [Medline: 18976879]
11. Butryn ML, Phelan S, Hill JO, Wing RR. Consistent self-monitoring of weight: a key component of successful weight loss maintenance. Obesity (Silver Spring) 2007 Dec;15(12):3091-3096 [FREE Full text] [doi: 10.1038/oby.2007.368] [Medline: 18198319]
12. Vanwormer JJ, French SA, Pereira MA, Welsh EM. The impact of regular self-weighing on weight management: a systematic literature review. Int J Behav Nutr Phys Act 2008;5:54 [FREE Full text] [doi: 10.1186/1479-5868-5-54] [Medline: 18983667]
13. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
14. Morgan PJ, Lubans DR, Collins CE, Warren JM, Callister R. 12-month outcomes and process evaluation of the SHED-IT RCT: an internet-based weight loss program targeting men. Obesity (Silver Spring) 2011 Jan;19(1):142-151 [FREE Full text] [doi: 10.1038/oby.2010.119] [Medline: 20523304]

15. Robroek SJ, Brouwer W, Lindeboom D, Oenema A, Burdorf A. Demographic, behavioral, and psychosocial correlates of using the website component of a worksite physical activity and healthy nutrition promotion program: a longitudinal study. J Med Internet Res 2010;12(3):e44 [FREE Full text] [doi: 10.2196/jmir.1402] [Medline: 20921001]
16. 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]
17. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
18. Stevens VJ, Funk KL, Brantley PJ, Erlinger TP, Myers VH, Champagne CM, et al. Design and implementation of an interactive website to support long-term maintenance of weight loss. J Med Internet Res 2008;10(1):e1 [FREE Full text] [doi: 10.2196/jmir.931] [Medline: 18244892]
19. 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]
20. Woodall WG, Buller DB, Saba L, Zimmerman D, Waters E, Hines JM, et al. Effect of emailed messages on return use of a nutrition education website and subsequent changes in dietary behavior. J Med Internet Res 2007;9(3):e27 [FREE Full text] [doi: 10.2196/jmir.9.3.e27] [Medline: 17942389]
21. Oakley CB, Bomba AK, Knight KB, Byrd SH. Evaluation of menus planned in Mississippi child-care centers participating in the Child and Adult Care Food Program. J Am Diet Assoc 1995 Jul;95(7):765-768. [doi: 10.1016/S0002-8223(95)00213-8] [Medline: 7797806]
22. Harvey-Berino J, Pintauro S, Buzzell P, Gold EC. Effect of internet support on the long-term maintenance of weight loss. Obes Res 2004 Feb;12(2):320-329. [doi: 10.1038/oby.2004.40] [Medline: 14981225]
23. Speck RM, Hill RK, Pronk NP, Becker MP, Schmitz KH. Assessment and outcomes of HealthPartners 10,000 Steps program in an academic work site. Health Promot Pract 2010 Sep;11(5):741-750. [doi: 10.1177/1524839908330745] [Medline: 19182264]
24. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res 2010;12(3):e29 [FREE Full text] [doi: 10.2196/jmir.1504] [Medline: 20663751]
25. Webber KH, Tate DF, Michael Bowling J. A randomized comparison of two motivationally enhanced Internet behavioral weight loss programs. Behav Res Ther 2008 Sep;46(9):1090-1095. [doi: 10.1016/j.brat.2008.06.008] [Medline: 18675402]
26. Estabrooks PA, Bradshaw M, Dzewaltowski DA, Smith-Ray RL. Determining the impact of Walk Kansas: applying a team-building approach to community physical activity promotion. Ann Behav Med 2008 Aug;36(1):1-12. [doi: 10.1007/s12160-008-9040-0] [Medline: 18607666]
27. Subar AF, Heimendinger J, Patterson BH, Krebs-Smith SM, Pivonka E, Kessler R. Fruit and vegetable intake in the United States: the baseline survey of the Five A Day for Better Health Program. Am J Health Promot 1995;9(5):352-360. [Medline: 10150767]
28. Emmons KM, Stoddard AM, Fletcher R, Gutheil C, Suarez EG, Lobb R, et al. Cancer prevention among working class, multiethnic adults: results of the healthy directions-health centers study. Am J Public Health 2005 Jul;95(7):1200-1205. [doi: 10.2105/AJPH.2004.038695] [Medline: 15933240]
29. Internet World Stats. Miniwatts Marketing Group. 2012 Feb 08. United States of America URL: http://www.internetworldstats.com/america.htm [accessed 2012-06-20] [WebCite Cache ID 68ZHSWSiw]
30. Burke LE, Styn MA, Glanz K, Ewing LJ, Elci OU, Conroy MB, et al. SMART trial: A randomized clinical trial of self-monitoring in behavioral weight management-design and baseline findings. Contemp Clin Trials 2009 Nov;30(6):540-551. [doi: 10.1016/j.cct.2009.07.003] [Medline: 19665588]
31. Carels RA, Young KM, Coit C, Clayton AM, Spencer A, Hobbs M. Can following the caloric restriction recommendations from the Dietary Guidelines for Americans help individuals lose weight? Eat Behav 2008 Aug;9(3):328-335. [doi: 10.1016/j.eatbeh.2007.12.003] [Medline: 18549992]
32. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J Med 2006 Oct 12;355(15):1563-1571 [FREE Full text] [doi: 10.1056/NEJMoa061883] [Medline: 17035649]
33. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102. [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
34. Honoré PA, Wright D, Berwick DM, Clancy CM, Lee P, Nowinski J, et al. Creating a framework for getting quality into the public health system. Health Aff (Millwood) 2011 Apr;30(4):737-745. [doi: 10.1377/hlthaff.2011.0129] [Medline: 21471496]
35. Koh HK, Sebelius KG. Promoting prevention through the Affordable Care Act. N Engl J Med 2010 Sep 30;363(14):1296-1299 [FREE Full text] [doi: 10.1056/NEJMp1008560] [Medline: 20879876]
36. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AA: automated assistance group
AAC: automated assistance + calls group
OO: observation-only group

Edited by G Eysenbach; submitted 03.11.11; peer-reviewed by J Whiteley; comments to author 15.01.12; revised version received 30.01.12; accepted 29.04.12; published 27.07.12.

Please cite as:
Greaney ML, Sprunck-Harrild K, Bennett GG, Puleo E, Haines J, Viswanath KV, Emmons KM
Use of Email and Telephone Prompts to Increase Self-Monitoring in a Web-Based Intervention: Randomized Controlled Trial
J Med Internet Res 2012;14(4):e96
URL: http://www.jmir.org/2012/4/e96/
doi:10.2196/jmir.1981
PMID:22842775

Original Paper

Preventing Smoking Relapse via Web-Based Computer-Tailored Feedback: A Randomized Controlled Trial

Iman Elfeddali[1]; Catherine Bolman[2], PhD; Math J.J.M Candel[3], PhD; Reinout W Wiers[4], PhD; Hein de Vries[1]

[1]Department of Health Promotion, School for Public Health and Primary Care (Caphri), Maastricht University, Maastricht, Netherlands
[2]Department of Psychology, Open University of the Netherlands, Heerlen, Netherlands
[3]Department of Methodology and Statistics, School for Public Health and Primary Care (Caphri), Maastricht University, Maastricht, Netherlands
[4]Department of Developmental Psychology, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Iman Elfeddali
Department of Health Promotion
School for Public Health and Primary Care (Caphri)
Maastricht University
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882436
Fax: 31 43 36 71 032
Email: iman.elfeddali@maastrichtuniversity.nl

Abstract

Background: Web-based computer-tailored approaches have the potential to be successful in supporting smoking cessation. However, the potential effects of such approaches for relapse prevention and the value of incorporating action planning strategies to effectively prevent smoking relapse have not been fully explored. The Stay Quit for You (SQ4U) study compared two Web-based computer-tailored smoking relapse prevention programs with different types of planning strategies versus a control group.

Objectives: To assess the efficacy of two Web-based computer-tailored programs in preventing smoking relapse compared with a control group. The action planning (AP) program provided tailored feedback at baseline and invited respondents to do 6 preparatory and coping planning assignments (the first 3 assignments prior to quit date and the final 3 assignments after quit date). The action planning plus (AP+) program was an extended version of the AP program that also provided tailored feedback at 11 time points after the quit attempt. Respondents in the control group only filled out questionnaires. The study also assessed possible dose–response relationships between abstinence and adherence to the programs.

Methods: The study was a randomized controlled trial with three conditions: the control group, the AP program, and the AP+ program. Respondents were daily smokers (N = 2031), aged 18 to 65 years, who were motivated and willing to quit smoking within 1 month. The primary outcome was self-reported continued abstinence 12 months after baseline. Logistic regression analyses were conducted using three samples: (1) all respondents as randomly assigned, (2) a modified sample that excluded respondents who did not make a quit attempt in conformance with the program protocol, and (3) a minimum dose sample that also excluded respondents who did not adhere to at least one of the intervention elements. Observed case analyses and conservative analyses were conducted.

Results: In the observed case analysis of the randomized sample, abstinence rates were 22% (45/202) in the control group versus 33% (63/190) in the AP program and 31% (53/174) in the AP+ program. The AP program (odds ratio 1.95, *P* = .005) and the AP+ program (odds ratio 1.61, *P* = .049) were significantly more effective than the control condition. Abstinence rates and effects differed per sample. Finally, the results suggest a dose–response relationship between abstinence and the number of program elements completed by the respondents.

Conclusion: Despite the differences in results caused by the variation in our analysis approaches, we can conclude that Web-based computer-tailored programs combined with planning strategy assignments and feedback after the quit attempt can be effective in preventing relapse 12 months after baseline. However, adherence to the intervention seems critical for effectiveness. Finally, our results also suggest that more research is needed to assess the optimum intervention dose.

Trial Registration: Dutch Trial Register: NTR1892; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1892 (Archived by WebCite at http://www.webcitation.org/693S6uuPM)

(J Med Internet Res 2012;14(4):e109) doi:10.2196/jmir.2057

KEYWORDS

Smoking relapse prevention; computer tailoring; multiple tailoring; planning strategies

Introduction

Smoking relapse rates can be extremely high (up to 90% in the first 3 months), and only 3%-5% of quitters maintain their quit attempt for 6 months or longer [1]. The role of risk factors for relapse (eg, low self-efficacy, the expectation of negative outcomes from quitting, negative affect, stress, and physical dependence) is quite well documented (see for examples [2-6]). However, a Cochrane review concluded that current smoking relapse prevention programs are not effective [7]. This ineffectiveness, combined with the alarming relapse rates, underlines the need for new, effective smoking relapse prevention strategies and programs. One potential explanation for these programs' lack of effectiveness is that quitters are not fully prepared for their cessation attempt and lack sufficient coping strategies to maintain their attempts successfully [8-10].

Adding action planning components to programs may be a promising strategy for improving smoking relapse prevention programs [8-12]. Planning strategies are already incorporated into many face-to-face and telephone smoking cessation counseling sessions (eg, smoking cessation courses provided by the Dutch Foundation for a Smoke-free Future in the Netherlands) [13]. Furthermore, the role of action planning strategies is acknowledged by integrative social cognitive models, such as the Integrated Change (I-Change) Model [14,15]. The I-Change Model distinguishes between two essential planning components: (1) preparatory plans as actions designed to prepare for the behavior change, and (2) coping plans as actions designed to maintain the new behavior by coping with challenging or difficult situations. Yet, as far as we know, these planning strategies have received little emphasis in Web-based prevention programs.

The Internet has proven to be a promising delivery mode for health-promoting and lifestyle-changing interventions (for instance, [12,16-18]). Since 91% of Dutch households have access to the Internet, it could potentially be used to reach large numbers of people [19]. Additionally, interactive, personalized Internet-based delivery modes, such as computer tailoring, have already demonstrated their potential to support smoking cessation [12,20-23]. These highly personalized approaches are assessment based and adapt their messages to the needs of the respondent [24,25]. Compared with nontailored messages, tailored messages are evaluated more positively, attract more attention, and are more likely to be read [12,16,24,26]. Computer-tailored health programs have shown to be promising tools for promoting healthy behavior in general [27-32] and smoking cessation specifically [12,20-23], with multiple tailoring moments being more effective than a single tailoring moment [32-34]. Few computer-tailored programs have explicitly targeted smoking relapse (see Borland et al [35] for an example). In line with earlier studies [31,33,34], Borland's study also indicated the surplus value of multiple tailoring moments and suggested a dose-response relationship between the number of feedback letters and smoking abstinence. Moreover, no studies have assessed the effects of using planning strategies in combination with multiple tailored feedback time points after the quit attempt.

In sum, the main goal of the Stay Quit for You (SQ4U) study was to assess the efficacy of two relapse prevention programs: (1) an action planning (AP) program that provided tailored feedback based on the baseline questionnaire and 6 preparatory and coping planning assignments, and (2) an action planning plus (AP+) program that extended the AP program by providing tailored feedback at 11 time points after the quit date. The efficacy of the programs was compared with that of a control group (with no intervention). Moreover, we aimed to assess possible dose–response relationships between abstinence and adherence to the number of program elements. First, we expected both programs to be more effective than the control condition in fostering continued abstinence 12 months after the start of the study (hypothesis 1). We expected the AP+ program to be the most effective. Moreover, we expected to find a dose–response relationship between continued abstinence and intervention dose (hypothesis 2). Finally, we will provide an overview of the respondents' program evaluations.

Methods

Ethics approval was obtained from the Medical Ethics Committee of Maastricht Academic Hospital and Maastricht University (MEC 08-3-003; NL21414.068.08). The study is registered with the Dutch Trial Register (NTR1892).

Respondents and Recruitment

We recruited smokers by placing ads in local newspapers, distributing 10,000 flyers in the city of Maastricht, and placing online ads on the websites of national health funds, a national news page, and the Dutch Foundation for a Smoke-free Future. The ads referred the respondents to our research website for more information. All data were gathered via the Web and there was no face-to-face contact between respondents and the study team. A software program randomly assigned a total of 2681 respondents to one of the three conditions using a simple randomization type (see design below). The enrollment and inclusion of respondents is presented in Figure 1. Respondents were eligible for participation when they met the baseline inclusion criteria (aged 18–65 years, smoked daily, willing to set a quit date within 1 month, and motivated to quit smoking) and agreed with the informed consent. The final sample consisted of 2031 respondents, of whom 566 (27.98%) responded to the 12-month follow-up measurement.

Figure 1. Enrollment and inclusion of respondents. NA = not applicable. [a] All respondents as randomly assigned, [b] modified sample excluding respondents who indicated after their quit date as well as at follow-up that they did not make a quit attempt during the study and respondents who quit too close to follow-up (see Analyses section in the Methods), [c] minimum dose sample additionally excluding respondents who did not adhere to at least one intervention element.

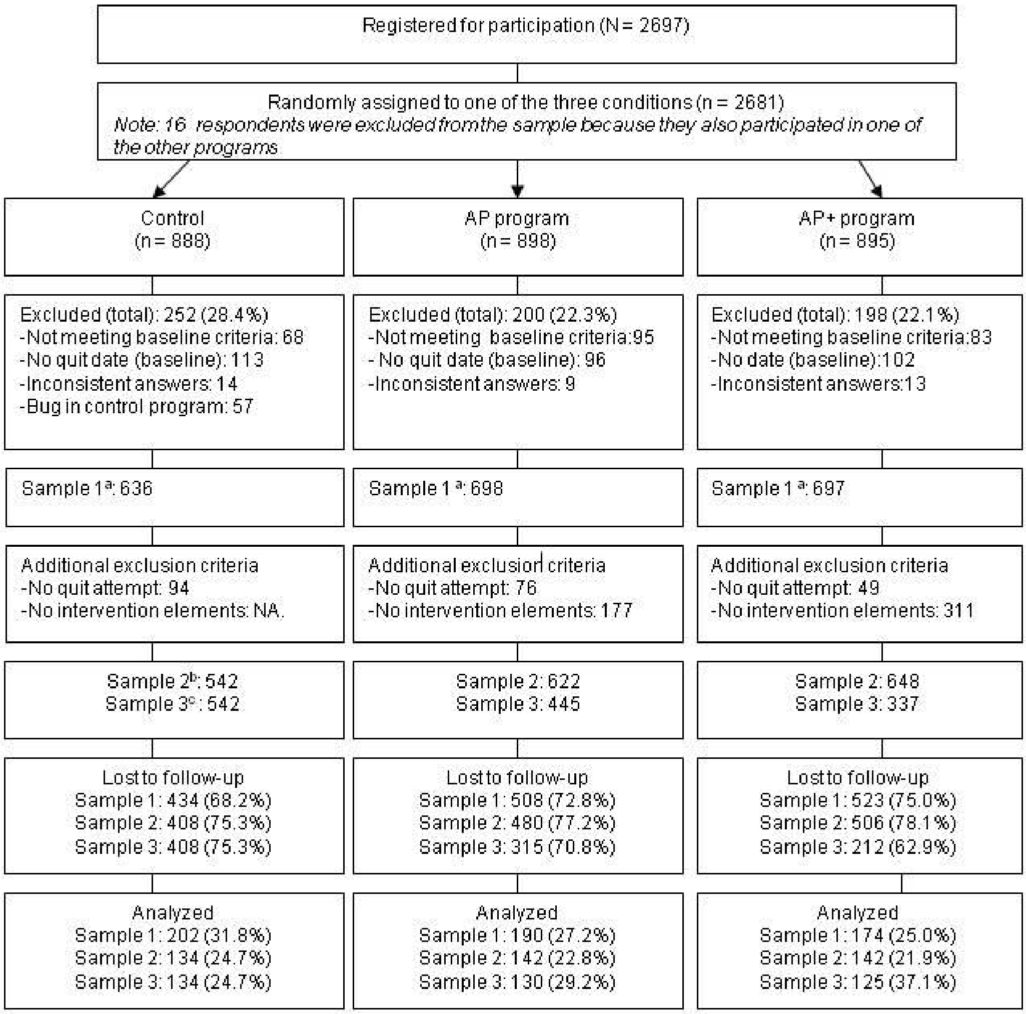


Figure 1: Enrollment and inclusion of respondents. NA = not applicable. [a] All respondents as randomly assigned, [b] modified sample excluding respondents who indicated after their quit-date as well as at follow-up that they did not make a quit-attempt during the study and respondents who quitted too close to follow-up (see Analyses section in the Methods), [c] minimum dose sample additionally excluding respondents who did not adhere to at least one intervention element.

Design and Procedure

The SQ4U study was a randomized controlled study with a control condition and two experimental conditions. Respondents registered via the research website and made their own login account (each time they were invited for participation they received this account in the invitation mail). After registration, respondents were randomly assigned to one of the three programs outlined in the Introduction. Respondents in the intervention groups were invited by email or text message (optional) to perform intervention tasks (eg, filling out a planning assignment). The same procedure was used to invite all respondents for the 12-month follow-up measurement. Reminder emails were sent when respondents did not respond to the follow-up measurement, which asked them to report on their smoking behavior using self-assessed questionnaires. Respondents who reported that they were abstinent were asked to take biochemical validation tests (see below). Respondents who completed all parts of their assigned SQ4U variant (including those in the control group) were eligible to win 1 of 20 prizes of €250. A more detailed description of the development, design, study course, biochemical validation, and intervention elements can be found elsewhere [36].

Biochemical Validation

We used cotinine assessments to validate reports of abstinence in a subsample (n = 70) of the respondents. This verified the nonsmoking status self-reported by respondents at the 12-month follow-up measurement. Only 1 (1%) respondent showed positive cotinine results and admitted to having smoked. Another respondent admitted to smoking when invited for biochemical validation. The smoking status for these 2 (3%) respondents was changed to relapse.

Intervention Materials

Baseline Computer-Tailored Feedback Before the Quit Attempt (AP and AP+ programs)

Feedback focused on respondents' perceptions of smoking and quitting and was based on previously tested effective computer tailoring programs that used the I-Change Model [12,20]. The feedback was intended to increase the respondents' motivation to quit smoking and to prepare them for the period after the quit attempt. It targeted the pros and cons of not smoking and relapses, provided tips to enhance or increase self-efficacy, offered suggestions for dealing with social influences, explained the importance of preparatory planning, and gave ideas for preparing for the quit attempt and dealing with stress.

Planning Strategy Assignments Before and After the Quit Attempt (AP and AP+ programs)

Since planning assignments have proven to be promising strategies to foster cessation and prevent relapse, the SQ4U programs incorporated assignments designed to motivate respondents to use these strategies to make a quit attempt (first 3 assignments, prior to quit date) and to prevent relapse (final 3 assignments, after quit date). Two assignments focused on making and carrying out preparatory plans; 2 more assignments addressed coping planning. Finally, 1 assignment focused on a specific coping plan (making an emergency plan for coping in case of a lapse) and 1 assignment required making a specific preparatory plan (planning a reward for abstaining from smoking for some time).

Tailored Feedback After the Quit Attempt (AP+ program only)

Studies that use ecological momentary assessments to gain in-depth, day-to-day information on the process of quitting have noted that low self-efficacy and negative affect preceded lapsing [37-40]. Therefore, we targeted self-efficacy, recovery self-efficacy, and negative affect in the feedback moments after the quit attempt. We also included feedback about the level of planning. Each measurement after the quit attempt targeted two of the four concepts mentioned. Tailored feedback about each of the four concepts was provided daily during the first 3 days (after the quit date), once a week during the rest of the first month, and once every 2 weeks until the third month. The feedback not only addressed the respondents' present state but was also ipsative, referring to earlier states (reporting changes in a positive or negative direction). Finally, all feedback targeted self-reported smoking behavior for the last three measurements and reported the progress of maintenance or decline over those three measurements.

Baseline Questionnaire

The following basic information was collected about each respondent:

- *Demographic variables* assessed age, gender, and educational level (1 = low [no education or only primary or primary-level vocational education], 2 = medium [secondary or secondary-level vocational education], and 3 = high [higher education]).
- *Perceived stress* was assessed by 5 items based on the Perceived Stress Scale [41] and was rated on a 5-point scale (1 = never, 5 = very often). The mean scale score was included in the analyses (Cronbach alpha = .75).
- *Level of depression* was measured by 10 items (eg, "I feel anxious") using the Center for Epidemiologic Studies Depression scale [42]. The answers were given on a 4-point scale (1 = never, 4 = all the time). The mean scale score was included in the analyses (alpha = .83).
- *Level of dependence* was assessed by 5 items (eg, the time when tobacco is smoked after awakening) based on an abbreviated Fagerström test [43,44]. The sum score was used in the analyses (0 = not dependent, 9 = very dependent).

The following questions about intention and motivational factors were based on questions used in previous studies [9,12,20,45-47]:

o *Attitudes* were measured on a 4-point scale (1 = no, 2 = yes, a bit, 3 = yes, 4 = yes, a lot/very) with four subscales measuring the following: (1) pros of not smoking, expressed by 9 statements representing positive expected outcomes (eg, "not smoking will save me money") (alpha = .73), (2) cons of not smoking, expressed by 9 statements representing negative expected outcomes (eg, "not smoking will cause me more stress") (alpha = .71), (3) pros of relapse, expressed by 4 statements representing positive expected outcomes (eg, "relapse will make me feel relieved") (alpha = .82), and (4) cons of relapse, expressed by 6 statements representing negative expected outcomes (eg, "relapse will make me feel weak") (alpha = .85).

o *Social support* was measured by 5 items that asked about the support received from significant others (partner, children, other family members, friends, or colleagues) on a 4-point scale (1 = no, 2 = yes, some, 3 = yes, moderate, and 4 = yes, a lot). A "not applicable" response was coded as missing. We created an index and used these scores in the analyses.

o *Modeling* was assessed by 5 items that measured the smoking status of the partner (yes, no, not applicable) and residential children, parents, colleagues, and friends (all, the majority, half, a minority, none, not applicable). The items were coded into three category scales (–1 = not smoking, 0 = not applicable, 1 = smoking). The items were summed and formed an index that we used in the analyses; smoking status of the partner was included separately in the analyses.

o *Self-efficacy* was assessed by 13 items that asked whether respondents thought they would be able to refrain from smoking in various smoking-related high-risk situations such as parties (alpha = .88). Recovery self-efficacy was assessed by 7 items that asked whether respondents thought they would be able to resume a smoking cessation program after a lapse (eg, after a week of abstinence) (alpha = .92). Answers were given on a 5-point scale (1 = totally disagree, 5 = totally agree) and the mean score was used in the analyses.

o *Preparatory planning* was assessed by 12 items (based on previous studies [8,9,12,46]) asking respondents to indicate whether they had plans to perform preparatory behaviors, such as removing all smoking materials, on a 5-point scale (1 = totally disagree, 5 = totally agree) (alpha = .72). The mean scale scores were used in the analyses.

Follow-up Measurements

Continued abstinence at 12 months after baseline was the primary outcome. In line with the definitions provided by Hughes et al [48], we defined continued abstinence as not smoking at all after the quit attempt. We measured it by asking whether the respondent had smoked since the quit date. Continued abstinence was coded as 0 = relapse and nonsmoking as 1 = abstinence.

Analyses

The most common approach to analyzing the effects of interventions is to include all respondents randomly assigned to the programs. The advantage to this approach is that it reflects the situation as it is likely to occur in practice (where low adherence levels and not following program protocols will also be common). This strategy also maintains the balance between randomly assigned groups and preserves sample sizes [49,50]. On the other hand, including all randomly assigned respondents is a very conservative approach that may be susceptible to type II errors [49,50] and may say little about the efficacy of a treatment, since respondents who did not receive the treatment or did not follow the protocol are still included [49,51]. A modified approach (which has been used more frequently in the last decades [52]) seems to be more suitable for gaining insight into the effects of programs among those who actually followed them. However, excluding respondents after randomization may threaten the randomization balance and neglect the fact that nonadherence (and even making or not making a quit attempt) may also be a result of the treatment [50]. Therefore, both approaches have their advantages and disadvantages. To provide a complete overview of our results, we analyzed three samples and report on the above-mentioned modified approach. This resulted in three different analytic approaches with three different samples:

- Sample 1 is the *sample as randomized*, including all respondents. Since this study focused on relapse prevention and therefore on continued abstinence, we also report on smaller samples:
- Sample 2, referred to as *the modified sample*, excluded respondents who indicated after their quit date as well as at follow-up that they did not make a quit attempt, and those who reported that they had quit less than 320 days before the follow-up measurement (this is after about 2 weeks after the first month in which they were supposed to make an attempt). We excluded the latter group because one of the inclusion criteria was that respondents needed to set a quit date within 1 month after inclusion. Including respondents who explicitly stated that they made their quit attempt much later or even too close to the follow-up measurement would result in differences between respondents in the time frame from the quit attempt to the follow-up measurement.
- Sample 3, referred to as the *minimum dose sample*, excluded respondents who did not adhere to at least one element of the program they were assigned to (ie, did not complete the prequit and postquit assignments) in addition to meeting the exclusion criteria in the modified sample.

For all three samples, we conducted an observed case analysis (only including respondents with follow-up data) and a conservative analysis in which we assumed that participants missing at follow-up had relapsed to smoking. We used a logistic regression analysis to compare the efficacy of the control condition and the SQ4U programs to foster continued smoking abstinence after 12 months. Differences in baseline factors (demographics, smoking-related factors, perceived stress, depression, intention strength, social influence, attitude, self-efficacy, recovery self-efficacy, and preparatory planning) were assessed using chi-square tests for categorical variables and analysis of variance for continuous variables. Attrition was analyzed by logistic regression and included baseline factors. To preserve power, we included only baseline factors on which the three conditions significantly differed at baseline as well as the factors that significantly predicted dropout at follow-up as covariates in the effect analyses. Listwise deletion with regard to missing values was used. Finally, chi-square tests were used to assess the existence of dose-response relationships within the SQ4U programs using the observed cases in sample 2 (the modified sample) to conform to the methods used by Borland et al [35].

Results

Sample Characteristics and Attrition Check

Table 1 presents the demographic and smoking-related characteristics at baseline and shows no significant differences between these variables in the three conditions. The mean age of the respondents was 41 (SD 11.80) years, and 1265 of the 2031 (62.28%) respondents were female. On average, respondents smoked 20 (SD 8.73) cigarettes per day, and 1887 (92.91%) had made previous attempts to quit. Furthermore, we found no significant (*P* > .05) differences between the three conditions on the baseline measurements of perceived stress, depression, intention strength, social influence, attitude, self-efficacy and recovery self-efficacy, and preparatory planning. (Table 1 shows only demographics and smoking-related factors.)

In the control condition, 434 of 636 (68.2%) were lost to follow-up versus 508 of 698 (72.8%) in the AP program and 523 of 697 (75.0%) in the AP+ program (Figure 1). Attrition analysis proved that dropout was more likely among respondents in the AP (odds ratio [OR] 1.26, *P* = .06) and AP+ (OR 1.44, *P* = .004) programs and respondents who were younger (OR 0.98, *P* < .001, 95% confidence interval [CI] 0.97–0.99), were male (OR 1.32, *P* = .02, 95% CI 1.05–1.66), had a lower education level (OR 2.00, *P* = .001, 95% CI 1.33–3.00), perceived fewer advantages to relapse (OR 0.80, *P* = .02, 95% CI 0.67–0.97), and had low levels of self-efficacy (OR 0.78, *P* = .001, 95% CI 0.67–0.91).

Table 1. Means and baseline differences between the three programs in demographic and smoking-related variables.

Characteristic	Overall (N = 2031)	Control (n = 636)	AP[a] (n = 698)	AP+[b] (n = 697)	*F* value/χ^2	*P* value
Female gender, n (%)	1265 (62.3%)	381 (59.9%)	442 (63.3%)	442 (63.4%)	2.2	.33
Age (years), mean (SD)	40.88 (11.80)	40.68 (11.81)	40.75 (11.48)	41.18 (12.12)	0.36	.70
Educational level, n (%)					6.4	.17
Low	207 (10.2%)	57 (9%)	86 (12%)	64 (9%)		
Medium	1130 (55.6%)	357 (56.1%)	371 (53.2%)	402 (57.7%)		
High	694 (34.2%)	222 (34.9%)	241 (34.5%)	231 (33.1%)		
Cigarettes smoked per day, mean (SD)	19.85 (8.73)	19.85 (8.39)	19.89 (9.36)	19.80 (8.41)	0.02	.98
Smoking duration (years), mean (SD)	24.81 (11.96)	24.61 (11.90)	24.79 (11.58)	25.01 (12.41)	0.19	.83
Previous quit attempts (yes), n (%)	1887 (92.9%)	588 (92.5%)	654 (93.7%)	645 (92.5%)	1.0	.61
Nicotine dependence, mean (SD)[d]	4.53 (2.18)	4.57 (2.21)	4.48 (2.13)	4.55 (2.19)	0.26	.77

[a] Action planning.

[b] Action planning plus.

[c] Degrees of freedom (DF) = 2 for all except for educational level (DF = 4).

[d] Sum score of abbreviated Fagerström test (0 = not dependent, 9 = very dependent).

Abstinence Rates 12 Months After Baseline

Table 2 presents the abstinence rates for samples 1, 2, and 3 when using observed cases alone, as well as when conducting conservative analyses and considering dropouts as relapsers. When using the observed case analysis strategy on sample 1, abstinence rates in the control group, the AP program, and the AP+ program were 22%, 33%, and 31%, respectively (see Table 2). In the sample where those who did not make a quit attempt conforming to the study protocol were excluded (sample 2), we found abstinence rates of 34%, 44%, and 37%, respectively. Finally, abstinence rates among only those who adhered to at least one of the intervention elements from the program they were assigned to (sample 3) showed abstinence rates of 34%, 46%, and 39%, respectively.

Table 2. 12-month abstinence rates per program for the three samples using observed and conservative analyses.

Sample	Observed				Conservative			
	n	n (%) abstinent			n	n (%) abstinent		
	Total	Control	AP[a]	AP+[b]	Total	Control	AP	AP+
1[c]	566	45 (22%)	63 (33%)	53 (31%)	2031	45 (7%)	63 (9%)	53 (8%)
2[d]	418	45 (34%)	63 (44%)	53 (37%)	1812	45 (8%)	63 (10%)	53 (8%)
3[e]	389	45 (34%)	60 (46%)	49 (39%)	1324	45 (8%)	60 (14%)	49 (15%)

[a] Action planning.

[b] Action planning plus.

[c] Including all respondents as randomly assigned.

[d] Modified sample excluding respondents who indicated after their quit date as well as at follow-up that they did not make a quit attempt during the study and respondents who quit too close to follow-up (see Analyses section in the Methods).

[e] Minimum dose sample additionally excluding those who adhered to none of the intervention elements of their SQ4U variant.

Main Effects of the AP and AP+ Programs With Correction for Covariates

Table 3 presents the main effects of the AP and AP+ programs compared with the control condition, while controlling for covariates (factors that predicted dropout according to the attrition analysis). The table shows these findings for observed case analyses as well as conservative analyses for sample 1 (including all respondents as randomly assigned), sample 2 (modified sample), and sample 3 (minimum dose sample). The AP program was significantly more effective in fostering abstinence than was the control condition in sample 1 (OR 1.95, *P* = .005), sample 2 (OR 1.71, *P* = .04), and sample 3 (OR 1.84, *P* = .02) when including observed cases only. The AP+ program was, however, only significant in sample 1 (OR 1.61, *P* =.049). In the conservative analyses, both programs were significantly more effective than the control condition only in sample 3 (AP: OR 1.72, *P* = .01; AP+: OR 1.76, *P* = .01). Finally, having high levels of self-efficacy was the only consistent predictor of continued abstinence in all these analyses.

Table 3. Regression 12-month continued abstinence on the AP[a] and AP+[b] program in sample 1 (observed cases, n = 559; conservative analysis, n = 1974), sample 2 (observed cases, n = 412; conservative analysis, n = 1757), and sample 3 (observed cases, n = 383; conservative analysis, n = 1297).

	Observed case analysis[c]			Conservative analysis[d]		
Variable	OR[e]	95% CI[f]	*P* value	OR	95% CI	*P* value
Sample 1 [g]						
Gender	1.15	0.77–1.71	.50	.91	0.64–1.29	.59
Age	1.01	0.99–1.03	.26	1.02	1.01–1.04	.002
Low education level (high[h])	1.61	0.75–3.43	.22	.76	0.40–1.43	.40
Pros of relapse	0.87	0.63–1.20	.39	1.00	0.76–1.32	.98
Self-efficacy	1.51	1.14–1.99	.004	1.64	1.29–2.07	<.001
AP program (control[h])	1.95	1.23–3.11	.005	1.38	0.92–2.08	.12
AP+ program (control[h])	1.61	1.00–2.60	.049	1.12	0.73–1.70	.61
Sample 2 [i]						
Gender	1.12	0.73–1.73	.59	.88	0.62–1.25	.48
Age	1.01	0.99–1.03	.28	1.03	1.01–1.04	.001
Low education level (high[h])	1.69	0.73–3.95	.22	.71	0.37–1.34	.29
Pros of relapse	0.81	0.57–1.15	.23	1.00	0.76–1.32	.99
Self-efficacy	1.39	1.03–1.89	.03	1.62	1.28–2.06	<.001
AP program (control[h])	1.71	1.03–2.83	.04	1.29	0.86–1.95	.23
AP+ program (control[h])	1.22	0.73–2.03	.44	.99	0.65–1.52	.98
Sample 3 [j]						
Gender	1.17	0.75–1.83	.49	.87	0.61–1.26	.47
Age	1.01	0.99–1.03	.31	1.02	1.01–1.04	.005
Low education level (high[h])	1.95	0.81–4.72	.14	.87	0.45–1.68	.68
Pros of relapse	0.86	0.60–1.23	.41	1.06	0.80–1.42	.68
Self-efficacy	1.43	1.04–1.96	.03	1.53	1.19–1.97	.001
AP program (control[h])	1.84	1.10–3.07	.02	1.72	1.13–2.61	.01
AP+ program (control[h])	1.36	0.80–2.29	.26	1.76	1.13-2.73	.01

[a] Action planning.

[b] Action planning plus.

[c] Sample including only respondents with follow-up data.

[d] Sample including missing data at follow-up as treatment failures.

[e] Odds ratio.

[f] Confidence interval.

[g] Including all respondents as randomly assigned.

[h] Reference category.

[i] Modified sample excluding respondents who indicated after their quit date as well as at follow-up that they did not make a quit attempt during the study and respondents who quit too close to follow-up (see Analyses section in the Methods).

[j] Minimum dose sample excluding those who did not adhere to at least one of the SQ4U elements.

Abstinence Rates per Program Stratified to Intervention Dose

Table 4 shows the 12-month continued abstinence rates per intervention dose for the modified sample (sample 2). As posed in hypothesis 2, the results revealed significant relationships between abstinence and the number of planning assignments using linear by linear association chi-square tests (AP: $\chi^2_1 = 7.4, P < .007$; AP+: $\chi^2_1 = 14.7, P < .001$) and feedback moments

(AP+: $\chi^2_1 = 24.5$, $P < .001$) and confirmed higher abstinence rates when more planning assignments or questionnaires after the quit attempt were completed. On average, respondents in the program groups adhered to 4 out of 6 planning assignments and 6 out of 11 feedback questionnaires.

Table 4. 12-month continued abstinence rates stratified by the number of planning assignments and feedback moments in the modified sample (sample 2).

		AP[a]		AP+[b]	
Stratification	Dose	n[c]	n (%) abstinent	n[c]	n (%) abstinent
Per number of assignments	0–1	27	6 (22%)	23	4 (17%)
	2–4	53	24 (45%)	49	13 (27%)
	5–6	62	33 (53%)	70	36 (51%)
	Total, mean (SD)		3.71 (2.00)		3.95 (1.93)
	χ^2_1		7.4		14.7
	P value		.007		<.001
Per number of feedback moments	0–5	NA[d]	NA	66	10 (15%)
	6–7	NA	NA	12	4 (33%)
	8–9	NA	NA	31	19 (61%)
	10–11	NA	NA	33	20 (61%)
	Total, mean (SD)		NA		5.77 (3.83)
	χ^2_1		NA		24.5
	P value		NA		<.001

[a] Action planning.

[b] Action planning plus.

[c] Only complete cases.

[d] Not applicable, as the AP program did not provide tailored feedback after the quit date.

Program Evaluation of the AP and AP+ Programs

Table 5 presents the program evaluation that was conducted 6 months after baseline. Respondents from the AP and AP+ programs gave the baseline-tailored feedback a positive evaluation. However, significantly more respondents in the AP+ program remembered the content of the baseline-tailored feedback and perceived the feedback as relevant and helpful for making the quit attempt. The planning assignments were perceived as useful by 87 of 164 (53%) respondents. Moreover, 724 (44%) of the respondents rated them as helpful for making a quit attempt and 51 (31%) agreed they were helpful in maintaining a quit attempt. Respondents from the AP and AP+ programs did not differ significantly on the evaluation of the planning assignments. The feedback after the quit attempt (short feedback moments) was rated as useful by 66 of 104 (63%) respondents, as helpful for making a quit attempt by 49 (47%), and as helpful for maintaining the quit attempt by 50 (48%).

Table 5. Program evaluation (conducted 6 months after baseline) by respondents from the AP[a] and the AP+[b] programs.

Evaluation item	Total (n = 248)	AP (n = 137)	AP+ (n = 111)	χ^2_2	*P* value
Baseline feedback					
Remembered the content				9.7	.008
Yes	142 (57.3%)	72 (53%)	70 (63%)		
Neutral	63 (25%)	32 (23%)	31 (28%)		
No	43 (17%)	33 (24%)	10 (9%)		
Perceived feedback as useful				4.4	.111
Yes	156 (62.9%)	79 (58%)	77 (69%)		
Neutral	68 (27%)	41 (30%)	27 (24%)		
No	24 (10%)	17 (12%)	7 (6%)		
Perceived feedback as relevant				8.8	.012
Yes	121 (48.8%)	56 (41%)	65 (59%)		
Neutral	98 (40%)	60 (44%)	38 (34%)		
No	29 (12%)	21 (15%)	8 (7%)		
Perceived feedback as understandable				3.8	.152
Yes	208 (83.9%)	110 (80.3%)	98 (88%)		
Neutral	34 (14%)	24 (18%)	10 (9%)		
No	6 (2%)	3 (2%)	3 (3%)		
Recognized own situation in feedback				0.2	.92
Yes	117 (47.2%)	63 (46%)	54 (49%)		
Neutral	101 (40.7%)	57 (42%)	44 (40%)		
No	30 (12%)	17 (12%)	13 (12%)		
Perceived feedback as credible				3.6	.16
Yes	170 (68.5%)	87 (64%)	83 (75%)		
Neutral	67 (27%)	43 (31%)	24 (22%)		
No	11 (4%)	7 (5%)	4 (4%)		
Feedback helped to make a quit attempt				6.7	.04
Yes	107 (43.1%)	51 (37%)	56 (50%)		
Neutral	73 (29%)	49 (36%)	24 (22%)		
No	68 (27%)	37 (27%)	31 (28%)		
Feedback helped to maintain quit attempt				3.7	.16
Yes	72 (29%)	33 (24%)	39 (35%)		
Neutral	65 (26%)	38 (28%)	27 (24%)		
No	111 (44.8%)	66 (48%)	45 (41%)		
Planning assignments	n = 164	n = 93	n = 71		
Perceived feedback as useful				4.0	.14
Yes	87 (53%)	49 (53%)	38 (54%)		
Neutral	62 (38%)	32 (34%)	30 (42%)		
No	15 (9%)	12 (13%)	3 (4%)		
Feedback helped to make a quit attempt				5.6	.06
Yes	72 (44%	37 (40%)	35 (49%)		
Neutral	57 (35%)	30 (32%)	27 (38%)		

Evaluation item	Total (n = 248)	AP (n = 137)	AP+ (n = 111)	χ^2_2	*P* value
No	35 (21%)	26 (28%)	9 (13%)		
Feedback helped to maintain quit attempt				5.1	.08
Yes	51 (31%)	27 (29%)	24 (34%)		
Neutral	56 (34%	27 (29%)	29 (41%)		
No	57 (35%)	39 (42%)	18 (25%)		
Feedback moments after quit attempt (n = 104)					
Perceived feedback as useful					
Yes	66 (63%)	NA[c]	66 (63%)		
Neutral	26 (25%)	NA	26 (25%)		
No	12 (12%)	NA	12 (12%)		
Feedback helped to make a quit attempt					
Yes	49 (47%)	NA	49 (47%)		
Neutral	32 (31%)	NA	32 (31%)		
No	23 (22%)	NA	23 (22%)		
Feedback helped to maintain quit attempt					
Yes	50 (48%)	NA	50 (48%)		
No	54 (52%)	NA	54 (52%)		

[a] Action planning.

[b] Action planning plus.

[c] Not applicable.

Discussion

The main goal of this study was to evaluate the efficacy of two Web-based smoking relapse prevention programs (AP and AP+) in the SQ4U study with regard to 12-month continued smoking abstinence. Furthermore, we aimed to assess dose-response associations. Despite the potential of the Internet, a recurrent problem in Internet trials is low adherence to the programs [53]. These low levels of adherence may lead to underestimations of intervention effects. This impression is strengthened by the fact that most studies of Internet interventions have shown dose-response relationships (see, for instance, [35,54-58]). Hence, an alternative strategy in which nonadherent respondents are excluded may be needed. Recently, a systematic review has discussed an increase in the literature of randomized controlled trials that report on a modified strategy [52]. To provide a complete overview of our results, we employed both approaches and reported on three samples: (1) the sample as randomly assigned, including all respondents, (2) a sample excluding those who did not make a quit attempt (modified sample), and (3) a sample additionally excluding those who did not adhere to at least one of the intervention elements (minimum dose sample). Observed case analyses (only including respondents with follow-up data) and conservative analyses (including all cases and coding respondents who were missing at follow-up as smoking relapsers) were conducted on all samples.

The results of all three observed case analyses (samples 1, 2, and 3) revealed significant effects in favor of the AP program. We can conclude that most of the analyses support our first hypothesis that planning strategies can be effective in preventing relapse among smokers who are motivated to quit smoking. The AP+ program, on the other hand, only proved to be significantly more effective than the control condition in the randomized sample (sample 1). The approach for sample 2 seems to have created a bias against the AP+ program by excluding respondents who did not make a quit attempt during the study, which resulted in more exclusions of relapsers (since nonquitters cannot be considered to be continued abstinent at follow-up) in the control condition than in the AP+ program (which consisted of elements designed to foster making a quit attempt). This approach may therefore have been too conservative. Furthermore, additional power analyses showed that 2623 respondents would have been needed to find significant differences in abstinence rate between the control group (45/134, 34%) and the AP+ program group (53/142, 37%). Therefore, the lack of significance in sample 2 may be due to power issues. Another possible explanation for lack of significant effect for the AP+ program in sample 2 and sample 3 may be that the program was too intensive and therefore resulted in an overload for the respondents that negatively influenced its efficacy. A comparable result was found in an earlier study conducted among vocational students, which tested the efficacy of a standard in-school program, a computer-tailored program, and a combined program [59]. This study found that the combined strategy did not have a surplus value. The authors suggested that combining strategies may lead to an overload of information, which may produce more negative effects.

Although this may also account for our findings, program evaluation data do not support this suggestion because the AP+ and AP programs both received a positive evaluation. A final explanation may be that the planning assignments and the feedback provided after the quit attempt may not have been an ideal combination to foster abstinence after all.

When conducting conservative analyses, we found the AP and AP+ programs to be significantly more effective than the control condition in the minimum dose sample (sample 3). However, given the large dropout rate, this conservative approach (coding missing participants as relapsers) may be too conservative. Additional descriptive analyses indicated that about 60% of the respondents who dropped out at follow-up were nonsmokers at their last visit, indicating that interpreting all missing cases as treatment failures may be too conservative, a finding also supported by others [60]. Finally, conservative analyses in the minimum dose sample (sample 3) may have biased the results in favor of the experimental conditions by excluding nonadherers (whose follow-up data are generally missing). Excluding these nonadherers results in fewer imputations of relapsers in the experimental conditions than in the control condition, as these respondents are already excluded because they did not adhere to the programs.

The abstinence rate of 22% in the control group is another notable finding, since self-quitters generally reach abstinence rates of only 3%-5% [1]. A possible explanation for this high abstinence rate may be that respondents in our control condition, who were highly motivated to quit smoking, sought additional help themselves when they did not receive the help they expected from our program. Since not all respondents reported on their use of additional help, we were not able to test this assumption with the available data. Moreover, the high motivation of these respondents may have played a role in their high abstinence rates.

Finally, the dose-response relationship between abstinence and the number of program elements suggests that a dose-response association may exist. This is because abstinence rates increased by up to 53% (33/62) in the AP program and 51% (36/70) in the AP+ program when doing 5 to 6 planning assignments, and up to 61% (20/33) in the AP+ program when filling out 10-11 feedback questionnaires (on which feedback was provided). These findings suggest, in line with hypothesis 2, that the efficacy of the programs depends on adherence to the program, also found in previous studies [35,61]. Our findings regarding dose-response relationships may, however, be attributed to the fact that respondents who relapsed after a few sessions discontinued the program, resulting in the finding that those who continued with the program were more successful. Caution is, therefore, needed when interpreting dose-response relationships: these relationships can be subject to selection biases, since the respondents are not randomly assigned to different doses [62]. Additional research that tests the assumption of a dose-response relationship in this context is needed in the form of an experimental design in which respondents are randomly assigned to groups with different doses.

The SQ4U study was subject to limitations. The first limitation is that the planning assignments and feedback moments were provided at fixed times, while the varying levels of adherence found in our study suggest that programs should perhaps provide support when the respondent needs it most (ie, by real-time support in difficult situations). Research is needed to explore the potential additional efficacy of such an approach. Second, the cut-off point for the minimum dose sample (sample 3) is not based on empirical findings and needs to be explored in additional studies. Third, because of medical ethics guidelines, we could not prevent respondents from using additional help to quit smoking. The use of additional help, however, may interfere with the effects of the programs and may be beneficial or counterproductive. Further research is needed to explore which additional aids may have positive or negative effects. Fourth, our study had a high rate of loss to follow-up (1465/2031, 72.1%), an issue that is very common in comparable studies [12,21,63,64]. Attrition may have been caused by factors such as spam filters or invalid email accounts or because people who have quit smoking do not want to be reminded of their past smoking behavior [65]. The latter is partly supported by our data, which showed that about 60% of the respondents who dropped out of the experimental programs were nonsmokers at their last visit. Finally, the conservative analyses (in which missing answers at follow-up were regarded as relapse) may be too conservative, as our data showed that about 60% of the respondents who dropped out of the experimental programs were nonsmokers at their last visit. Furthermore, attrition analyses indicated that dropout was more likely in the experimental programs. Consequently, the conservative analyses can be subject to biases strengthening the relation between the experimental programs and relapse, while the actual relation is one between the programs and dropout. Therefore, caution is needed with interpreting these results.

Aside from these limitations, our SQ4U study is the first to test the efficacy of incorporating planning assignments in a Web-based computer-tailored relapse prevention program and of combining planning assignments with multiple tailored feedback moments after the quit date. The study reflected on the findings from different samples, and the results pointed out the importance of using planning strategies and tailored feedback moments after the quit date for smokers who are motivated to quit. Previous studies indicated that a lack of preparatory planning is associated with smoking relapse [8,9]. Our current study further illuminates the role of planning by showing that managing behavior using action planning strategies in relapse prevention programs fosters abstinence. Additional research is needed to determine the optimum dose for reaching best effects and which planning strategies are most effective for which groups. Since some of the respondents in the experimental programs did not adhere to even one intervention element, insight into the predictors of adherence and into strategies to facilitate adherence are also needed.

Acknowledgments

This study was funded by the Dutch Organization for Health Research and Innovation (grant number 6130-0030). We would like to thank Claire Jeukens and Verina Servranck for their contribution to the development and piloting of the SQ4U program materials. We would also like to thank Stivoro for their help in recruiting respondents, as well as the respondents for their participation in the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT eHealth V1.6.1 checklist [66].

[PDF File (Adobe PDF File), 558KB - jmir_v14i4e109_app1.pdf]

References

1. Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. Addiction 2004 Jan;99(1):29-38. [Medline: 14678060]
2. Dijkstra A, Borland R. Residual outcome expectations and relapse in ex-smokers. Health Psychol 2003 Jul;22(4):340-346. [Medline: 12940389]
3. Kenford SL, Smith SS, Wetter DW, Jorenby DE, Fiore MC, Baker TB. Predicting relapse back to smoking: contrasting affective and physical models of dependence. J Consult Clin Psychol 2002 Feb;70(1):216-227. [Medline: 11860048]
4. Killen JD, Fortmann SP, Kraemer HC, Varady AN, Davis L, Newman B. Interactive effects of depression symptoms, nicotine dependence, and weight change on late smoking relapse. J Consult Clin Psychol 1996 Oct;64(5):1060-1067. [Medline: 8916636]
5. Marlatt GA, Gordon JA. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. New York, NY: Guilford Press; 1985.
6. Shiffman S, Waters AJ. Negative affect and smoking lapses: a prospective analysis. J Consult Clin Psychol 2004 Apr;72(2):192-201. [doi: 10.1037/0022-006X.72.2.192] [Medline: 15065954]
7. Hajek P, Stead LF, West R, Jarvis M, Lancaster T. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2009(1):CD003999. [doi: 10.1002/14651858.CD003999.pub3] [Medline: 19160228]
8. Elfeddali I, Bolman C, Candel MJ, Wiers RW, De Vries H. The role of self-efficacy, recovery self-efficacy, and preparatory planning in predicting short-term smoking relapse. Br J Health Psychol 2012 Feb;17(1):185-201. [doi: 10.1111/j.2044-8287.2011.02032.x] [Medline: 22107073]
9. Hoving EF, Mudde AN, de Vries H. Predictors of smoking relapse in a sample of Dutch adult smokers; the roles of gender and action plans. Addict Behav 2006 Jul;31(7):1177-1189. [doi: 10.1016/j.addbeh.2005.09.002] [Medline: 16242246]
10. O'Connell KA, Hosein VL, Schwartz JE, Leibowitz RQ. How does coping help people resist lapses during smoking cessation? Health Psychol 2007 Jan;26(1):77-84. [doi: 10.1037/0278-6133.26.1.77] [Medline: 17209700]
11. van Osch L, Lechner L, Reubsaet A, Wigger S, de Vries H. Relapse prevention in a national smoking cessation contest: effects of coping planning. Br J Health Psychol 2008 Sep;13(Pt 3):525-535. [doi: 10.1348/135910707X224504] [Medline: 17631695]
12. Te Poel F, Bolman C, Reubsaet A, de Vries H. Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months. Health Educ Res 2009 Dec;24(6):930-940 [FREE Full text] [doi: 10.1093/her/cyp036] [Medline: 19574405]
13. van der Plas AGM, Hilberink S, Hermans M, Breteler M. Evaluatie van de Millennium Campagne 'Dat kan ik ook!' en regionale cursussen stoppen met roken: Resultaten en predictoren van succes [Evaluation of the Millenium Campaign and Regional Smoking Cessation Programs: Results and Predictors of Success]. Nijmegen: Nijmegen Institute for Scientifist-Practitioners in Addiction; 2001.
14. de Vries H, Kremers SP, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]
15. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]
16. Brug J, Campbell M, van Assema P. The application and impact of computer-generated personalized nutrition education: a review of the literature. Patient Educ Couns 1999 Feb;36(2):145-156. [Medline: 10223019]
17. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
18. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]

19. Anonymous. Statline: vrije tijd en cultuur; internet toegang huishoudens 2005-2010 [Statline: spare time and culture; internet access households 2005-2010] [database online]. The Hague/Heerlen: Centraal Bureau voor de Statistiek; 2011.
20. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. Health Educ Res 1998 Jun;13(2):207-214 [FREE Full text] [Medline: 10181019]
21. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [doi: 10.2196/jmir.7.1.e2] [Medline: 15829474]
22. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
23. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. Addiction 2005 May;100(5):682-688. [doi: 10.1111/j.1360-0443.2005.01093.x] [Medline: 15847626]
24. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. Patient Educ Couns 1999 Feb;36(2):193-203. [Medline: 10223023]
25. Kreuter MW, Strecher VJ, Glassman B. One size does not fit all: the case for tailoring print materials. Ann Behav Med 1999;21(4):276-283. [Medline: 10721433]
26. Brug J, Oenema A, Campbell M. Past, present, and future of computer-tailored nutrition education. Am J Clin Nutr 2003 Apr;77(4 Suppl):1028S-1034S [FREE Full text] [Medline: 12663313]
27. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998;13(4):623-649. [doi: 10.1080/08870449808407422]
28. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [Medline: 14672383]
29. Kreuter MW, Caburnay CA, Chen JJ, Donlin MJ. Effectiveness of individually tailored calendars in promoting childhood immunization in urban public health centers. Am J Public Health 2004 Jan;94(1):122-127. [Medline: 14713709]
30. 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]
31. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
32. Smeets T, Brug J, de Vries H. Effects of tailoring health messages on physical activity. Health Educ Res 2008 Jun;23(3):402-413 [FREE Full text] [doi: 10.1093/her/cyl101] [Medline: 17032705]
33. Dijkstra A, De Vries H, Roijackers J, van Breukelen G. Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions. Health Psychol 1998 Nov;17(6):513-519. [Medline: 9848801]
34. Brug J, Glanz K, Van Assema P, Kok G, van Breukelen GJ. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. Health Educ Behav 1998 Aug;25(4):517-531. [Medline: 9690108]
35. Borland R, Balmford J, Hunt D. The effectiveness of personally tailored computer-generated advice letters for smoking cessation. Addiction 2004 Mar;99(3):369-377. [doi: 10.1111/j.1360-0443.2003.00623.x] [Medline: 14982550]
36. Elfeddali I, Bolman C, de Vries H. SQ4U - a computer tailored smoking relapse prevention program incorporating planning strategy assignments and multiple feedback time points after the quit-attempt: development and design protocol. Contemp Clin Trials 2012 Jan;33(1):151-158. [doi: 10.1016/j.cct.2011.09.008] [Medline: 22015574]
37. Gwaltney CJ, Shiffman S, Balabanis MH, Paty JA. Dynamic self-efficacy and outcome expectancies: prediction of smoking lapse and relapse. J Abnorm Psychol 2005 Nov;114(4):661-675. [doi: 10.1037/0021-843X.114.4.661] [Medline: 16351387]
38. Shiffman S, Balabanis MH, Gwaltney CJ, Paty JA, Gnys M, Kassel JD, et al. Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. Drug Alcohol Depend 2007 Dec 1;91(2-3):159-168. [doi: 10.1016/j.drugalcdep.2007.05.017] [Medline: 17628353]
39. Ferguson SF, Shiffman S, Gwaltney CJ. Does reducing withdrawal severity mediate nicotine patch efficacy? A randomized controlled trial. J Consult Clin Psychol 2004;74:1153-1161. [doi: 10.1037/0022-006X.74.6.1153]
40. Shiffman S, Paty JA, Gnys M, Kassel JA, Hickcox M. First lapses to smoking: within-subjects analysis of real-time reports. J Consult Clin Psychol 1996 Apr;64(2):366-379. [Medline: 8871421]
41. Cohen S, Williamson G. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S. editors. The Social Psychology of Health. Beverly Hills, CA: Sage Publications; 1988:31-67.
42. Cole JC, Rabin AS, Smith TL, Kaufman AS. Development and validation of a Rasch-derived CES-D short form. Psychol Assess 2004 Dec;16(4):360-372. [doi: 10.1037/1040-3590.16.4.360] [Medline: 15584795]
43. Fagerstrom KO, Heatherton TF, Kozlowski LT. Nicotine addiction and its assessment. Ear Nose Throat J 1990 Nov;69(11):763-765. [Medline: 2276350]
44. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
45. Dijkstra A, de Vries H. Self-efficacy expectation with regard to different tasks in smoking cessation. Psychol Health 2000;15(4):501-511. [doi: 10.1080/08870440008402009]

46. Hilberink SR, Jacobs JE, Schlösser M, Grol RP, de Vries H. Characteristics of patients with COPD in three motivational stages related to smoking cessation. Patient Educ Couns 2006 Jun;61(3):449-457. [doi: 10.1016/j.pec.2005.05.012] [Medline: 16157462]
47. Mudde A, Willemsen M, Kremers S, de Vries H. Meetinstrumenten voor onderzoek naar roken en stoppen met roken [Measurements for Studies About Smoking and Smoking Cessation]. Den Haag: Stivoro; 2006.
48. Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE. Measures of abstinence in clinical trials: issues and recommendations. Nicotine Tob Res 2003 Feb;5(1):13-25. [Medline: 12745503]
49. Heritier SR, Gebski VJ, Keech AC. Inclusion of patients in clinical trial analysis: the intention-to-treat principle. Med J Aust 2003 Oct 20;179(8):438-440 [FREE Full text] [Medline: 14558871]
50. Wertz R. Intention to treat: once randomized, always analyzed. Clin Aphasiol 1995;23:57-64.
51. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999 Sep 11;319(7211):670-674. [Medline: 10480822]
52. Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 2010;340:c2697 [FREE Full text] [Medline: 20547685]
53. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. Patient Prefer Adherence 2008;2:57-65 [FREE Full text] [Medline: 19920945]
54. An LC, Perry CL, Lein EB, Klatt C, Farley DM, Bliss RL, et al. Strategies for increasing adherence to an online smoking cessation intervention for college students. Nicotine Tob Res 2006 Dec;8 Suppl 1:S7-12. [Medline: 17491165]
55. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216. [doi: 10.1080/14622200500055319] [Medline: 16036277]
56. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8 Suppl 1:S59-S67. [Medline: 17491172]
57. Eysenbach G. Issues in evaluating health websites in an Internet-based randomized controlled trial. J Med Internet Res 2002 Dec;4(3):E17 [FREE Full text] [doi: 10.2196/jmir.4.3.e17] [Medline: 12554548]
58. Steele RM, Mummery WK, Dwyer T. Examination of program exposure across intervention delivery modes: face-to-face versus internet. Int J Behav Nutr Phys Act 2007;4:7 [FREE Full text] [doi: 10.1186/1479-5868-4-7] [Medline: 17352817]
59. Ausems M, Mesters I, van Breukelen G, De Vries H. Effects of in-school and tailored out-of-school smoking prevention among Dutch vocational school students. Health Educ Res 2004 Feb;19(1):51-63 [FREE Full text] [doi: 10.1093/her/cyg001] [Medline: 15020545]
60. Lachin JM. Worst-rank score analysis with informatively missing observations in clinical trials. Control Clin Trials 1999 Oct;20(5):408-422. [Medline: 10503801]
61. Fiore MC, Jaen CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. US Department of Health and Human Services, Public Health Service. 2008 May. Clinical Practice Guideline: Treating Tobacco Use and Dependence: 2008 Update URL: http://www.ahrq.gov/clinic/tobacco/treating_tobacco_use08.pdf [accessed 2012-07-22] [WebCite Cache ID 69Lp07Hry]
62. McGowan HM, Nix RL, Murphy SA, Bierman KL, Conduct Problems Prevention Research Group. Investigating the impact of selection bias in dose-response analyses of preventive interventions. Prev Sci 2010 Sep;11(3):239-251. [doi: 10.1007/s11121-010-0169-2] [Medline: 20180151]
63. 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]
64. Tate DF, Zabinski MF. Computer and Internet applications for psychological treatment: update for clinicians. J Clin Psychol 2004 Feb;60(2):209-220. [doi: 10.1002/jclp.10247] [Medline: 14724928]
65. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
66. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AP: Action planning (program)
AP+: Action planning plus (program)
CI: confidence interval
I-Change Model: Integrated Change Model
OR: odds ratio
SQ4U: Stay Quit for You

Edited by G Eysenbach; submitted 19.01.12; peer-reviewed by R Borland, J McClure, N Zhang, L Vervoort; comments to author 05.02.12; revised version received 29.04.12; accepted 24.05.12; published 20.08.12.

Please cite as:
Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H
Preventing Smoking Relapse via Web-Based Computer-Tailored Feedback: A Randomized Controlled Trial
J Med Internet Res 2012;14(4):e109
URL: http://www.jmir.org/2012/4/e109/
doi:10.2196/jmir.2057
PMID:22903145

Tutorial

Design Considerations in Developing a Text Messaging Program Aimed at Smoking Cessation

Michele L Ybarra[1], MPH, PhD; Jodi Summers Holtrop[2], PhD; A Tülay Bağci Bosi[3], PhD; Salih Emri[4], MD

[1]Center for Innovative Public Health Research, San Clemente, CA, United States

[2]Department of Family Medicine, Michigan State University, Lansing, MI, United States

[3]Department of Public Health, Hacettepe University, Ankara, Turkey

[4]Department of Chest Diseases, Faculty of Medicine, Hacettepe University, Ankara, Turkey

Corresponding Author:
Michele L Ybarra, MPH, PhD
Center for Innovative Public Health Research
555 El Camino Real
A347
San Clemente, CA, 92672
United States
Phone: 1 8773026858 ext 801
Fax: 1 877 360 7627
Email: Michele@innovativepublichealth.org

Abstract

Background: Cell phone text messaging is gaining increasing recognition as an important tool that can be harnessed for prevention and intervention programs across a wide variety of health research applications. Despite the growing body of literature reporting positive outcomes, very little is available about the design decisions that scaffold the development of text messaging-based health interventions. What seems to be missing is documentation of the thought process of investigators in the initial stages of protocol and content development. This omission is of particular concern because many researchers seem to view text messaging as the intervention itself instead of simply a delivery mechanism. Certainly, aspects of this technology may increase participant engagement. Like other interventions, however, the content is a central driver of the behavior change.

Objective: To address this noted gap in the literature, we discuss the protocol decisions and content development for SMS Turkey (or *Cebiniz birakin diyor* in Turkish), a smoking cessation text messaging program for adult smokers in Turkey.

Methods: Content was developed in English and translated into Turkish. Efforts were made to ensure that the protocol and content were grounded in evidence-based smoking cessation theory, while also reflective of the cultural aspects of smoking and quitting in Turkey.

Results: Methodological considerations included whether to provide cell phones and whether to reimburse participants for texting costs; whether to include supplementary intervention resources (eg, personal contact); and whether to utilize unidirectional versus bidirectional messaging. Program design considerations included how messages were tailored to the quitting curve and one's smoking status after one's quit date, the number of messages participants received per day, and over what period of time the intervention lasted.

Conclusion: The content and methods of effective smoking cessation quitline programs were a useful guide in developing SMS Turkey. Proposed guidelines in developing text messaging-based behavior change programs are offered.

(J Med Internet Res 2012;14(4):e103) doi:10.2196/jmir.2061

KEYWORDS

Smoking cessation; mHealth; text messaging

Introduction

Cell phone text messaging is gaining increasing recognition as an important tool with a wide variety of health research applications. A search on the US Department of Health and Human Services' Research Portfolio Online Reporting Tools (RePORT) website [1] using the keyword *text messaging* reveals over 40 studies currently funded that have a text messaging component. Surely, this is just a subset of the larger field of text messaging-based health programming, funded both privately and publicly, across the world, yet it demonstrates the public health field's enthusiasm for this new technology.

Research studies have used text messaging as a data collection tool (eg, principal investigator [PI]: Boushey, 5U01CA130784 [2]; PI: Mundt, 1R43MH086152 [3]) and as a medication adherence enhancer (eg, PI: Belzer, 5U01HD040463 [4]). Perhaps the most interactive application of text messaging in the health arena is to change health behavior. Text messaging has been tested as the delivery mechanism for the main intervention content (eg, PI: Bull, 1R21MH083318 [5]; PI: Olson, 1R21HS018214 [6]) or for booster content that enhances content delivered online (eg, PI: Cornelius, 5R21NR011021 [7]). Data are still emerging and, at this stage, are preliminary. Nonetheless, reviews of the available literature find reason for optimism. A recent review reported 16 randomized controlled trials that involved text messaging, 10 of which reported significant improvement in their outcome measures; the remaining 6 reported positive trends [8]. Strong acceptability among intervention participants also was noted. Another recent review, focused specifically on health behavior change interventions, reported 9 sufficiently powered studies. Of these, 8 reported results supportive of a conclusion that text messaging can deliver content that affects behavior change [9].

Despite the growing body of literature documenting the outcomes of text messaging-based (sometimes also called mobile health or mHealth) interventions, very little is available about the development of these interventions. Owens and colleagues reported their experience developing a self-harm intervention to be delivered via text messaging. Based on feedback from service users and providers, the researchers chose a model whereby the participants created their own content, which then could be sent on demand in times of crisis [10]. Whittaker and colleagues reported the steps they followed to develop a multimedia smoking cessation program, including intensive focus group testing and pilot testing [11]. This latter study is a particularly useful guide in how to engage one's user audience during the intervention development stage. What seems to be missing, however, are scientific papers that describe the practical decisions taken in developing the first drafts of the content of a text messaging program. As Cole-Lewis and Kershaw noted [9], many researchers seem to view text messaging as the intervention, yet it is simply the delivery mechanism. Certainly, there are benefits of text messaging communication that may increase the engagement and salience of the program content for participants. Like other interventions, however, it is the content, not the delivery mechanisms (eg, Internet, in-person), that is the central driver of behavior change.

Methods

To address this noted gap in the literature, we describe the program development for SMS Turkey. Compared with the United States, where 23% of men and 18% of women are current smokers [12], an estimated 44% of men and 12% of women smoke daily in Turkey [13]. Despite Turkey's high smoking prevalence rate, over half of smokers report a desire to quit, and 45% have made a quit attempt in the past year. To invigorate cessation rates, smoking cessation programs need to be easily accessible and have high reach. With an estimated 62 million cell phones in Turkey [14], there is sufficient reason for optimism about the feasibility of a text messaging-based smoking cessation program in this setting. We were funded by the US National Institutes of Health's Fogarty International Center (R01TW007918) to develop and test SMS Turkey, a smoking cessation program delivered via text messaging. The program was designed in 2007–2008 and was created for adults seriously thinking about quitting smoking and living in Ankara, Turkey. Here, we describe the initial content and protocol development. Findings not only document the methodological development of SMS Turkey, but also provide direction for other researchers endeavoring to create mHealth behavior change programs.

Results

Methodological Considerations

Prior to developing the content, there are several methodological decisions to be made, many of which are determined by timeline and budget. One consideration is whether the program will include bidirectional messaging, whereby the participant provides input via text messaging that is then captured by the intervention software and responded to [9]. This can increase participants' commitment to the study, and feelings that they are playing an active role in the intervention rather than passively receiving messages. It also can serve as an opportunity to tailor messaging or collect real-time data about the participant's status (eg, current smoking behavior, level of cravings). On the other hand, the programming of the software to receive messages is much more complicated and therefore more costly and time intensive to develop than a unidirectional program that sends but does not receive and respond to participant messages. For pilot projects, it may be best to develop a unidirectional program and have research staff execute any bidirectional messaging through manual means. Once feasibility is determined, then a more complicated and costly program can be developed at the next stage. In the case of SMS Turkey, we chose to develop a unidirectional software program because of the intervention's pilot nature.

Another important consideration is whether the research program will provide phones or text messaging plans for participants; or reimburse participants' costs associated with receiving text messages. This decision should be based on the study goals. Some studies are aimed at harnessing the mobility of technology: for example, using cell phones to consistently reach unstable populations (eg, the marginally housed). In this case, it may be appropriate to give participants phones. In other studies, the motivation is to take advantage of the explosive increase in text messaging among adolescents and adults [15]. In this case, researchers need to be clear that the intervention is not intended for everyone (indeed, it is unlikely that there is one unique program that will address the needs and interests of all people at risk for a particular outcome), but rather for those who have adopted text messaging. In this case, the eligibility criteria should include having a text-capable phone, as well as an unlimited text messaging plan. If the target population is not using text messaging, then this mode may not be the best way to reach and engage the population.

In SMS Turkey, participants were required to have a cell phone and have used text messaging in the past year. It is free to receive text messages in Turkey; otherwise, we also would have required participants to be enrolled in an unlimited text messaging plan. If we had instead given participants cell phones or reimbursed them for costs associated with receiving program messages, we believe the findings had the potential to be adversely affected in three important ways. First, feedback from the participants about their study experience might not have reflected the views of the intended audience. Second, a study focus would necessarily have become management of the phones themselves. Third, the resulting data would have been less informative for potential scale-up of the intervention because agencies that might adopt the intervention are not likely to have resources to provide cell phones.

Design Considerations

Although text messaging interventions are relatively new, effective interventions using other modalities likely exist that can be used to guide the design. In smoking cessation, quitlines (ie, cessation counseling delivered via telephone) are widely available and known to be an effective method of counseling that reaches many smokers [16-18]. Quitlines were an especially amenable guide for SMS Turkey because they are grounded in behavior change theory and use ongoing contact between an interventionist (phone counselor) and participant (smoker), similar to how participants may interact with the information received through the text messages in an ongoing fashion to affect behavior change. Also, the proactive nature of the counselor calling the participant is more similar to text messaging-based communication compared with the reactive nature of most Internet-based interventions that rely on participants to log on to a website for information [19].

Another early decision that needs to be made is the overall study length: over what period of time will messages be sent? The intervention needs to be long enough to affect behavior change without being so long as to cause participants to lose interest and drop out. SMS Turkey was developed to be a 6-week program: 2 weeks of prequit and 4 weeks of postquit messages. This was based on the length of most quitline programs [20-22] and on successful implementation of a text messaging-based smoking cessation program in New Zealand of similar length [23]. Future interventions may explore longer program periods.

Another decision is the number of messages that will be sent per day or per week. This should be based on the target population (eg, how many messages do they receive in a typical day?) and on the literature associated with the target behavior. Similar to the overall intervention length, it is important to deliver a sufficient intensity of messages to affect behavior change, while not overwhelming participants to the point where they no longer read the messages. Our development survey suggested that almost half of adult smokers in our target population texted daily [24]. Therefore, we felt comfortable that a daily schedule that fluctuated based on the quitting curve would be appropriate. For smoking cessation, telephone-based programs' sessions increase in frequency around the quit date, and then gradually reduce in frequency as the participant gets further along the relapse curve associated with cessation [25]. We tried to create a similar experience in SMS Turkey: for the first 2 weeks leading into cessation, participants received three messages per day. As participants got closer to the quit day, we sent five messages per day. On the quit day and the following day, we sent eight messages each day. For the next 2 days, six messages were sent; this was reduced to five messages on the next day. We sent four messages on each of the last 2 days of their first postquit week. For the next 2 weeks, participants received two messages per day. In the final week, messages were pared down to one message per day.

Decisions about tailoring also need to be made at this stage. Tailoring uses information that an individual has provided about his or her circumstances to personalize the information that the participant receives to affect behavior change [26]. Tailoring increases the self-relevance of material, thereby increasing the likelihood that participants will be motivated to act on the material [27-30]. A recent review of computer-based health behavior change interventions suggests that the more dynamically the program is tailored, the stronger the efficacy data are likely to be [31]. Each point of tailoring results in more content needing to be written, however, so the number of points the program is tailored on should be weighed against the amount of time and budget the research team has. For researchers who are interested in concrete examples of how tailoring may be applicable to their program, Strecher and colleagues provide useful direction on their website [32].

In SMS Turkey, we chose to create different content paths for participants based on their progress along the quitting curve [16,21]. Previous data suggest that most smokers who relapse will do so within the first 2 days after quitting; at 7 days, the relapse curve begins to bottom out [25]. Therefore, we created paths for participants who were quit 2 days after quit day versus those who were smoking; and for those who were quit 7 days after quit day versus those who were smoking. The pathing could be done manually by research staff or automatically by the software program using a bidirectional data collection and response system. Due to time and financial constraints in the pilot study, research staff contacted participants at each time point and then manually pathed the participant to the applicable messages based on his or her response.

Researchers also will need to decide whether the program will have unique messages throughout the program or whether some messages will be repeated. This is particularly relevant if the intervention has a relapse path for those who are unsuccessful in enacting the targeted behavior change: do these participants repeat the previous content, or do they receive a different set of messages? In SMS Turkey, we decided to create unique messages such that participants would receive new messages across the program and paths.

A last decision is whether to enhance the text messaging program with other outreach efforts: this might include telephone calls from a counselor or an interactive website that reinforces important concepts. The interactive website option may be particularly important if the intervention is trying to demonstrate a new skill, such as how to use a condom in a sexual health intervention. The potential benefits need to be weighed against potential costs and how additional components may affect the potential scale-up of the program. Given the resource-limited setting in Turkey, we wanted to develop a stand-alone program that did not require the funding of in-person support. Because comparatively fewer people had access to the Internet than to text messaging [24], we also decided that an online add-on would not be useful. Therefore, we decided that the SMS Turkey program would rely solely on the text messaging content.

Drafting the Content

The content should be guided by experts in the field and grounded in a theoretical model. Ideally too, formative research should be conducted within the target population to identify any unique challenges they face, as well as to gather feedback about the content as it is being developed. In SMS Turkey, we conducted a quantitative survey among 150 adult smokers living in Ankara to better understand the technology use and smoking experiences of adults who were seriously thinking about quitting smoking in the next 30 days and living in Turkey's capital city [24].

Smoking cessation quitlines rely heavily on cognitive behavioral therapy [33,34]. Cognitive behavioral therapy focuses on altering the individual's way of thinking (cognitive processes) and acting (behavioral actions). Components include identifying new behaviors to be substituted for smoking-related activities, making a commitment to quitting, considering consequences of continued smoking, seeking information about smoking, controlling cues that may trigger the urge to smoke, and rewarding oneself for not smoking [35]. Additional program components include self-efficacy [36-39] (eg, reinforcing beliefs that the person is capable of quitting, while also teaching the person how to set goals to master quitting) and relapse prevention [40-45] (eg, helping participants identify potential triggers and coping strategies). Smoking cessation guidelines also recommend a combination of pharmacotherapy and behavioral strategies [33,46,47]. In SMS Turkey, the content provided cognitive and behavioral strategies. Instead of also providing pharmacotherapy, however, the program encouraged participants who smoked 10 or more cigarettes a day to separately talk with their health care provider.

Once the literature has been consulted and formative research conducted, a content map needs to be drafted. For SMS Turkey, five stages of messages emerged from the quantitative survey, a review of the literature, and the research team's clinical experiences [16,17,24,25,41,42,48-50]. The content aim for each stage was as follows. (1) *Prequit messages*: clarify reasons for quitting; understand smoking patterns and tempting situations, triggers, and urges; practice altering smoking patterns; learn about pharmacotherapy options and obtain prescription or medication; prepare to quit (eg, preparing environment, seeking support). (2) *Early quit*: identify common difficulties and discomforts (eg, what to expect and how to deal with them); emphasize use of coping strategies and to not over-think the decision to quit; emphasize self-identified reasons for and benefits of quitting; manage issues and side effects of medication; increasing pleasurable and other activities. (3) *Late quit*: become vigilant to recognize relapse in a different way (eg, situations, confidence); learn how to deal with issues that arise as a nonsmoker (eg, handle stress, moods); learn about weight gain and preventing weight gain; learn to reward and care for oneself; learn how to think of oneself as a nonsmoker; learn the benefits of a nonsmoking lifestyle. (4) *Slip/relapse*: know that many slip and how to get back on track; clarify reasons for quitting and recommit; learn about what didn't work and new strategies. (5) *Encouragement*: encourage those unable to quit with this attempt to try quitting again in the future; learn the norms of quitting and the quitting process.

As shown in Figure 1, all participants received the prequit and early quit messages. According to our tailoring plan, at day 2 postquit, participants were pathed to more early quit messages if they continued to be quit, or pathed into slip/relapse messages if they were smoking. Participants who spontaneously shared that they had tried but just cannot quit at day 2 were given the option to opt-out and end the program immediately, or to be pathed to the Encouragement arm. Content was tailored again at day 7. Participants who were quit at day 2 and continued to be quit at day 7 moved into late quit messages; participants who were quit at day 2 but were then smoking at day 7 were moved into slip/relapse and then moved into late quit; those who were smoking at both day 2 and day 7 were moved into encouragement messages. We included this last, abbreviated arm for people who continued to struggle with quitting in order to be respectful of the participant's individual quitting processes. If participants are finding it extremely difficult to quit, sending them text messages about how well they are doing with quitting, or even how they should "not quit quitting" might be perceived as shaming and may disenfranchise them during what might be a teachable moment. Instead, we tried to capitalize on this opportunity by helping participants frame their quitting attempt as a positive step forward in the sometimes long quitting process.

As shown in Table 1, we crafted eight different message types. Participants need to receive messages that provide specific and actionable strategies to cope with the discomfort of quitting. It also is important to educate smokers about the benefits of quitting to motivate them through difficult periods. We created an algorithm for each message type across each of the program stages (see Table 2). Armed with this guide for message creation, we produced a draft and then iterated the program content.

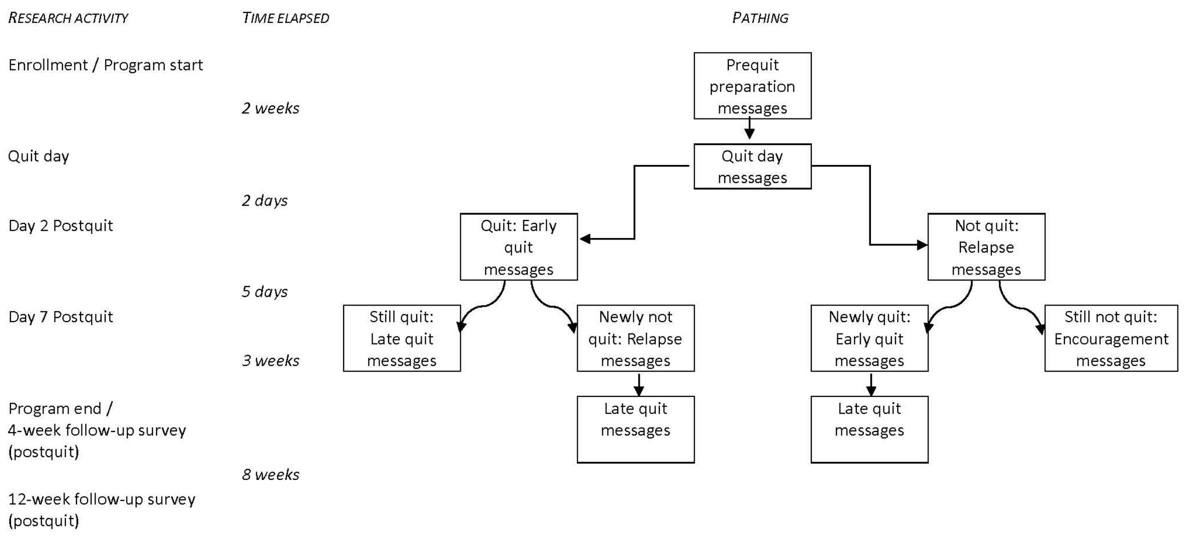
Table 1. Quit message types.

Message type	Description
Preparing to quit	Describes steps to take in preparing to quit smoking, including preparing oneself mentally and physically, and preparing one’s own and others’ environment
Benefits of quitting	Describes the health-related, social, and financial benefits of becoming a nonsmoker
Coping and coping strategies	Describes and encourages the effectiveness and use of cognitive and behavioral strategies to avoid smoking during a craving or impulse to smoke
Discomfort and difficulties	Discusses discomfort associated with the quitting process and how the participant may see his or her discomfort as normal and how to cope with such discomfort
Encouragement	Offers motivation and support to the participant to continue with quitting
Nicotine replacement therapy and pharmacotherapy	Encourages consideration and proper use of pharmacotherapy aids for use with the program in quitting smoking, such as nicotine replacement and other medications
Quitting skills	Teaches specific skills to aid in the quitting process
Relapse	Discusses the norms of slipping and how to get back on track; Clarifies reasons for quitting and to recommit; Teaches the participant to learn about what didn't work and new strategies

Table 2. Algorithm for cessation messages.

Message type	Number of messages per quitting stage					
	Prequit	Quit day and day 2	Early quit	Late quit	Relapse	Encouragement to try quitting again later
Preparing to quit	17	0	0	0	0	0
Benefits of quitting	4	2	3	19	3	2
Coping and coping strategies	14	4	11	6	7	1
Discomfort and difficulties	1	6	3	1	2	0
Encouragement	3	2	5	3	5	2
Nicotine replacement therapy and pharmacotherapy	3	2	2	1	2	1
Quitting skills	2	0	0	6	0	0
Relapse	0	0	0	0	5	0

Figure 1. SMS Turkey text messaging program flow.



We developed the content using the following guidelines.

Guideline 1

Write the messages to flow like a discussion across the day. For example, in the prequit stage, we encouraged people to start a smoking diary. The following are three messages that followed each other consecutively across the day:

Message 1: *When and why do you smoke? Start a smoking diary. Keep track of when you smoke, what you're doing (the activity), how you're feeling, and your craving (from 1 to 3).*

Message 2: *Write down a list of reasons why you want to quit smoking. Put it where you can see it.*

Message 3: *How is your smoking diary going? Put the "diary" paper on your cigarette pack with a rubber band. Every time you have a cigarette, fill out 1 line.*

The content of each new message builds on the previous message just as a conversation would.

Similarly, messages across days can refer to each other. For example, in the prequit stage, we tell participants to make a plan to reward themselves for quitting:

> *Write down a list of rewards for yourself—plan what you are going to do for yourself after you've quit for 1 week, for 2 weeks, for a whole month.*

Two days before their 1-week quit anniversary, participants receive the following message:

> *Have you been rewarding yourself for not smoking? Look back at your Rewards List—what did you write down on your list for your 1-week special reward? Too fun!*

Guideline 2

In the initial stages of message creation, the number of characters is not important. Certainly, brevity is the aim, but it is not useful to keep vigilant and stay below 160 characters (the limit for one text message), as the content itself is changing. This can be the last step before finalizing the content.

Guideline 3

Color-coding the messages in a spreadsheet program provides a useful visual cue to ensure that the message types are well spaced across days.

A Example of the Finished Product

To provide a concrete example of a finalized pool of text messages that are ready for subsequent testing, we show messages in the encouragement path in Table 3.

Table 3. Example of SMS Turkey content: the encouragement path.

Timing	Message text
Day 1a	Most smokers try to quit 6–7 times before they quit for good. Don't quit quitting!
Day 1b	It's a great thing that you've tried to quit smoking. You learned some things that you can apply to the next time you try to quit. What worked? What didn't?
Day 2a	Quitting smoking is the single most important step you can take to improve your health.
Day 2b	Medicines that treat craving can double your success. Try medicine next time you quit. If you used medicine, try a different one next time. Ask your doctor.
Day 3a	Smokers live an average of 7–12 years less than nonsmokers. Consider quitting again!
Day 3b	Whatever you decide about smoking, believe in yourself. You CAN quit smoking if you put your mind to it and have a plan for success.

Cultural and Cost Considerations

As mHealth research begins to extend beyond developed countries, cultural considerations in designing content and programming should be taken into account. The team leading the content development resided in the United States. The messages, therefore, were originally written in English. To ensure that the messages were culturally relevant and salient, a bilingual person translated the SMS Turkey messages from English into Turkish, paying particular attention to capturing cultural meaning and context. Then, a separate team member back-translated the messages into English. This typical procedure for translation and back-translation helps to ensure that the essence of the content is captured in the translation. The messages also were reviewed by students at Hacettepe University for understandability and credibility.

As noted above, a consideration during the protocol development stage is whether and how to reimburse costs that the participant may incur as a result of participating in the intervention. In Turkey, and most countries aside from the United States, receiving text messages is free. Thus, participants can take part in the program and receive as many text messages as possible without incurring a cost. As such, text messaging-based health programs are as free to participants as web-based health programs in countries such as Turkey.

Discussion

As text messaging becomes more common in research and intervention delivery, it will be important to document not only outcomes, but also content and protocol development procedures. This is particularly true as researchers struggle to see text messaging as the delivery mechanism, rather than the intervention itself [9]. For this reason, this paper focuses specifically on the development plan of SMS Turkey (called *Cebiniz birakin diyor* in Turkey).

Program development includes careful attention to not only the theoretical underpinnings of the behavior change strategy and associated content, but also the design, including how messages will be tailored, how participants may be pathed to different content based on their smoking status, and the number of messages participants will receive per day. The program must also attend to the technology use of the target population by including elements that fit that group. If the target population is not using text messaging, alternative delivery modes should be considered instead.

Effective health behavior change programs are guided by strong theoretical models [20,51,52]. Recent reviews have suggested, however, that many available text messaging-based behavior change programs may not be using theoretically based intervention strategies that have proven utility in producing behavior change [53]. Even though the technology is new, existing programs developed for more traditional environments can likely serve as an applicable guide when designing the content. In this study, we found that the theory and protocols underpinning telephone-based smoking cessation programs were useful guides for developing similar SMS Turkey program components. Indeed, this previous work in the field of smoking cessation informed the SMS Turkey content map, which guided the tone of the messages, and stage (eg, prequit versus early quit) and type of messages (eg, coping strategies, relapse) delivered over the course of the program and the participant's quit progress.

Although the intervention context behavior was smoking cessation, this development strategy is generalizable to other text messaging-based behavior change interventions. It should be noted that, as with many pilot projects, budget and timeline restrictions limited exploring all possible aspects of program development and tailoring. Certainly too, although we used theoretical models in developing the intervention, we based the program more generally on the key areas of behavior change found to be effective in quitline counseling. Our intent was to adapt successful quitline interventions to this new modality rather than start anew and develop a completely new theoretical perspective.

Acknowledgments

The project described was supported by Award Number R01TW007918 from the Fogarty International Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Fogarty International Center or the US National Institutes of Health.

Conflicts of Interest

None declared.

References

1. US Department of Health and Human Services. 2012 Jan 16. National Institutes of Health Research Portfolio Online Reporting Tools (RePORT) URL: http://projectreporter.nih.gov/reporter.cfm [accessed 2012-07-09] [WebCite Cache ID 69FdZoxRh]
2. Boushey CJ. National Cancer Institute. Improving Dietary Assessment Methods Using the Cell Phone and Digital Imaging [Grant no. 5U01CA130784-04] URL: http://projectreporter.nih.gov/project_info_description.cfm?aid=7920075 [accessed 2012-06-21] [WebCite Cache ID 68ak2zMo9]
3. Mundt JC. National Institute of Mental Health. Automated Interactive Text Messaging to Access Adolescent Depression [Grant no. 1R43MH086152-01A1] URL: http://projectreporter.nih.gov/project_info_description.cfm?projectnumber=1R43MH086152-01A1 [accessed 2012-06-21] [WebCite Cache ID 68akTXkDk]
4. Belzer ME. Eunice Kennedy Shriver National Institute of Child Health & Human Development. Adolescent Medicine Trials Network for HIV/AIDS Interve* [Grant no. 5U01HD040463-07] URL: http://projectreporter.nih.gov/project_info_description.cfm?aid=7195770 [accessed 2012-06-21] [WebCite Cache ID 68akQ0TSM]
5. Bull S. National Institute of Mental Health. Text Messaging for Abstinence and HIV Risk Prevention [Grant no. 1R21MH083318-01] URL: http://projectreporter.nih.gov/project_info_description.cfm?aid=7336217 [accessed 2012-06-21] [WebCite Cache ID 68akg42UC]
6. Olson A. Agency for Healthcare Research and Quality. Healthy Teens Txt Me: IT to Change Teen Health Risk Behaviors [Grant no. 1R21HS018214-01] URL: http://projectreporter.nih.gov/project_info_description.cfm?aid=7774003 [accessed 2012-06-21] [WebCite Cache ID 68aki3mTJ]
7. Cornelius JB. National Institute of Nursing Research. 2012 Jun 21. Feasibility of Delivering a Text Messaging Enhanced Adolescent HIV Curriculum [Grant no. 1R21NR011021-01] URL: http://projectreporter.nih.gov/project_info_description.cfm?icde=0&aid=7554587 [accessed 2012-07-06] [WebCite Cache ID 68x535Hsl]
8. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. J Telemed Telecare 2011;17(1):41-48. [doi: 10.1258/jtt.2010.100322] [Medline: 21097565]
9. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010 Apr;32(1):56-69. [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
10. Owens C, Farrand P, Darvill R, Emmens T, Hewis E, Aitken P. Involving service users in intervention design: a participatory approach to developing a text-messaging intervention to reduce repetition of self-harm. Health Expect 2011 Sep;14(3):285-295. [doi: 10.1111/j.1369-7625.2010.00623.x] [Medline: 20860777]
11. Whittaker R, Maddison R, McRobbie H, Bullen C, Denny S, Dorey E, et al. A multimedia mobile phone-based youth smoking cessation intervention: findings from content development and piloting studies. J Med Internet Res 2008;10(5):e49 [FREE Full text] [doi: 10.2196/jmir.1007] [Medline: 19033148]
12. Pleis JR, Ward BW, Lucas JW. Summary health statistics for U.S. adults: National Health Interview Survey, 2009. Vital Health Stat 10 2010 Dec(249):1-207. [Medline: 21905346]
13. Ministry of Health Turkey. 2010 Jul. Global Adult Tobacco Survey (GATS): Fact Sheet Turkey 2008 URL: http://www.who.int/tobacco/surveillance/en_tfi_gats_turkey_factsheet_2009.pdf [accessed 2012-01-16] [WebCite Cache ID 64k4Bwyu9]
14. Central Intelligence Agency. 2011. The World Factbook: Turkey URL: https://www.cia.gov/library/publications/the-world-factbook/geos/tu.html [accessed 2012-01-16] [WebCite Cache ID 64k59Q5rd]
15. Lenhart A, Ling R, Campbell S, Purcell K. Teens and mobile phones. Washington, DC: Pew Internet & American Life Center; 2010 Apr. URL: http://www.pewinternet.org/~/media//Files/Reports/2010/PIP-Teens-and-Mobile-2010-with-topline.pdf [accessed 2012-06-21] [WebCite Cache ID 68aePBmJX]
16. Zhu SH, Anderson CM, Tedeschi GJ, Rosbrook B, Johnson CE, Byrd M, et al. Evidence of real-world effectiveness of a telephone quitline for smokers. N Engl J Med 2002 Oct 3;347(14):1087-1093 [FREE Full text] [doi: 10.1056/NEJMsa020660] [Medline: 12362011]
17. Stead LF, Perera R, Lancaster T. Telephone counselling for smoking cessation. Cochrane Database Syst Rev 2006(3):CD002850. [doi: 10.1002/14651858.CD002850.pub2] [Medline: 16855992]
18. Stead LF, Perera R, Lancaster T. A systematic review of interventions for smokers who contact quitlines. Tob Control 2007 Dec;16 Suppl 1:i3-i8 [FREE Full text] [doi: 10.1136/tc.2006.019737] [Medline: 18048627]
19. Velicer WF, Prochaska JO, Fava JL, Laforge RG, Rossi JS. Interactive versus noninteractive interventions and dose-response relationships for stage-matched smoking cessation programs in a managed care setting. Health Psychol 1999 Jan;18(1):21-28. [Medline: 9925042] [doi: 10.1037/0278-6133.18.1.21]
20. Lichtenstein E, Zhu SH, Tedeschi GJ. Smoking cessation quitlines: an underrecognized intervention success story. Am Psychol 2010;65(4):252-261. [doi: 10.1037/a0018598] [Medline: 20455619]
21. Zhu SH, Tedeschi GJ, Anderson CM, Pierce JP. Telephone counseling for smoking cessation: what's in a call? J Couns Dev 1996;75(2):93-102. [doi: 10.1002/j.1556-6676.1996.tb02319.x]

22. Cummins SE, Bailey L, Campbell S, Koon-Kirby C, Zhu SH. Tobacco cessation quitlines in North America: a descriptive study. Tob Control 2007 Dec;16 Suppl 1:i9-15 [FREE Full text] [doi: 10.1136/tc.2007.020370] [Medline: 18048639]
23. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
24. Ybarra ML, Bağci Bosi AT, Bilir N, Holtrop JS, Korchmaros J, Emri S. Interest in technology-based and traditional smoking cessation programs among adult smokers in Ankara, Turkey. Tob Induc Dis 2011;9:10 [FREE Full text] [doi: 10.1186/1617-9625-9-10] [Medline: 21806793]
25. Zhu SH, Stretch V, Balabanis M, Rosbrook B, Sadler G, Pierce JP. Telephone counseling for smoking cessation: effects of single-session and multiple-session interventions. J Consult Clin Psychol 1996 Feb;64(1):202-211. [Medline: 8907100]
26. Kreuter MW, Skinner CS. Tailoring: what's in a name? Health Educ Res 2000 Feb;15(1):1-4 [FREE Full text] [Medline: 10788196]
27. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication With Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
28. Hines D, Fretz AC, Nollen NL. Regular and occasional smoking by college students: personality attributions of smokers and nonsmokers. Psychol Rep 1998 Dec;83(3 Pt 2):1299-1306. [Medline: 10079727]
29. Strecher VJ, Marcus A, Bishop K, Fleisher L, Stengle W, Levinson A, et al. A randomized controlled trial of multiple tailored messages for smoking cessation among callers to the cancer information service. J Health Commun 2005;10 Suppl 1:105-118. [doi: 10.1080/10810730500263810] [Medline: 16377603]
30. Strecher VJ, Kreuter M, Den Boer DJ, Kobrin S, Hospers HJ, Skinner CS. The effects of computer-tailored smoking cessation messages in family practice settings. J Fam Pract 1994 Sep;39(3):262-270. [Medline: 8077905]
31. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221. [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
32. University of Michigan Center for Health Communications Research. 2008. Tailoring URL: http://chcr.umich.edu/tailoring.php [accessed 2012-07-06] [WebCite Cache ID 68x5oPugm]
33. Hajek P, Stead LF, West R, Jarvis M, Lancaster T. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2009(1):CD003999. [doi: 10.1002/14651858.CD003999.pub3] [Medline: 19160228]
34. Ossip-Klein DJ, McIntosh S. Quitlines in North America: evidence base and applications. Am J Med Sci 2003 Oct;326(4):201-205. [Medline: 14557735] [doi: 10.1097/00000441-200310000-00010]
35. Rosen CS. Is the sequencing of change processes by stage consistent across health problems? A meta-analysis. Health Psychol 2000 Nov;19(6):593-604. [Medline: 11129363] [doi: 10.1037/0278-6133.19.6.593]
36. O'Hea EL, Boudreaux ED, Jeffries SK, Carmack Taylor CL, Scarinci IC, Brantley PJ. Stage of change movement across three health behaviors: the role of self-efficacy. Am J Health Promot 2004;19(2):94-102. [Medline: 15559709] [doi: 10.4278/0890-1171-19.2.94]
37. Fagan P, Eisenberg M, Frazier L, Stoddard AM, Avrunin JS, Sorensen G. Employed adolescents and beliefs about self-efficacy to avoid smoking. Addict Behav 2003 Jun;28(4):613-626. [Medline: 12726779] [doi: 10.1016/S0306-4603(02)00227-7]
38. Tucker JS, Ellickson PL, Klein DJ. Smoking cessation during the transition from adolescence to young adulthood. Nicotine Tob Res 2002 Aug;4(3):321-332. [doi: 10.1080/14622200210142698] [Medline: 12215241]
39. Borrelli B, Hogan JW, Bock B, Pinto B, Roberts M, Marcus B. Predictors of quitting and dropout among women in a clinic-based smoking cessation program. Psychol Addict Behav 2002 Mar;16(1):22-27. [Medline: 11934082] [doi: 10.1037//0893-164X.16.1.22]
40. Stöffelmayr B, Wadland WC, Pan W. An examination of the process of relapse prevention therapy designed to aid smoking cessation. Addict Behav 2003 Sep;28(7):1351-1358. [Medline: 12915175] [doi: 10.1016/S0306-4603(02)00250-2]
41. Wadland WC, Stöffelmayr B, Berger E, Crombach A, Ives K. Enhancing smoking cessation rates in primary care. J Fam Pract 1999 Sep;48(9):711-718. [Medline: 10498078]
42. Wadland WC, Soffelmayr B, Ives K. Enhancing smoking cessation of low-income smokers in managed care. J Fam Pract 2001 Feb;50(2):138-144. [Medline: 11219562]
43. Ramsay J, Hoffmann A. Smoking cessation and relapse prevention among undergraduate students: a pilot demonstration project. J Am Coll Health 2004;53(1):11-18. [doi: 10.3200/JACH.53.1.11-18] [Medline: 15266725]
44. Larimer ME, Palmer RS, Marlatt GA. Relapse prevention. An overview of Marlatt's cognitive-behavioral model. Alcohol Res Health 1999;23(2):151-160. [Medline: 10890810]
45. Marlatt GA, George WH. Relapse prevention: introduction and overview of the model. Br J Addict 1984 Sep;79(3):261-273. [Medline: 6595020] [doi: 10.1111/j.1360-0443.1984.tb00274.x]
46. Lancaster T, Hajek P, Stead LF, West R, Jarvis MJ. Prevention of relapse after quitting smoking: a systematic review of trials. Arch Intern Med 2006 Apr 24;166(8):828-835. [doi: 10.1001/archinte.166.8.828] [Medline: 16636207]
47. Silagy C, Lancaster T, Stead L, Mant D, Fowler G. Nicotine replacement therapy for smoking cessation. Cochrane Database Syst Rev 2004(3):CD000146. [doi: 10.1002/14651858.CD000146.pub2] [Medline: 15266423]

48. Brown R. Comorbidity treatment: skills training for coping with depression and negative moods. In: Abrams D, Niaura RS, Brown R, Emmons KM, Goldstein MG, Monti PM, editors. The Tobacco Dependence Handbook: A Guide to Best Practices. New York, NY: Guilford Press; 2003.
49. Holtrop JS, Stommel M, Corser W, Holmes-Rovner M. Predictors of smoking cessation and relapse after hospitalization for acute coronary syndrome. J Hosp Med 2009 Mar;4(3):E3-E9. [doi: 10.1002/jhm.415] [Medline: 19301384]
50. Fiore MC, Jaen CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, Henderson, PN, et al. US Department of Health and Human Services, Public Health Service. 2008 May. Clinical Practice Guideline: Treating Tobacco Use and Dependence: 2008 Update URL: http://www.surgeongeneral.gov/tobacco/treating_tobacco_use08.pdf [accessed 2012-01-16] [WebCite Cache ID 64k5nbAPi]
51. Johnson WD, Diaz RM, Flanders WD, Goodman M, Hill AN, Holtgrave D, et al. Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men. Cochrane Database Syst Rev 2008(3):CD001230. [doi: 10.1002/14651858.CD001230.pub2] [Medline: 18646068]
52. Gaglio B, Smith TL, Estabrooks PA, Ritzwoller DP, Ferro EF, Glasgow RE. Using theory and technology to design a practical and generalizable smoking reduction intervention. Health Promot Pract 2010 Sep;11(5):675-684. [doi: 10.1177/1524839908324778] [Medline: 19116418]
53. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285. [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]

Abbreviations

PI: principal investigator

Edited by G Eysenbach; submitted 25.01.12; peer-reviewed by R Whittaker; accepted 10.03.12; published 24.07.12.

Please cite as:
Ybarra ML, Holtrop JS, Bağci Bosi AT, Emri S
Design Considerations in Developing a Text Messaging Program Aimed at Smoking Cessation
J Med Internet Res 2012;14(4):e103
URL: http://www.jmir.org/2012/4/e103/
doi:10.2196/jmir.2061
PMID:22832182

Original Paper

Improving Diabetes Care for Young People With Type 1 Diabetes Through Visual Learning on Mobile Phones: Mixed-Methods Study

Dag Helge Frøisland[1,2], MD; Eirik Årsand[3,4], PhD; Finn Skårderud[1,5], MD, PhD

[1]Research Centre for Child and Youth Competence Development, Lillehammer University College, Lillehammer, Norway
[2]Faculty of Medicine, University of Oslo, Oslo, Norway
[3]Norwegian Centre for Integrated Care and Telemedicine (NST), University Hospital of North Norway, Tromsø, Norway
[4]Department of Computer Science, University of Tromsø, Tromsø, Norway
[5]Department of Special Needs Education, University of Oslo, Oslo, Norway

Corresponding Author:
Eirik Årsand, PhD
Norwegian Centre for Integrated Care and Telemedicine (NST)
University Hospital of North Norway
PO Box 35
Tromsø, 9038
Norway
Phone: 47 992 43 592
Fax: 47 77 75 40 98
Email: eirik.arsand@telemed.no

Abstract

Background: Only 17% of Norwegian children and adolescents with diabetes achieve international treatment goals measured by glycated hemoglobin (HbA_{1c}). Classic patient–physician consultations seem to be poorly adapted to young children. New strategies that are better attuned to young people to improve support of adolescents' self-management of diabetes need to be tested and evaluated.

Objective: (1) To explore how applications for mobile phones can be used in follow-up of adolescents with type 1 diabetes, and (2) to use the findings to guide further development of the applications and as a basis for future studies.

Method: We pilot tested two mobile phone applications: (1) an application that contained a picture-based diabetes diary to record physical activity and photos taken with the phone camera of food eaten, where the phone also communicated with the glucometer by Bluetooth technology to capture blood glucose values, and (2) a Web-based, password-secured and encrypted short message service (SMS), based on access using login passwords received via SMS to be used by participants to send messages to their providers when they faced obstacles in everyday life, and to send educational messages to the participants. At the end of the 3-month pilot study, 12 participants (7 girls and 5 boys) aged 13–19 years completed semistructured interviews. The participants had a mean HbA_{1c} value of 8.3 (SD 0.3), mean age of 16.2 (SD 1.7) years, mean body mass index of 23.3 (SD 3.2) kg/m^2, and mean diabetes duration of 7.5 (SD 4.6) years. We applied three additional measurements: change in metabolic control as measured by HbA_{1c}, the System Usability Scale, and diabetes knowledge.

Results: From the interviews, three main categories emerged: visualization, access, and software changes. Participants appreciated the picture-based diary more than the SMS solution. Visualization of cornerstones in diabetes self-care (ie, diet, insulin dosage, physical activity, and pre- and postprandial glucose measurements all transformed into one picture) in the mobile diary was found to be an important educational tool through reflections in action. This led to a change in participants' applied knowledge about the management of their disease. Additional measurements supplemented and supported the qualitative findings. However, changes in HbA_{1c} and participants' theoretical knowledge as tested by a 27-item questionnaire, based on a national health informatics' diabetes quiz, before and after the intervention were not statistically significant (P = .38 and P = .82, respectively, paired-samples *t* test). Participants suggested additional functionality, and we will implement this in the design of the next software generation.

Conclusion: Participants reported an increased understanding of applied knowledge, which seem to positively affect diabetes self-care. Visual impressions seem well adapted to the maturation of the adolescent brain, facilitating the link between theoretical knowledge and executive functions. SMS gave the adolescents a feeling of increased access and security. Participants gave valuable input for further development of these applications.

(J Med Internet Res 2012;14(4):e111) doi:10.2196/jmir.2155

KEYWORDS

Diabetes; adolescents; user-centered design; education; eHealth; mHealth; mobile phones; short message service; SMS; qualitative research; triangulation of methods

Introduction

Diabetes

Type 1 diabetes is one of the most common chronic illnesses of childhood. The incidence of diabetes in Norway is one of the highest in the world, reported as 32 new cases per 100,000 person-years [1]. Type 1 diabetes often requires rigorous daily routines and a high level of self-management, which is seen as a complex task. Treatment goals might be achieved through frequent *blood glucose measurements* as well as through *insulin medication* tailored to *food intake* and *physical activity*. These elements are often described as the cornerstones of diabetes treatment. A significant increase in blood glucose measurement was found in a study using reminders and cueing, social media communication, and gamification [2]. However, in a review of telemedicine, Farmer et al concluded that intervention “with or without telemedicine is only likely to be helpful when test results are linked to educational or behavioral advice and changes in clinical management” [3]. Clinical guidelines highlight the importance of education; nevertheless, this seems to be a missing functionality in most telemedicine interventions [4].

Adolescence and Diabetes

Adolescence as a transitional phase in human life challenges both young people and their caregivers [5]. Physicians and caregivers as well as adolescents with diabetes often experience frustration and powerlessness when the adolescents repeatedly return to their clinics with poor metabolic control.

Only 17% of Norwegian children and adolescents with diabetes achieve international treatment goals measured by glycated hemoglobin (HbA_{1c}) [6]. As researchers, we need to uncover the reasons behind such treatment results and explore ways to improve them. One approach is to develop developmentally appropriate means of supporting adolescents’ self-management of diabetes. Consequently, we developed two mobile phone applications as tools for diabetes self-management. The research described in this paper is an evaluation of these tools.

Education

The development of competence (ie, knowledge and skills) among people living with diabetes is integral to their effective management of the disease. In modern health care, providers are advised to use new technology and educational models based on learning theory, health models, and instructional design theory when delivering patient education [7]. Diabetes self-care is highly dependent on coordination of thoughts and behavior—that is, executive functions. Skills necessary for such coordination are selective attention, decision making, voluntary response, inhibition, and working memory [8]. Motor and sensory brain areas mature first; the primary visual cortex matures early, while areas involved in executive functions mature later. These development patterns might facilitate the application of modern technology using visual imaging in this particular patient group [9]. Visual impressions have been shown to improve understanding and self-management in patients with chronic diseases [10,11].

Information and Communication Technology

Information and communication technology (ICT) has developed rapidly over the last decades. Reports from Statistics Norway show that among adolescents aged 13-19 years, more than 90% use the Internet for more than 2 hours every day; 95% have their own mobile phone, and more than half of them accessed the Web through their mobile phone daily [12]. ICT is gradually being applied in health care and eases the flow of information between providers and their patients [13-15]. These technologies include Internet, email, and mobile phone applications, and are often referred to as electronic health or eHealth [16].

The use of ICT to facilitate health care has traditionally been dominated by PC-based technology [14]. Many ICT studies show overall positive results and indicate that ICT-based interventions improve health care utilization, health behaviors, attitudes, skills, and knowledge [14]. Studies involving mobile phones used by adults have been published [15,17-20], but there are few reports of the use of mobile phone-based tools among children and adolescents with type 1 diabetes.

When a new ICT system is developed, user-involved design of patient-operated systems is advocated to promote useful applications [21]. Based on this philosophy, the Norwegian Centre for Integrated Care and Telemedicine (NST) has developed several mobile applications based on user-participatory design processes.

Aim

The aim of this study was to evaluate adolescent patients’ experiences with two different mobile phone applications used for diabetes care. We wanted to examine whether an intervention using information and communication tools could affect disease management measured by metabolic control and through qualitative methods. We also wanted the adolescents to use their experiences to guide the product developers by giving advice on further improvements. Finally, we wanted to propose new hypotheses for further research on approaches that diabetes teams might apply to facilitate the use of young people’s competencies.

Methods

Design

This study used triangulation of methods to provide details about the phenomenon studied that would not be available with the use of one method alone. Interviews were recorded. Field notes were taken systematically during the semistructured interviews and used as an additional data source. Data on metabolic control measured by HbA_{1c} were collected before and at the end of the intervention. HbA_{1c} was analyzed at the local hospitals using Bayer DCA 2000 (Tarrytown, NY, USA; normal reference range 3.4%-6.1%). Usability data were collected using the System Usability Scale (SUS), a questionnaire related to human-computer interaction consisting of a simple 10-item scale based on a 5-point Likert scale [22]. A SUS score above 58 is regarded as above average, and a SUS score above 80 is regarded as high and a score where participants are likely to recommend the product to friends [23]. Additional data were collected before and at the end of the intervention period through a 27-item questionnaire based on the Norwegian National Health Informatics' diabetes quiz [24].

Intervention

The intervention period was 3 months. This included an instruction day when the participants were introduced to an HTC Touch 2 mobile phone (HTC Corporation, Bellevue, WA, USA), on loan from NST, and the two novel diabetes software applications: a mobile phone-based diabetes diary called Diamob and the Diabetes Message System, a short message service (SMS) based on the Secure Health Dialogue system [25]. Participants could use the Diamob application as much as they wanted, but 2 periods of recording data, each lasting 3 days, before the consultation were mandatory. Participants came for a consultation with the research team (midway) to discuss use of the applications and to participate in a reflection-in-action talk about the mandatory 3-day recordings in the Diamob application. When the participants browsed through the pictures, self-reflection was guided by emphasis on recording images of food, level of activity, insulin dosages, and pre- and postprandial glucose measurements. Two physicians from the outpatient clinics were responsible for the Diabetes Message System responses. At the end of the 3-month period, the participants met with the research team for a semistructured interview.

Technological Applications

Application 1

The prototype Few Touch application [17] was optimized by NST to include a camera-based dietary capture and a feedback component. This modified version of the Few Touch application is referred to as the Diamob application and targets communication between the patient and the health care team about carbohydrate evaluation and insulin dosages. The application was used by adolescents with type 1 diabetes to document the food they ate. They had to choose one of four pictograms describing the physical activity in which they planned to participate or had already been engaged in (Figure 1).

The page for the first step guided the participants to a screen prompting them for the insulin dosage suitable for the food they planned to eat. The next step was to photograph the food using the mobile phone's camera. The phone communicated with the glucometer by Bluetooth technology to capture blood glucose values (Figure 2). One and a half hours after the meal, the phone reminded the users to measure their postprandial blood glucose.

The picture produced by the Diamob application incorporated relevant pre- and postprandial blood glucose values, insulin dosage given, and information about the participant's physical activity (Figure 3).

Figure 1. The first step in the Diamob application, consisting of pictograms specifying planned physical activity. This leads the user to the next step.

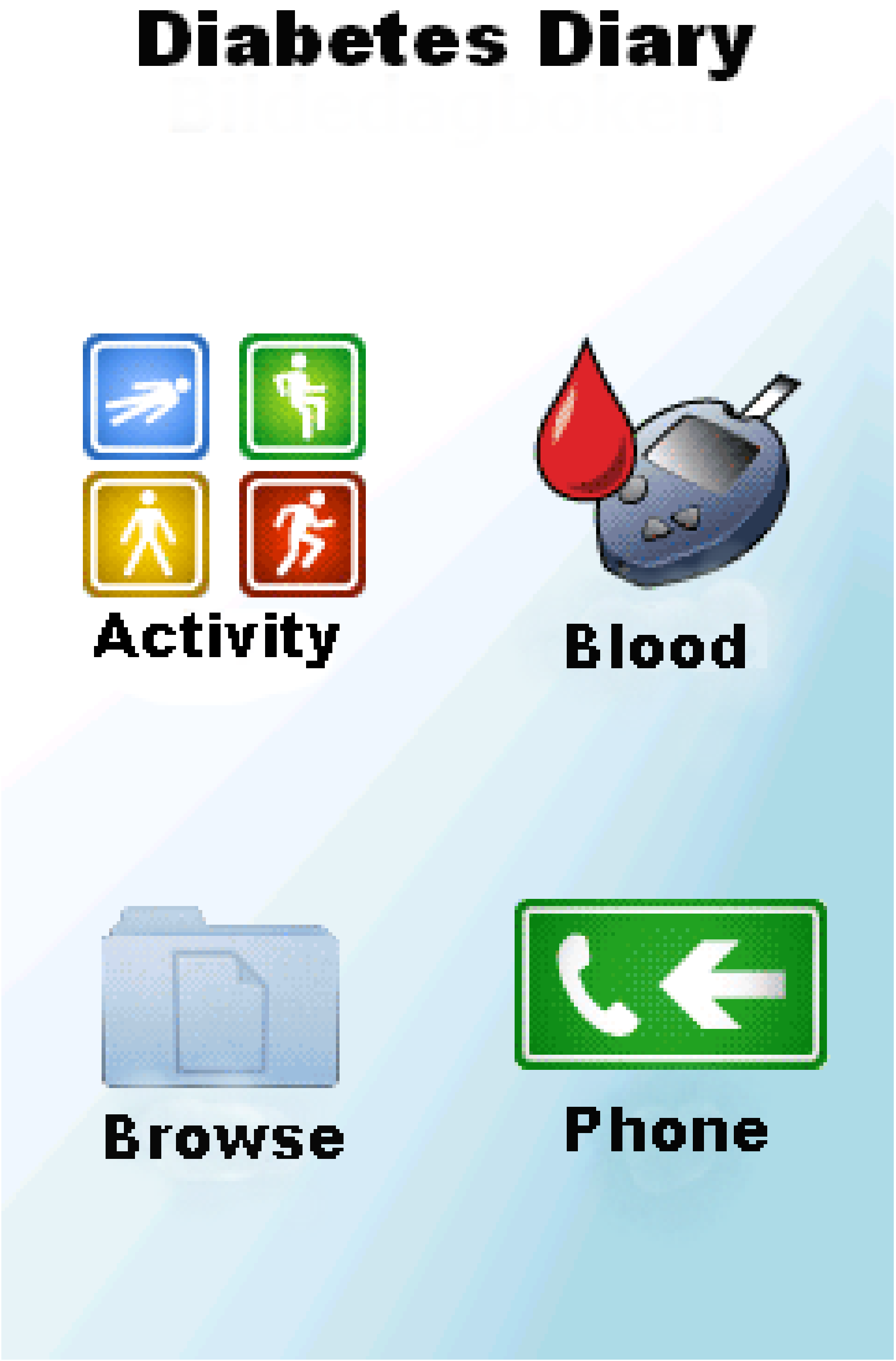
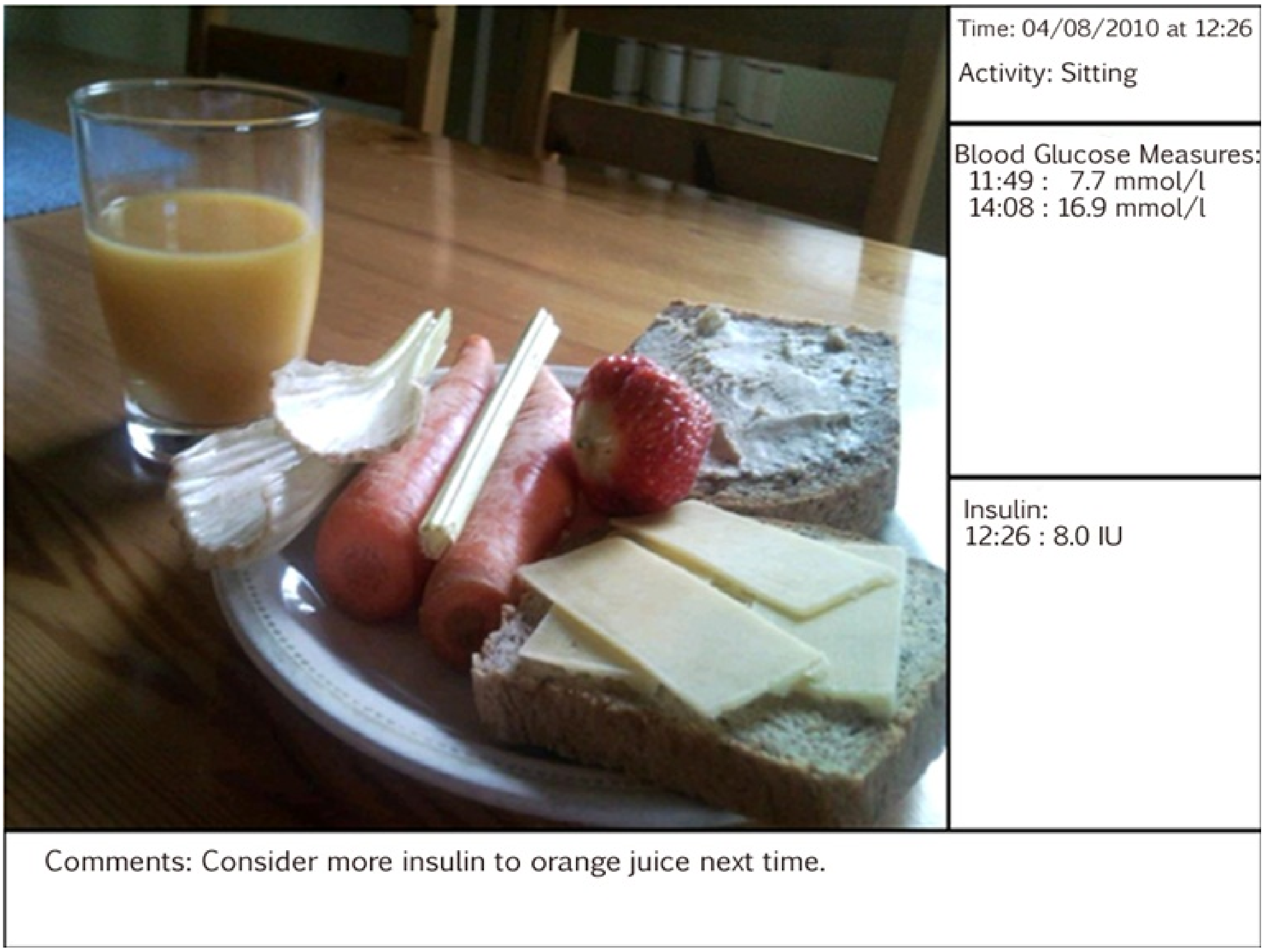


Figure 2. The Bluetooth transmitter (left) sends glucose measurements to the phone as soon as the test strip is extracted from the glucometer.

Figure 3. Example of a picture produced by the Diamob application. Pictures were available for users to browse through on the phone and could also be uploaded to computers and used during consultations.



Time: 04/08/2010 at 12:26

Activity: Sitting

Blood Glucose Measures:
11:49 : 7.7 mmol/l
14:08 : 16.9 mmol/l

Insulin:
12:26 : 8.0 IU

Comments: Consider more insulin to orange juice next time.

Application 2

Due to strict Norwegian laws regarding security and privacy of health communication, we created a functional Web-based encrypted SMS based on the Secure Health Dialogue SMS with the help of a commercial software company (WTW AS, Tiller, Norway) [25]. This Web-based SMS system is referred to as the Diabetes Message System. It is based on access using login passwords received via SMS to ensure that health information is not compromised. We invited the participants to use this application to send messages to their providers when they faced obstacles in everyday life. The Diabetes Message System application was also tailored to send educational messages to the participants. An example of such messages is “Reasons for high blood sugar: Not enough insulin, intake of food rich in carbohydrates, infections, fever.”

Sample

We recruited a convenience sample of adolescents aged between 13 and 19 years from two pediatric clinics in Eastern Norway, Innlandet Hospital Trust. Inclusion criteria were a diagnosis of type 1 diabetes at least 1 year prior to the start of the study and current HbA_{1c} of less than 10.0%. We set the upper HbA_{1c} limit to avoid bias from potential psychological maladjustment associated with poor metabolic control. We enlisted 12 participants: 7 girls and 5 boys. Of the 12 interviewees, 1 withdrew from the use of the application halfway through the intervention for personal reasons. We interviewed all 12 participants at the end of the study period.

Interviews

We developed a semistructured interview guide to elicit responses on the topics described in the aims of the study. The guide included questions to identify different experiences with the implemented technology. The interviews took place in a meeting room at the outpatient clinics. They lasted between 45 and 90 minutes. Recordings were later transcribed, including notes on nonverbal aspects of the communication, such as pauses and laughter.

Analytical Procedures

The qualitative data analysis was based on transcribed interviews, supported by the interview moderator’s field notes. The analytic approach was based on qualitative description [26], influenced by phenomenology and hermeneutics [27]. The analysis entailed an inductive process, focusing on the participants’ experience with the mobile device in an attempt to make sense of their personal experiences with the device. The understanding of another person’s situated experience of the world requires interpretation, thus making hermeneutics an important part of the approach.

According to a bottom-up principle, all the interviews were carefully read and reread with the aim of discriminating units of meaning from the transcripts. Such units were assigned codes. Elements of texts with the same codes were identified and extracted to code sheets. Units of meaning were first identified, and condensed units of meaning were then developed into categories. Finally, the codes and categories were organized hierarchically.

To increase validity and reduce possibilities of idiosyncratic coding, two of the authors conducted the analysis separately [28]. The categories were discussed and negotiated until consensus was achieved. In the final stage of the process, these concepts were tested by comparing them with the transcripts, according to a top-down principle, to ensure validation of the extraction. Through this process, the final categories emerged. To ensure that the ensuing concepts were well represented in the data, the research team discussed the developing analysis to add credibility to the study and ensure agreement on the main themes.

Statistical Analyses

Results are presented as means with 1 standard deviation (SD). Paired-samples *t* test analyses were used to compare pre- versus postinterventional data. Significance was defined as $P < .05$. We used SPSS version 18.0 (IBM Corporation, Somers, NY, USA) for analyses.

Ethical Considerations

The adolescents and their parents gave written consent for the study according to Norwegian requirements. The study was approved by the regional committee for medical research ethics (ref. 2009/773b). As compensation, participants could ask for a refund of the costs of their Internet subscription to cover expenses related to accessing the Web-based SMS application

Results

A total of 7 girls and 5 boys participated. Their mean HbA_{1c} was 8.3 (SD 0.3) before the intervention. Mean age was 16.2 (SD 1.7) years, mean disease duration was 7.5 (SD 4.6) years, and mean body mass index was 23.3 (SD 3.2) kg/m^2. All the participants used insulin pumps, compared with 75% in the background population. All participants used a mobile phone daily, and 8 (67%) reported accessing the Internet via a mobile phone daily. The participants were asked to commit to completing 2 sets of diabetes diary records, each covering a continuous period of 3 days, during the intervention. This would yield approximately 24 pictures for each participant (288 pictures). During the study, 691 pictures were downloaded (mean 50, minimum 25, maximum 94).

The overall result from the qualitative data demonstrated that the adolescents found both mobile applications useful as a support to their diabetes self-management. They appreciated the picture based diabetes application more than the SMS solution. Table 1 give a brief overview of the process of analyzing the interviews, from which three main categories emerged: visualization, access, and software changes. The quotes used to illustrate these themes use fictitious names for the participants.

Table 1. Brief overview of the process of analyzing the interviews.

Theme	Codes and condensed meaning	Category and concept final theme	Findings and hypothesis
Functionality of mobile phone diary Communication facility	To see the coherence of treatment Communication through pictures To visually identify unhealthy food	Visualization	Diamob seems to be more an educational tool to enhance understanding of diabetes self-care than just a communication tool. This might be due to brain development among children and adolescents, and because practical executive competence is facilitated by visual memory rather than theoretical facts.
SMS[a] Educational messages	Closeness Empowerment Easy access to information	Access	SMS is a useful and preferred tool for adolescents to enhance contact with the health care system, and a nonencrypted version is favored.
Technical advice and challenges	Technical challenges Future development	Software changes	Changes to the two applications need to be made to increase functionality.

[a] Short message service.

Visualization

All the adolescent participants used the verb “to see” in relation to the first application’s functionality. They reported a better visual understanding of the cornerstones of diabetes self-management: food intake, insulin dosage, physical activity, and blood glucose measurements. The participants reported that the pictures of food they had consumed, integrating pre- and postprandial glucose measurements and insulin dosages as well as information on physical activity, gave them a visual and tangible understanding of how physical activity, food intake, and the insulin dosage interact and affect postprandial glucose measurements.

> *Before, I really thought that the blood sugar was one thing and giving insulin was one thing and eating was one thing, but now I see more all three of them as a whole, that they all belong together. Because if not all three of them come together, I feel it’s like I miss a part of the puzzle.* [Tom]
>
> *It is just to browse back in the picture diary and look at how much [insulin] I actually needed to the food I had eaten, that is an advantage...I have learned to*

> *think about what to eat and the value of measuring my blood sugar before I eat and think about the insulin dosage in relation to these factors, then to measure again 1.5 hours later in order to evaluate if the dosage of insulin was correct.* [Ann]

> *The Diamob application helped the most...I became tougher in taking insulin doses. Because I saw how the glucose measurements changed and the value of giving enough insulin.* [Erik]

The participants also said that they preferred the mobile phone-based diabetes diary to paper-based diaries, and that it provided an incentive for communication about their diabetes self-management both with their parents and with their health care providers.

> *I think it is a lot easier to understand and to have it explained when I can see things.* [Oda]

> *When I met Mom or Dad I could show them what I had done during the day instead of writing. I liked that a lot.* [Beth]

The adolescents also reported that seeing the pictures of their own food in the Diamob application gave them a visual understanding of their own unhealthy diet.

> *I just photographed the food I usually eat, but I thought during the process that I should apply more healthy eating habits, because I saw I had a lot of unhealthy canteen food in school.* [Erik]

Access

The participants experienced Web-based SMS as a positive instrument for bidirectional contact with health care providers. It provided a safety net that gave them a sense of protection because it made it easy for them to access their physician with questions and concerns.

> *I liked the project and the follow-up. I could send an SMS whenever I wanted. I got an answer within half an hour. I especially liked the SMS—in the Netherlands, where I lived prior to this, I knew I could call, but I like the message system a lot better.* [Jan]

> *The fact that you have someone to support you—someone who knows the subject, and if you get into difficulties you can get an answer—it gives a certain feeling of security.* [Oda]

All participants greatly appreciated the possibility of contacting their health care provider and receiving an immediate response. They indicated that it gave them a feeling of closeness to the health care practitioner.

> *It is usually not easy to get in touch with your doctor...it was nice, because usually you call in and you are told he is for lunch or are busy and then it ends up you don't do anything.* [Tom]

The participants also reported feeling empowered in that they could access the health care practitioner so readily.

> *It has been pretty good to know that if I have an issue, then I can just send a message...Instead of calling Mom or Dad and ask them to call [the physician], and when they have the answer it might be an answer to something I was not wondering about.* [Beth]

The participants reported that they appreciated having access to information through the mobile phone. They described the educational text messages as useful in increasing their understanding of diabetes self-management. However, they also stated that they appreciated simple and practical self-management advice more than large amounts of information that was not relevant to their immediate situation.

> *It is more about those messages and the information. It has been practical advice, easy to understand, simple facts that are very nice to know. It is better to have it in such small portions instead of reading a lot of information, then everything is poorly read and poorly understood. I liked the way the information was given.* [Oda]

> *...it is easier to access the information in this way, than listening to a doctor who talks about it for 5 hours.* [Kristin]

Despite their appreciation for the accessibility of the device, several participants indicated that the Web-based SMS system was unduly cumbersome.

> *But it was a bit cumbersome to have to remember a code [to get access].* [Jan]

> *But what is cumbersome is that you have to access that Internet browser on the mobile. I would prefer to send normal SMS on the phone...that would make it even easier if you could access it using the usual SMS [on the phone].* [David]

Software Changes

Another category of data was the software changes that the participants recommended. Although the participants described the applications as highly usable in general, they also identified particular problems that required resolution. They reported technical problems that we attributed to the immaturity of the applications and the chosen mobile platform. Some of the participants explained that they experienced problems and found it troublesome to use the applications.

> *The Diamob app didn't work at the end of the project. The glucometer with Bluetooth worked, but batteries ran out of power quickly.* [Jan]

> *Yes, the project was tailored to me, but it could have been better on the glucose transmission [from the glucometer to the phone] because it didn't work all the time.* [David]

The participants agreed that the applications were highly usable, but requested the possibility of downloading the applications to their own mobile phones in the future. With 1 exception, the participants indicated that if the applications were available for downloading to their own phones, they would continue to use the applications. They contributed technical advice on how the applications could be improved.

> *It would give a better overview as a matter of fact, when you think about it, when you have to go from picture to picture, but if you had all 3 days in 1 page,*

> *as a complete overview of 1 day, then I think it would be better.* [Oda]
>
> *I didn't like the diagram thing [in the Diamob apps]. It was a mess, and I didn't like that it could not be tailored to each patient...And for instance I think it is stupid that it [the pictograms for activity] only marks sitting, lying, standing, and training, but maybe you don't do the physical activity, and then, after an hour, it is time to go for a run.* [Eva]

Additional Measurements

Qualitative data were supported by glycemic control, SUS scores, and knowledge tests. Metabolic control had improved (ie, lowered HbA_{1c}) in 7 participants at the end of the study, but was unchanged in 2 participants. During the study period, 3 of the participants increased their HbA_{1c} (Table 2). Statistically, we found no significant changes in HbA_{1c} over the course of the intervention (Table 3). None of the participants who completed the 3-month intervention had severe hypoglycemia or severe ketoacidosis or were hospitalized during the intervention period.

The mean score on the SUS was 73 (SD 22) (Table 2); 10 participants had a high SUS score: mean 81 (SD 10). These 10 reported positive experiences with the 2 mobile applications. However, 2 participants had a considerably lower SUS score than the others (each scoring 30).

All the participants had relatively high scores on the theoretical diabetes knowledge tests, with a mean of 22 (SD 2) (maximum score 27; Table 2). Scores were nearly the same prior to the project and at the end; no statistical differences were found (Table 3).

Table 2. Participants' glycated hemoglobin (HbA_{1c}) and knowledge test score (maximum score 27), before and after the intervention, and System Usability Scale (SUS) score (maximum score 100) after the intervention.

Participant	HbA_{1c} %		Knowledge test score		SUS score
	Before	After	Before	After	
Jan	8.6	9.0	18	20	67.5
David	6.6	6.4	22	21	87.5
Eva	8.6	7.6	11	NA[a]	30
Beth	8.3	8.3	23	23	87.5
Kristin	8.0	8.4	21	22	65
Oda	7.4	7.2	23	23	87.5
Tom	9.4	9.1	21	24	30
Erik	9.2	8.3	23	24	92.5
Ann	8.6	8.6	23	27	80
Emma	7.1	7.0	25	21	95
Ingrid	8.9	8.5	18	18	75
Trond[b]	9.0	9.7	24	20	77.5

[a] Eva did not take the postintervention knowledge test, and this is marked as not available.

[b] Trond did not complete the whole intervention for personal reasons unrelated to the study and was interviewed at the end.

Table 3. Glycated hemoglobin (HbA_{1c}) values, knowledge test scores, and System Usability Scale score before and after the intervention.

Measure	Mean	SD	No.	*P* value (paired-samples *t* test)
HbA 1c				.38
Before	8.3	0.9	12	
After	8.1	0.9	12	
Knowledge test score				.82
Before	22.0	2.3	12	
After	22.1	2.5	11	
System Usability Scale score				
All participants	73.0	22.1	12	
High scorers[a]	81.5	10.2	10	

[a] Mean score excluding the 2 participants who scored low on the System Usability Scale (score of 30).

Discussion

The most surprising finding was the reports of a new visual understanding of diabetes treatment. Visual impression and consultations based on reflection in action seemed to have a startling effect on self-reported perception of the coherence of the cornerstones of diabetes treatment: diet, insulin dosage, physical activity, and blood glucose measurements. Most of the adolescents considered the mobile picture diary to be superior to paper-based systems and preferred the SMS application as a convenient means of communicating with their health care providers.

Limitations

A limitation of the study is that our sample might be biased. Those who volunteered to participate might have been more comfortable with the technology than those who did not. The intervention period was short, and some technical problems arose. Although the sample is small, statistical analyses were performed on some of the variables. Low numbers of participants and the short intervention period reduce the potential for proving statistical significance. A control group would have strengthened this potential. The technical solutions had limitations in regard to both applications and mobile phones. The Diamob application was only compatible with the HTC 2 mobile phone, and participants disliked that phone's touch-sensitive user interface. Furthermore, they experienced problems with Bluetooth transfer of the glucose measurements.

Visualization

The last 20 years of diabetes education have reflected an increased emphasis on integrated educational strategies and collaboration with the patient [29,30]. Future mHealth applications should implement the educational benefits of visualization with the report of factors that enhance the frequency of glucose measurements [2].

Despite several large multicenter randomized studies on diabetes treatment, we have limited knowledge about why adherence to diabetes treatment, especially among young people, is so difficult. Applications like the Diamob application may help to bridge the gap between theoretical knowledge and lack of execution of treatment competencies in different contexts in patients' everyday life. The use of meal images and the reflection-in-action method contributes to the change in applied knowledge and ease of communication during consultation [31]. Health care personnel need to recognize patients' existing competencies, experiences, and preferences to be able to deliver education and health services tailored to the people they serve. Simple visual tools designed by young people in their own personal settings seem important to improve patients' comprehension, recall, and adherence [10,11], which our study also supports. In their review article of visual impression in patient education, Houts et al [10] asserted that pictures with cultural relevance that were designed by patients themselves are suitable for increasing understanding and adherence to treatment [10]. The Diamob application creates such personal pictures, and the widespread use of mobile phones among adolescents in Norway should encourage the use of such applications in this population [12].

To perform well in self-care, patients depend on well-developed executive functions: the ability to store important information, to keep in mind a plan to carry out in the future, and to inhibit impulses. Work has associated these behaviors with the frontal lobes, with maturation late in the adolescent period [8]. Magnetic resonance imaging studies have shown that maturation of the visual cortex and the subsequent pruning of neural connections happen early in childhood, compared with the frontal parts of the brain, where centers for cognitive functions such as planning and advanced thinking are located [8,9].

It has been shown that children with type 1 diabetes have slightly lower cognitive function than healthy children, and that they perform less well than controls on global measures of both intelligence and neuropsychological skills such as attention and executive functions [32]. The lack of maturation and pruning in the adolescent brains might explain the mismatch between theoretical competence and execution of practical skills found in our study. Professionals educating young people with diabetes need to be aware that they are less able than adults to link theoretical knowledge with cognitive and executive functions.

Access

Young people rarely approach health care providers by themselves, and often communicate to physicians and other practitioners through their parents. It has been claimed that questions considered difficult to ask in a face-to-face dialogue, or that should not be postponed until the next clinical appointment, are facilitated by text messaging [33]. As in the Sweet Talk study [34], the participants in our study reported using the SMS solution relatively rarely; nevertheless, they appreciated its function as a safety net and a feeling of increased access to their health providers. Before the study, the participating professionals expressed a fear that SMS would be overused, but these assumptions proved unfounded. Our Web-based, password-secured, encrypted message system prevented leaks of health information through mobile phone communication. Such systems might be appreciated all over the world to ensure that health information is not compromised.

Global penetration of mobile phones is increasing rapidly, with a penetration rate of up to 90% in high-income countries. This suggests the potential for using mobile phone applications in the treatment of young people. Future research into patients' own experiences of this feeling of access and its significance for treatment results, quality of life, and a sense of empowerment is recommended.

Technical Considerations

Technical problems were reported by the participants in the study and need to be taken into consideration for future versions of this application. Novel hardware and software already on the market based on other mobile platforms may ease the transfer challenges, overriding the need for Bluetooth technology. We appreciated the suggestions for additional functionality, for instance, new presentation software, and will try to implement this when designing the next software generation. A whole body of literature supports participatory design as mutually beneficial for both end users and developers [35-37].

Additional Findings

It seems that with a few reservations, the findings from the qualitative interviews were supported by the additional methods used to evaluate this study. The triangulation of methods, with the use of semistructured interviews, SUS, and knowledge questionnaires in relation to metabolic control (HbA_{1c}), strengthened this study. This provides additional information and allows for greater accuracy. An additional strength of the study is that it gave a voice to the experiences and concerns of the adolescents themselves. Previous work has shown that our health care system is poorly designed to meet the needs of patients with chronic diseases, and that Web-based programs and ICT tools are useful in meeting these patients' requirements [38,39]. Our findings contribute to this literature [20,40].

Conclusion

The integration of cornerstones in diabetes treatment into a picture-based diary gave young people a better understanding of their diabetes treatment. Furthermore, it is more an educational tool than a communication tool. We hypothesize that the reported effect of the picture-based diary is partly due to early maturation of the visual cortex of the brain. Members of diabetes care teams need to take into account that the child and adolescent brain is immature, a work in progress. It is likely that young people are less capable than we previously thought of converting different theoretical facts related to diabetes into applied knowledge in their daily lives.

Participants reported the Web-based encrypted SMS application for mobile phones as a sought-after tool that gave them a feeling of access and security. Participants gave valuable input for further development of these applications.

Simple information and communication tools like the applications in this study should be further developed and tested in larger-scale studies to investigate their role as a mediator for increased understanding and better self-care for diabetes in the adolescent population.

Acknowledgments

The authors wish to express their gratitude to the teenagers participating in the study, and to the diabetes nurses and doctors at the participating centers. Further, we are grateful for the support from the system development team at NST and to WTW for valuable help. This work was supported by a grant from Novo Nordic in Norway and the Norwegian Diabetes Association. We also appreciate the financial support from Professor Knut Dahl-Jorgensen's research group in the developmental process.

Authors' Contributions

DHF designed the study, collected and analyzed the data, and drafted the manuscript. FS contributed to data analysis, interpretation of results, and drafting and revising the manuscript. EÅ contributed to conceptualization and design of the applications and to drafting and revising the manuscript.

Conflicts of Interest

None declared.

References

1. Joner G, Stene LCM, Njolstad PR, Sovik O. Marked jump in incidence of childhood onset diabetes in Norway after ten years of stable incidence: prospective nationwide study 1999-2003. In: Diabetologia. 2005 Aug Presented at: 41st Annual Meeting of the European Association of Study of Diabetes; 09.10.2005; Athens,Greece.

2. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
3. Farmer A, Gibson OJ, Tarassenko L, Neil A. A systematic review of telemedicine interventions to support blood glucose self-monitoring in diabetes. Diabet Med 2005 Oct;22(10):1372-1378. [doi: 10.1111/j.1464-5491.2005.01627.x] [Medline: 16176199]
4. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
5. Blomquist KB, Brown G, Peersen A, Presler EP. Transitioning to independence: challenges for young people with disabilities and their caregivers. Orthop Nurs 1998;17(3):27-35. [Medline: 9847815]
6. The Norwegian Childhood Diabetes Registry. Annual Report 2010. Oslo: Oslo University Hospital; 2011.
7. Keulers B, Keulers M, Scheltinga M, Spauwen P. Evidence-based development of interactive patient education programs: a guide for achieving optimal computer-based patient education. Eur J Plast Surg 2006 Dec;29(4):169-175.
8. Blakemore SJ, Choudhury S. Development of the adolescent brain: implications for executive function and social cognition. J Child Psychol Psychiatry 2006;47(3-4):296-312. [doi: 10.1111/j.1469-7610.2006.01611.x] [Medline: 16492261]
9. Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, et al. Dynamic mapping of human cortical development during childhood through early adulthood. Proc Natl Acad Sci U S A 2004 May 25;101(21):8174-8179 [FREE Full text] [doi: 10.1073/pnas.0402680101] [Medline: 15148381]
10. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006 May;61(2):173-190. [doi: 10.1016/j.pec.2005.05.004] [Medline: 16122896]
11. Peregrin T. Picture this: visual cues enhance health education messages for people with low literacy skills. J Am Diet Assoc 2010 May;110(5 Suppl):S28-S32. [doi: 10.1016/j.jada.2010.03.005] [Medline: 20399294]
12. Statistics Norway, Statistisk Sentralbyrå. Norwegian Media Barometer 2011: We Watch More Television URL: http://www.ssb.no/medie_en/ [accessed 2012-05-28] [WebCite Cache ID 67zpWGeH7]
13. Avison D, Young T. Time to rethink health care and ICT? Commun ACM 2007 Jun;50(6):69-74. [doi: 10.1145/1247001.1247008]
14. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. Interactive information technology in diabetes care. J Gen Intern Med 2006 Feb;21(2):105-110. [doi: 10.1111/j.1525-1497.2005.00310.x] [Medline: 16390512]
15. Blake H. Innovation in practice: mobile phone technology in patient care. Br J Community Nurs 2008 Apr;13(4):160, 162-160, 165. [Medline: 18595303]
16. Pagliari C, Sloan D, Gregor P, Sullivan F, Detmer D, Kahan JP, et al. What is eHealth (4): a scoping exercise to map the field. J Med Internet Res 2005;7(1):e9 [FREE Full text] [doi: 10.2196/jmir.7.1.e9] [Medline: 15829481]
17. Årsand E. The Few Touch Digital Diabetes Diary: User-Involved Design of Mobile Self-Help Tools for People With Diabetes [doctoral thesis]. Tromsø: University of Tromsø; 2009. URL: http://munin.uit.no/handle/10037/2762 [accessed 2012-06-22] [WebCite Cache ID 68bcp2CtG]
18. Farmer A, Gibson O, Hayton P, Bryden K, Dudley C, Neil A, et al. A real-time, mobile phone-based telemedicine system to support young adults with type 1 diabetes. Inform Prim Care 2005;13(3):171-177. [Medline: 16259856]
19. Ferrer-Roca O, Cárdenas A, Diaz-Cardama A, Pulido P. Mobile phone text messaging in the management of diabetes. J Telemed Telecare 2004 Oct;10(5):282-285. [Medline: 15494086]
20. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
21. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. Inform Health Soc Care 2008 Sep;33(3):158-169. [doi: 10.1080/17538150802457562] [Medline: 18850399]
22. Brooke J. Redhatch Consulting Ltd. SUS: A Quick and Dirty Usability Scale URL: http://www.usabilitynet.org/trump/documents/ [accessed 2012-05-03] [WebCite Cache ID 67Np9fYuw]
23. Sauro J. Measuring Usability. 2011 Feb 2. Measuring Usability With the System Usability Scale (SUS) URL: http://www.measuringusability.com/sus.php [accessed 2012-05-03] [WebCite Cache ID 67Nq6fOyy]
24. Norwegian Health Informatics. Norsk Helseinformatikk AS. 2011. Diabetes Quiz 1, 2, 3, 4 URL: http://nhi.no/forside/quiz/ [accessed 2012-05-03] [WebCite Cache ID 67NqBiOiy]
25. WTW AS. 2011. HelseRespons URL: http://helserespons.no/ [accessed 2012-05-03] [WebCite Cache ID 67Nr0IRSQ]
26. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [Medline: 10940958]
27. Giorgi A, Giorgi B. Phenomenology. In: Smith JA, editor. Qualitative Psychology: A Practical Guide to Research Methods. 1st edition. London: SAGE; 2003:25-50.
28. Kvale S, Brinchmann S. Det kvalitative forskningsintervjuet. 2nd edition. Oslo: Gyledendal Akademiske; 2009.

29. Funnell MM, Anderson RM. Patient empowerment: a look back, a look ahead. Diabetes Educ 2003;29(3):454-8, 460, 462 passim. [Medline: 12854337]
30. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2009 Jan;32 Suppl 1:S87-S94. [doi: 10.2337/dc09-S087] [Medline: 19118294]
31. Schon DA. Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions (Higher Education Series). San Francisco, CA: Jossey-Bass; 1990.
32. Gaudieri PA, Chen R, Greer TF, Holmes CS. Cognitive function in children with type 1 diabetes: a meta-analysis. Diabetes Care 2008 Sep;31(9):1892-1897. [doi: 10.2337/dc07-2132] [Medline: 18753668]
33. Franklin V, Waller A, Pagliari C, Greene S. "Sweet Talk": text messaging support for intensive insulin therapy for young people with diabetes. Diabetes Technol Ther 2003;5(6):991-996. [doi: 10.1089/152091503322641042] [Medline: 14709202]
34. Franklin VL, Greene A, Waller A, Greene SA, Pagliari C. Patients' engagement with "Sweet Talk" - a text messaging support system for young people with diabetes. J Med Internet Res 2008 Jun;10(2):e20 [FREE Full text] [doi: 10.2196/jmir.962] [Medline: 18653444]
35. Schuler D, Namioka A. Participatory Design: Principles and Practices. 1st edition. Hillsdale, NJ: L Erlbaum Associates; 1993.
36. Greenbaum JM, Kyng M. From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In: Greenbaum JM, Kyng M, editors. Design at Work: Cooperative Design of Computer Systems. Hillsdale, NJ: L Erlbaum Associates; 1991.
37. Kensing F, Blomberg J. Participatory design: issues and concerns. Comput Support Coop Work 1998;7(3):167-185. [doi: 10.1023/A:1008689307411]
38. Ralston JD, Revere D, Robins LS, Goldberg HI. Patients' experience with a diabetes support programme based on an interactive electronic medical record: qualitative study. BMJ 2004 May 15;328(7449):1159. [doi: 10.1136/bmj.328.7449.1159] [Medline: 15142919]
39. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. JAMA 2002 Oct 9;288(14):1775-1779. [Medline: 12365965]
40. Blake H. Mobile phone technology in chronic disease management. Nurs Stand 2008 Nov;23(12):43-46. [Medline: 19093357]

Abbreviations

HbA1c: glycated hemoglobin
ICT: information and communication technology
NST: Norwegian Centre for Integrated Care and Telemedicine
SMS: short message service
SUS: System Usability Scale

Edited by G Eysenbach; submitted 04.05.12; peer-reviewed by R Istepanian, D Cox, L Diehl; comments to author 25.05.12; revised version received 25.06.12; accepted 23.07.12; published 06.08.12.

Please cite as:
Frøisland DH, Årsand E, Skårderud F
Improving Diabetes Care for Young People With Type 1 Diabetes Through Visual Learning on Mobile Phones: Mixed-Methods Study
J Med Internet Res 2012;14(4):e111
URL: http://www.jmir.org/2012/4/e111/
doi:10.2196/jmir.2155
PMID:22868871

Viewpoint

Making Sense of Mobile Health Data: An Open Architecture to Improve Individual- and Population-Level Health

Connie Chen[1], BA; David Haddad[2], MSc; Joshua Selsky[2,3], BSC; Julia E Hoffman[4], PsyD; Richard L Kravitz[5], MSc, MD; Deborah E Estrin[2,3], PhD; Ida Sim[2,6], MD, PhD

[1]School of Medicine, University of California San Francisco, San Francisco, CA, United States
[2]Open mHealth, San Francisco, CA, United States
[3]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, United States
[4]National Center for Post-Traumatic Stress Disorder, VA Palo Alto Healthcare System, Menlo Park, CA, United States
[5]Division of General Medicine and Center for Healthcare Policy and Research, University of California Davis, Sacramento, CA, United States
[6]Division of General Internal Medicine, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

Corresponding Author:
Ida Sim, MD, PhD
Division of General Internal Medicine
Department of Medicine
University of California San Francisco
Suite 308
1545 Divisadero Street
San Francisco, CA, 94143-0320
United States
Phone: 1 14155148657
Fax: 1 4155028627
Email: ida.sim@ucsf.edu

Abstract

Mobile phones and devices, with their constant presence, data connectivity, and multiple intrinsic sensors, can support around-the-clock chronic disease prevention and management that is integrated with daily life. These mobile health (mHealth) devices can produce tremendous amounts of location-rich, real-time, high-frequency data. Unfortunately, these data are often full of bias, noise, variability, and gaps. Robust tools and techniques have not yet been developed to make mHealth data more meaningful to patients and clinicians. To be most useful, health data should be sharable across multiple mHealth applications and connected to electronic health records. The lack of data sharing and dearth of tools and techniques for making sense of health data are critical bottlenecks limiting the impact of mHealth to improve health outcomes. We describe Open mHealth, a nonprofit organization that is building an open software architecture to address these data sharing and "sense-making" bottlenecks. Our architecture consists of open source software modules with well-defined interfaces using a minimal set of common metadata. An initial set of modules, called InfoVis, has been developed for data analysis and visualization. A second set of modules, our Personal Evidence Architecture, will support scientific inferences from mHealth data. These Personal Evidence Architecture modules will include standardized, validated clinical measures to support novel evaluation methods, such as n-of-1 studies. All of Open mHealth's modules are designed to be reusable across multiple applications, disease conditions, and user populations to maximize impact and flexibility. We are also building an open community of developers and health innovators, modeled after the open approach taken in the initial growth of the Internet, to foster meaningful cross-disciplinary collaboration around new tools and techniques. An open mHealth community and architecture will catalyze increased mHealth efficiency, effectiveness, and innovation.

(J Med Internet Res 2012;14(4):e112) doi:10.2196/jmir.2152

KEYWORDS

Mobile health; software tools; software engineering; open access to information; open architecture; open source; evaluation methodology; data analysis; data visualization

Introduction

Today, mobile phones are in nearly every pocket, with an estimated 83% of Americans owning a mobile phone and 35% of US mobile phone subscribers possessing a smartphone [1]. Mobile phones and devices—with their constant presence, connectivity, and multiple intrinsic sensors—can be easily integrated into daily life to effectively support chronic disease prevention and management.

Early research suggests that mobile health (mHealth) applications can empower individuals to track and manage their own health, thus improving user-centered outcomes. A recent randomized controlled trial of WellDoc, a diabetes management app that prompts users via short message service text messages to check and record their blood sugar values, showed a significant reduction in glycated hemoglobin among users at 1 year (1.9% in the treatment group versus 0.7% for usual care, $P < .001$), as well as a 20% reduction in emergency department use and hospitalization [2,3]. Text message reminders have also been shown to promote smoking cessation, improve attendance at medical appointments, increase knowledge about prenatal care, and encourage sunscreen use. Patients are increasingly using mobile phones to track their own health measures, ranging from blood sugar to vital signs to exercise and food intake [4-10].

Despite the promise of these preliminary findings, the evidence base for mHealth remains sparse and methodologically weak [11]. Anecdotally, the rates of reuse for mobile applications remain very low [12]. With nearly 12,000 health-related apps available, and more being created every day, the continued proliferation of mHealth apps runs the risk of simply creating confusion [13]. It is predicted that the number of mobile app downloads will reach 142 million by 2016, generating billions of real-world data points on patient health experiences and outcomes [14].

Unfortunately, the mHealth ecosystem lacks modular tools and techniques for drawing meaning and scientifically valid inferences from the masses of collected data. Without the development of more sophisticated and effective tools for data visualization and analysis, legitimate questions remain regarding mHealth’s projected impact on chronic disease management and prevention. In considering how the mHealth ecosystem might need to evolve to achieve maximum impact, we can draw lessons from the success of the Internet’s open architecture and its ability to support both open and closed proprietary applications. In contrast, the closed, stovepipe architecture of electronic health records yields a cautionary tale about the deleterious effects of highly closed ecosystems.

In this paper, we describe Open mHealth, a nonprofit organization that is building an open software architecture for mHealth and catalyzing an open community of developers, clinicians, researchers, and entrepreneurs to build and reuse Open mHealth modules across a broad range of mHealth applications, disease conditions, and user populations. Over time, the open architecture’s functionality and robustness will grow through reuse and community validation. Our postulate is that progress in mHealth will be best served by a dynamic, open, multidisciplinary community that innovates collaboratively on an open architecture.

How mHealth Data Contribute to Three Essential Feedback Loops

mHealth applications are rich sources of passive and actively collected data. These mHealth data are integral to three essential feedback loops for improving health outcomes: (1) patient-facing feedback to guide patients’ self-care (eg, how does taking the medication topiramate affect my pain?), (2) clinician-directed summary data to guide clinical decision making for individual patients (eg, how do the side effects and therapeutic benefits of topiramate balance for my patient?), and (3) research evidence to improve clinical care for groups of patients (populations) (eg, in patients with neuropathic pain, does topiramate reduce pain intensity and improve quality of life?).

These three feedback loops are powerful channels by which mHealth data can improve health outcomes. However, mHealth data tend to have lots of bias, noise, variability, and gaps, such that it is difficult to make sense of the data and extract relevant features and patterns to drive information through the feedback loops. Lack of visualization tools to help end users understand collected data and absence of analysis tools for generating robust clinical evidence remain significant impediments threatening to limit the impact of mobile technology on health outcomes.

In addition, the disaggregation of data across siloed applications and devices hinders patient-specific analysis. For example, a diabetic patient might find herself recording her insulin use, nutrition intake, exercise, blood sugar values, and mood in five separate mHealth applications. Without a shared architecture for data analysis, the patient and clinician would encounter significant friction in trying to correlate the blood sugar values with corresponding diet, exercise, or medication data. Without being able to determine what was driving a suboptimal blood sugar value, the clinician would not be able to make a fully informed adjustment to the patient’s management plan and might eventually discourage her patients from sharing this type of uninterpretable data.

On top of data aggregation and analysis, there must be visual displays that help users—both patients and providers—understand the meaning of their data. The lack of heterogeneous-data analysis tools among mHealth applications is similarly limiting the use of these data for clinical research [15]. This limitation presents a very high opportunity cost. Unlike traditional randomized controlled studies, which are costly, slow, and generate estimates of average treatment effects, trials of mHealth applications can be conducted as time series and n-of-1 studies for individual patients, enabling researchers to estimate with a high degree of granularity within-individual correlations between clinical interventions, specific patient behaviors, and health outcomes [16].

Open mHealth: An Open Architecture to Improve Individual- and Population-Level Health Outcomes

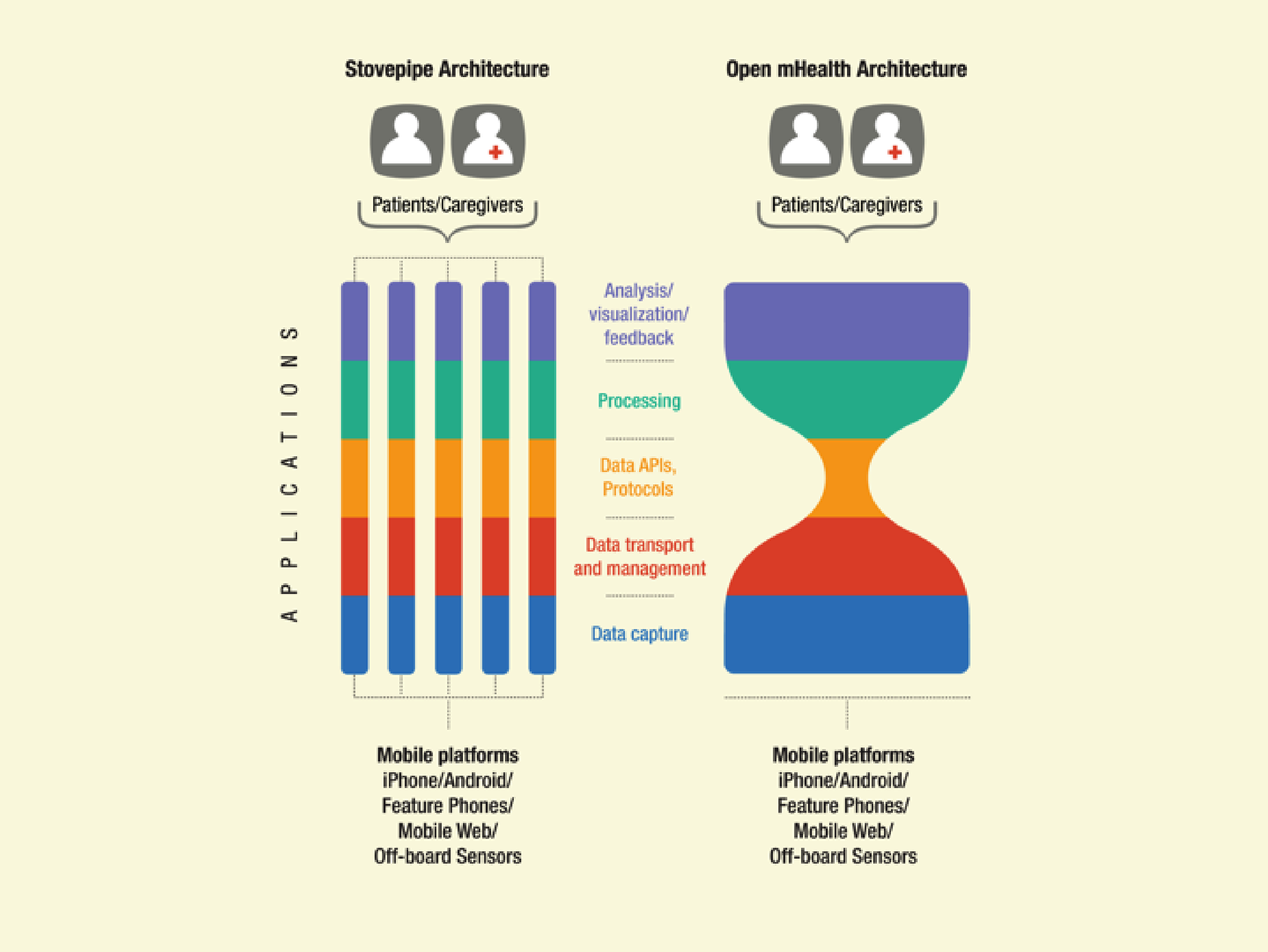
Open mHealth addresses the gap between the reality of fragmented mHealth applications and the full promise of mHealth powering the three feedback loops of personal care, clinical decision making, and research evidence in a virtuous cycle. Features of the needed solution include the following:

- Community: must be multidisciplinary, safe, and collaborative
- Iteration: delivers efficient reuse through collaborative cycles of development
- Flexible architecture: recognizes both the limits and the utility of existing closed systems and is designed to maximize participation from all players
- Shared learning: uses the strongest appropriate methods, matched to the evidence needs and the rapid pace of technological advances in mHealth
- Scalable solutions: offers mass customization of applications and evidence, from personal to population.

As is the case for the Internet, shared modules with open application programming interfaces (APIs) around a minimal set of common standards meet these needs for an open community. The Internet has what is called an hourglass architecture, from which it derived much of its success. In this architecture, a common communications protocol acts as a simple point of commonality at the narrow waist. This allows innovation to flourish through open interfaces, or APIs, both above and below the waist (Figure 1).

More recent examples of successful open software communities are Apache, Eclipse, and Mozilla. These communities spawned huge, lucrative industries through collaborative development that blended both proprietary and open components. mHealth is ripe for such open treatment.

Figure 1. Stovepipe versus open architectures: mHealth apps (left) are built independently with little sharing of data, methods, or learning. In contrast, the Internet has an hourglass architecture (right), in which a common protocol, transmission control protocol/Internet protocol (TCP/IP), acts as a simple point of commonality at the narrow waist that allows innovation to flourish through open application programming interfaces (APIs) both above and below the waist. Open mHealth aims to catalyze the mHealth ecosystem from a siloed architecture to an hourglass architecture to increase the scale and effectiveness of mHealth.



InfoVis

Open mHealth is catalyzing a decentralized, innovative community committed to developing sharable mHealth tools with open APIs that allow independently developed software components to be mixed and matched, swapped, and shared like Lego blocks (Figure 2) [17]. To begin, we developed InfoVis, which is the architectural scaffolding for data analysis and visualization building blocks that the Open mHealth community is creating, combining, evaluating, and adapting.

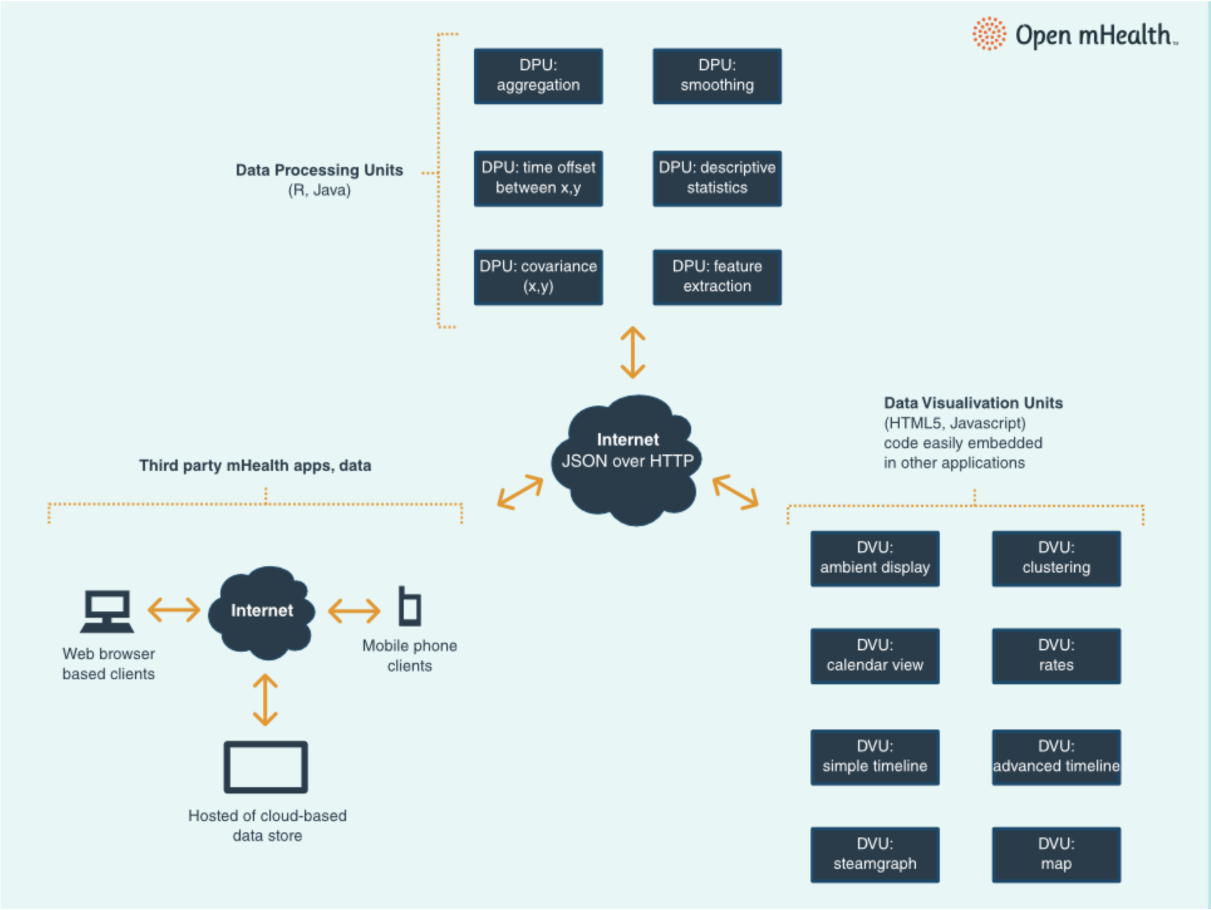
Open mHealth's architecture mimics the natural structure of the honeycomb. The foundational framework is a common set of principles and APIs that enable reusable software modules—or individual pieces of the honeycomb—to be built into and upon the underlying structure in a plug-and-play fashion. The architecture enables additional modules to be easily added and pieced together, facilitating the growth of the entire honeycomb and strengthening the overall structure.

The basic types of reusable software components in Open mHealth are data processing units (DPUs) and data visualization units (DVUs). DPUs are the building blocks for extracting relevant features from data streams, whereas DVUs enable the presentation of those features and patterns. Data storage units are components that manage the input and output of data to DPUs and DVUs, and are specific to particular data storage solutions (eg, a Health Insurance Portability and Accountability Act [HIPAA]-compliant cloud storage vendor). For any particular application, the DPUs, DVUs, and data storage units are embedded in a plug-and-play fashion within that application's running system, which can range from the Android operating system, for example, to full-featured platforms such as that of AT&T [18].

Each DPU and DVU does one task and can be composed to produce higher-level functions. For example, low-level DPUs can transform time series of on-phone and other sensor measurements (eg, accelerometer) into time series of user states (eg, sitting, walking, or driving). Midlevel DPUs compute clinically relevant metrics (eg, 6-minute walk test) [19]. Higher-level DPUs process and fuse one or more metrics (eg, activity metrics with self-report data) to come up with health markers for a person's state (eg, functional status). Such hierarchical analyses that transform lower-level data streams into higher-order markers will reduce the need for self-reports, thus mitigating the challenge of user engagement.

Open mHealth components can be incorporated into applications as libraries or can be invoked using JavaScript object notation over hypertext transfer protocol if they are developed with a Web service wrapper. We encourage component developers to support both library- and Web services-based approaches to accommodate application-specific preferences. All components must follow interoperability specifications that set forth common patterns of implementation and methods for data interchange. These specifications, which are continually being refined and are available at [20], follow these principles: using modern open source industry standards where possible; using lightweight interoperability standards; using declarative semantics with allowance for multiple bindings to multiple reference standards; and allowing standards to emerge through community patterns of use rather than imposition. For example, the data input to a DVU is at a minimum specified by a payload ID in the Open mHealth namespace (omh). This payload ID will be in the form of a string (eg, omh:serum-sodium) that is intentionally light on required formatting or semantic standards. This is to encourage and facilitate rapid exploration and innovation. As individual components gain traction with the community, external IDs can be used to map the payload ID to external existing health standards using a uniform resource name (URN) to the BioPortal server [21], an approach that is similar to the Substitutable Medical Apps, Reusable Technologies (SMART) platform [22] (eg, the URN of Logical Observation Identifiers Names and Codes [LOINC] code [23] for serum sodium values would be purl.bioontology.org/ontology/LNC/2951-2). Developers may choose to map to zero or more external standards, and all component IDs will be indexed for search functions that will be available in the Open mHealth code repository.

Figure 2. InfoVis: Third-party data applications and data stores save and manage data. Data processing units (DPUs) are the building blocks for extracting relevant features and patterns from data streams. Data visualization units enable visual presentations to be created from the extracted data features. Each DPU and DVU completes one task, and then can be composed for higher functions. HTTP = hypertext transfer protocol, JSON = JavaScript object notation.



An Illustrative Use Case

An excellent use case that demonstrates the power of our architectural approach is self-care tools for posttraumatic stress disorder (PTSD). PTSD Coach is a mobile app conceived by the US Department of Veterans Affairs and Department of Defense to help PTSD patients manage acute distress symptoms through education, connection with personal and public resources, self-assessment, and personalized, interactive tools rooted in cognitive behavioral principles [24].

Management of PTSD presents an ideal use case for augmentation with mHealth tools, as many patients who need care may not seek in-person assistance due to stigma, logistical barriers, or lack of problem recognition [25]. While standard face-to-face treatments for PTSD have been found to be quite effective, mobile apps can provide a convenient, location-independent, anonymous alternative to standard care. Even those individuals who are receiving PTSD care may experience distress in the week that passes between treatment sessions. For them, mHealth can provide just-in-time tools, including crisis management strategies, wherever and whenever they are needed. Primary reliance on mobile devices for Internet access is becoming increasingly common [26], but existing PTSD Web resources are not optimized for mobile use.

Whether used between clinic visits or for independent self-management, PTSD Coach supports skill acquisition for coping with acute symptoms (eg, guided relaxation or progressive muscle relaxation); self-assessment for improved problem recognition and self-monitoring; and education aimed at increasing knowledge about PTSD and its effective treatments, decreasing stigma, providing messages of normalization, and increasing likelihood of entering care. Links to support, both national and personal, improve the individual’s chances of entering care if it is warranted. Due to data sensitivity concerns, PTSD Coach was built as a stand-alone application for patient self-care with no transmission of data to or from the app for clinician involvement. Open mHealth collaborated with the PTSD Coach team to develop a version called PTSD Explorer that captures and reports user-reported and other data back to a HIPAA-compliant server. PTSD Explorer will be integrated into the Veterans Administration’s electronic health record in a future phase of this project. Open mHealth’s approach to electronic and personal health record integration is not yet defined but will follow principles aligned with the “substitutable apps” approach described by Mandl and Kohane [27].

To help clinicians make sense of PTSD Explorer patient data for use in direct clinical care, we developed InfoVis data processing and visualization modules, some of which are generically usable across disease conditions. Data input and output formats of each InfoVis unit are specified as part of the DPU or DVU interface. Open mHealth’s modular open approach facilitated rapid exploration and iterative innovation to support participatory design of PTSD Explorer with a team of clinical psychologists and psychiatrists, allowing for quick and easy configuration of new dataviews for various clinicians, cohorts, and conditions and over time.

Our development process for PTSD DPUs and DVUs explicitly involved abstracting common processing functions that would be reusable across multiple disease conditions. For example, one DVU we built for PTSD Explorer displays continuous data over time, which is a disease-independent function that can be chained with other components to yield more complex, disease-specific visualizations (eg, of PTSD Checklist scores or blood glucose values over time). We are now using the DPUs and DVUs built for the PTSD use case to generate patient-facing visualizations of various self-reported measures of chronic pain. Because the Open mHealth community will reuse and adapt these DPUs, DVUs, and their interfaces over time, multiple approaches to processing and visualizing PTSD and pain data will coexist and will be reused or not depending on their effectiveness and value for both disease-independent and disease-specific usage.

Over time, actual usage and demonstrated value of InfoVis components across the range of all Open mHealth projects will drive convergence on common interface and semantic usage standards. DPUs and DVUs will process data from data storage units that access a wide range of third-party data applications and stores. In this modular way, Open mHealth will build a strong, community-sourced open component architecture to complement proprietary innovations to maximize the overall impact of mHealth.

Personal Evidence Architecture

To further make sense of mHealth data streams, we have also designed a personal evidence architecture based on (1) standardized, validated clinical measures, 2) ways of collecting and interpreting these measures over time (such as statistical and graphical methods for time series analysis), and (3) use of an n-of-1 trial structure to reduce bias.

Standardization of Clinical Measures in mHealth

To aggregate data collected across multiple mHealth applications and n-of-1 studies, we must first adopt a standardized clinical vocabulary. As the basis for our personal evidence architecture, we are incorporating measures from a Patient Reported Outcomes Measurement Information System (PROMIS) put forth by the National Institutes of Health [28]. These PROMIS measures are a system of patient-reported health outcome assessments for physical, mental, and social well-being. These measures are broadly validated, having been widely used as primary or secondary end points in clinical studies of treatment effectiveness across disease conditions.

In addition to traditional clinical measures, there remains a need to develop and validate measures specifically for mHealth, which enables self-reported data to be collected several times a day, rather than once every few months. One of the greatest benefits of mHealth will be using passively collected data to estimate and predict health outcomes, and we anticipate a flurry of activity around the design and validation of these approaches. We are exploring a shared set of metadata tags to capture the contextual variables about how data are collected to ensure that data and evidence can make sense together, as well as separately.

Standardized measures and vocabulary specific to mHealth would enable the aggregation of high-value mHealth data, greatly expanding its potential to advance the public good. Data access is a priority of the US government’s Open Government Initiative, as exemplified by the flagship Community Health Data Initiative at the US Department of Health and Human Services [29]. Combining mHealth data with other community health data streams, such as the US Department of Veterans Affairs’ Blue Button personal health data initiative, would catalyze the information ecosystem, expanding the potential use and applicability of mHealth data in guiding clinical decision-making, performance improvement, and community public health initiatives.

N-of-1 Study Design

An n-of-1 study is a single-patient crossover trial in which an individual patient is randomly assigned to alternative interventions over time [30]. N-of-1 trials are most readily applicable to conditions that are chronic and to treatments that have a short onset and rapid washout. In contrast to anecdotal observations, n-of-1 trials can be used to identify effective approaches for an individual patient with enhanced scientific rigor [30-32].

Hundreds of n-of-1 studies have been completed and have been shown to be a rigorous means by which to generate personalized evidence [33]. However, n-of-1 trials have not gained much traction with clinicians, patients, and the scientific community at large because of the perception that such trials demand too much time and effort from clinicians and patients [34,35]. Mobile devices are well suited to overcome these barriers, as they facilitate data collection with minimal effort required by the physician and friction by the user.

Using our Personal Evidence Architecture, patients and clinicians (either together or independently) will be able to define a question, set up a study using an n-of-1 study template, and run the study on any mHealth app using our data analysis modules. Textbox 1 shows an example of this. By engaging patients in their own care, n-of-1 studies can enhance shared decision making, support better patient–clinician communication, and foster commitment to treatment, leading to better adherence. Furthermore, the results of n-of-1 trials can be aggregated using Bayesian methods, informing care of populations beyond the n-of-1 trial participants themselves. This would flip the traditional direction of research inference on its head, aggregating individual-level evidence to get at population-level evidence, rather than the other way around.

Textbox 1. Jack’s N-of-1 Trial.

Jack is a 55-year-old man with chronic back pain of moderate severity. He is taking Vicodin 5/500 tablets several times daily, but the pills make him sleepy and he’s not sure they do much. Jack has been followed closely by Nurse Practitioner (NP) Erlich for several years. He is randomly assigned to the Trialist and decides to design an n-of-1 trial comparing Vicodin 5/500 5 tabs daily with acetaminophen 500 mg 5 tabs daily. Working with NP Erlich, he decides on 1-week treatment periods for a total of 6 weeks. In addition to “pain interference” (a mandatory outcome), he creates “longest time, in minutes, able to sit continuously at work” as his customized outcome.

These choices are programmed into the Trialist. Beginning the next day, Jack is notified to start acetaminophen. He is also reminded at random intervals to note how long he has been sitting and to what extent he is experiencing discomfort. Once weekly he reports on mandatory outcomes. The process continues for 6 weeks, with the Trialist signaling Jack to switch at intervals. Based on Jack’s n-of-1 trial results plus priors supplied at the beginning of the study, the Trialist reports that there is a <30% chance that Vicodin is superior to acetaminophen with respect to prolonged sitting and only a 10% chance for reduced pain interference. Jack decides to go with plain Tylenol. He does well, and 6 months later he tells Ms Erlich he received a promotion at work.

Building the Open mHealth Community

Open mHealth is leading several ongoing projects, including work with the US federal government on PTSD and chronic pain as discussed above, to demonstrate the efficacy of an open architecture and the value of a broad, open community. For each project, we consider and aim to support all three feedback loops of self-care, clinical decision making, and research evidence to maximize clinical and scientific impact. These projects exemplify the value of joint technical and health innovation, since the high-level features are determined by what is clinically relevant (eg, mobility correlates of chronic pain), while the lower-level features are determined by what is technically feasible (accelerometer data from onboard phone sensors).

Open mHealth, and by extension mHealth, would be more successful with more projects where health innovators and developers can jointly develop tools and methods that are then shared through an open architecture. To catalyze this community, we are (1) convening capacity-building workshops, to increase the number of health innovators using Open mHealth, (2) holding developer connection events to galvanize the developer community, and (3) creating self-governing working groups to advance work in key topic areas.

Our paramount community engagement goal is to make it as easy and as worthwhile as possible for health and technology innovators to use and contribute to Open mHealth and to advance overall mHealth impact and effectiveness. While we do not expect all of mHealth to be open, we hope to foster a commons for sharing and learning that is inclusive of proprietary components and approaches, to allow health innovators and entrepreneurs to focus on their unique market offerings while increasing the validity, robustness, and efficiency of shared components and methods. In addition, adoption of our open architecture and Personal Evidence Architecture components will help generate more evidence that can be pooled and shared across studies, resulting in a stronger, more cohesive evidence base on mHealth efficacy for personalized care.

Open mHealth is the only organization dedicated to scaling effective mHealth solutions through an open architecture and an open collaborative community drawn from both the health and technology realms. Open mHealth is different from other organizations such as AT&T and Aetna that are developing open end-to-end platforms for mHealth applications, in that Open mHealth modules are embeddable within applications developed on those platforms. Open mHealth is not a competing platform but a source of shared components that can be compatible with AT&T and other open and closed platforms. Open mHealth is also different from organizations such as the mHealth Alliance or the nascent National Institutes of Health-led mHealth Public–Private Partnership, which are dedicated to scaling mHealth through public–private partnerships but not through a technical infrastructure. Open mHealth’s overall approach and the specific software developed for InfoVis and Personal Evidence Architecture are applicable within those partnerships, and Open mHealth continues to be active in collaborating with the myriad other organizations in the mHealth space.

Conclusion

While mHealth holds great promise, disappointment in health information technology has been commonplace, with hype cycles that come and go, punctuated by successful but ultimately limited pilots. At this juncture, improving health outcomes faster and at lower cost is essential. We look to adapt the model of one of the most successful innovations of all time—the Internet—to Open mHealth to seed and catalyze methods and techniques for maximal improvement of individual and population health through a vibrant and open mHealth community.

Acknowledgments

Open mHealth is a project of the Tides Center, and is funded by the Robert Wood Johnson Foundation and the California HealthCare Foundation. The authors would like to thank Ingrid Madden and Marc Schwartz for their contributions to this paper.

Conflicts of Interest

None declared.

References

1. Smith A. Americans and their cell phones. Washington, DC: Pew Internet & American Life Project; 2011 Aug 15. URL: http://pewinternet.org/~/media//Files/Reports/2011/Cell%20Phones%202011.pdf [accessed 2012-08-01] [WebCite Cache ID 69avRbEqc]
2. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942. [doi: 10.2337/dc11-0366] [Medline: 21788632]
3. Dolan B. mobihealthnews. 2011 Dec 6. Medicaid Patients Reduce Hospitalizations With WellDoc URL: http://mobihealthnews.com/15116/medicaid-patients-reduce-hospitalizations-with-welldoc/ [accessed 2012-04-27] [WebCite Cache ID 67EL6amm7]
4. Whittaker R, Borland R, Bullen C, Lin RB, McRobbie H, Rodgers A. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2009(4):CD006611. [doi: 10.1002/14651858.CD006611.pub2] [Medline: 19821377]
5. Obermayer JL, Riley WT, Asif O, Jean-Mary J. College smoking-cessation using cell phone text messaging. J Am Coll Health 2004;53(2):71-78. [doi: 10.3200/JACH.53.2.71-78] [Medline: 15495883]
6. Haug S, Meyer C, Schorr G, Bauer S, John U. Continuous individual support of smoking cessation using text messaging: a pilot experimental study. Nicotine Tob Res 2009 Aug;11(8):915-923. [doi: 10.1093/ntr/ntp084] [Medline: 19542517]
7. 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]
8. Vaughan C. Office of Communications, California State University San Marcos. 2011 Nov 1. San Diego Researchers First to Report Positive Impact of Text4Baby Program URL: http://www.text4baby.org/templates/beez_20/images/HMHB/SD_press_release.pdf [accessed 2012-04-27] [WebCite Cache ID 67ELAQUxN]
9. Armstrong AW, Watson AJ, Makredes M, Frangos JE, Kimball AB, Kvedar JC. Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. Arch Dermatol 2009 Nov;145(11):1230-1236. [doi: 10.1001/archdermatol.2009.269] [Medline: 19917951]
10. Krishna S, Boren SA. Diabetes self-management care via cell phone: a systematic review. J Diabetes Sci Technol 2008 May;2(3):509-517 [FREE Full text] [Medline: 19885219]
11. Déglise C, Suggs LS, Odermatt P. Short message service (SMS) applications for disease prevention in developing countries. J Med Internet Res 2012;14(1):e3 [FREE Full text] [doi: 10.2196/jmir.1823] [Medline: 22262730]
12. Suthoff B. Localytics. 2011 Jan 31. First Impressions Matter! 26% of Apps Downloaded in 2010 Were Used Just Once URL: http://www.localytics.com/blog/2011/first-impressions-matter-26-percent-of-apps-downloaded-used-just-once/ [accessed 2012-04-27] [WebCite Cache ID 67ELGTTaB]
13. Dolan B. mobihealthnews. 2012 Jul 11. An Analysis of Consumer Health Apps for Apple's iPhone 2012 URL: http://mobihealthnews.com/research/an-analysis-of-consumer-health-apps-for-apples-iphone-2012/ [accessed 2012-07-30] [WebCite Cache ID 69YHjptQ9]
14. iHealthBeat. 2011 Dec 1. 44M Mobile Health Apps Will Be Downloaded in 2012, Report Predicts URL: http://www.ihealthbeat.org/articles/2011/12/1/44m-mobile-health-apps-will-be-downloaded-in-2012-report-predicts.aspx [accessed 2012-04-27] [WebCite Cache ID 67ELKYPWa]
15. Kanas G, Morimoto L, Mowat F, O'Malley C, Fryzek J, Nordyke R. Use of electronic medical records in oncology outcomes research. Clinicoecon Outcomes Res 2010;2:1-14. [Medline: 21935310]
16. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 2004;82(4):661-687. [doi: 10.1111/j.0887-378X.2004.00327.x] [Medline: 15595946]
17. Estrin D, Sim I. Health care delivery. Open mHealth architecture: an engine for health care innovation. Science 2010 Nov 5;330(6005):759-760. [doi: 10.1126/science.1196187] [Medline: 21051617]
18. AT&T. 2012. AT&T ForHealth URL: http://www.att.com/gen/press-room?pid=18708 [accessed 2012-07-05] [WebCite Cache ID 68we1flAu]
19. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002 Jul 1;166(1):111-117 [FREE Full text] [Medline: 12091180]
20. GitHub Inc. 2012. Open mHealth GitHub Repository URL: https://github.com/openmhealth [accessed 2012-07-05] [WebCite Cache ID 68weBvS34]
21. National Center for Biomedical Ontology. 2012. BioPortal URL: http://bioportal.bioontology.org/ [accessed 2012-07-05] [WebCite Cache ID 68weIxWSC]
22. Office of the National Coordinator for Health Information Technology. 2012. Substitutable Medical Apps, Reusable Technologies: SMART Intro URL: http://www.smartplatforms.org/ [accessed 2012-07-05] [WebCite Cache ID 68weNPGDb]
23. Regenstrief Institute, Inc. 2012. Logical Observation Identifiers Names and Codes (LOINC) URL: http://loinc.org/ [accessed 2012-07-05] [WebCite Cache ID 68weR6RrF]

24. US Department of Veterans Affairs. 2012 Mar 28. Mobile App: PTSD Coach URL: http://www.ptsd.va.gov/public/pages/PTSDCoach.asp [accessed 2012-04-27] [WebCite Cache ID 67ELQQNA0]
25. Hoge CW, Castro CA, Messer SC, McGurk D, Cotting DI, Koffman RL. Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. N Engl J Med 2004 Jul 1;351(1):13-22 [FREE Full text] [doi: 10.1056/NEJMoa040603] [Medline: 15229303]
26. Smith A. Smartphones as an internet appliance. Washington, DC: Pew Internet & American Life Project; 2011 Jul 11. URL: http://pewinternet.org/Reports/2011/Smartphones/Section-2/Smartphones-as-an-internet-appliance.aspx [accessed 2012-04-27] [WebCite Cache ID 67ELVGVGK]
27. Mandl KD, Kohane IS. No small change for the health information economy. N Engl J Med 2009 Mar 26;360(13):1278-1281 [FREE Full text] [doi: 10.1056/NEJMp0900411] [Medline: 19321867]
28. US National Institutes of Health. Patient Reported Outcomes Measurement Information System (PROMIS) URL: http://www.nihpromis.org/?AspxAutoDetectCookieSupport=1 [accessed 2012-07-05] [WebCite Cache ID 68weV13SY]
29. US Department of Health and Human Services. Community Health Data Initiative URL: http://www.hhs.gov/open/plan/opengovernmentplan/initiatives/initiative.html [accessed 2012-04-27] [WebCite Cache ID 67ELZwfNF]
30. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. Med Care 2011 Aug;49(8):761-768. [doi: 10.1097/MLR.0b013e318215d90d] [Medline: 21478771]
31. Nikles CJ, Yelland M, Glasziou PP, Del Mar C. Do individualized medication effectiveness tests (n-of-1 trials) change clinical decisions about which drugs to use for osteoarthritis and chronic pain? Am J Ther 2005;12(1):92-97. [Medline: 15662296]
32. Scuffham PA, Nikles J, Mitchell GK, Yelland MJ, Vine N, Poulos CJ, et al. Using N-of-1 trials to improve patient management and save costs. J Gen Intern Med 2010 Sep;25(9):906-913. [doi: 10.1007/s11606-010-1352-7] [Medline: 20386995]
33. Scuffham PA, Yelland MJ, Nikles J, Pietrzak E, Wilkinson D. Are N-of-1 trials an economically viable option to improve access to selected high cost medications? The Australian experience. Value Health 2008;11(1):97-109. [doi: 10.1111/j.1524-4733.2007.00218.x] [Medline: 18237364]
34. Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to N-of-1 trials? Insiders' perspectives and a look to the future. Milbank Q 2008 Dec;86(4):533-555. [doi: 10.1111/j.1468-0009.2008.00533.x] [Medline: 19120979]
35. Kravitz RL, Paterniti DA, Hay MC, Subramanian S, Dean DE, Weisner T, et al. Marketing therapeutic precision: Potential facilitators and barriers to adoption of n-of-1 trials. Contemp Clin Trials 2009 Sep;30(5):436-445. [doi: 10.1016/j.cct.2009.04.001] [Medline: 19375521]

Abbreviations

API: application programming interface
DPU: data processing unit
DVU: data visualization unit
HIPAA: Health Insurance Portability and Accountability Act
LOINC: Logical Observation Identifiers Names and Codes
PROMIS: Patient Reported Outcomes Measurement Information System
PTSD: posttraumatic stress disorder
SMART: Substitutable Medical Apps, Reusable Technologies

Edited by G Eysenbach; submitted 01.05.12; peer-reviewed by B Macleod; comments to author 25.06.12; revised version received 05.07.12; accepted 06.07.12; published 09.08.12.

Please cite as:
Chen C, Haddad D, Selsky J, Hoffman JE, Kravitz RL, Estrin DE, Sim I
Making Sense of Mobile Health Data: An Open Architecture to Improve Individual- and Population-Level Health
J Med Internet Res 2012;14(4):e112
URL: http://www.jmir.org/2012/4/e112/
doi:10.2196/jmir.2152
PMID:22875563

Review

The Therapeutic Relationship in E-Therapy for Mental Health: A Systematic Review

Madalina Sucala[1,2], PhD; Julie B Schnur[1], PhD; Michael J Constantino[3], PhD; Sarah J Miller[1], PsyD; Emily H Brackman[1]; Guy H Montgomery[1], PhD

[1]Department of Oncological Sciences, Mount Sinai School of Medicine, New York, NY, United States
[2]Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania
[3]Department of Psychology, University of Massachusetts Amherst, Amherst, MA, United States

Corresponding Author:
Madalina Sucala, PhD
Department of Oncological Sciences
Mount Sinai School of Medicine
Box 1130
1425 Madison Avenue
New York, NY, 10029
United States
Phone: 1 212 659 5504 ext 85504
Fax: 1 212 659 5479
Email: madalina.sucala@mssm.edu

Abstract

Background: E-therapy is defined as a licensed mental health care professional providing mental health services via e-mail, video conferencing, virtual reality technology, chat technology, or any combination of these. The use of e-therapy has been rapidly expanding in the last two decades, with growing evidence suggesting that the provision of mental health services over the Internet is both clinically efficacious and cost effective. Yet there are still unanswered concerns about e-therapy, including whether it is possible to develop a successful therapeutic relationship over the Internet in the absence of nonverbal cues.

Objective: Our objective in this study was to systematically review the therapeutic relationship in e-therapy.

Methods: We searched PubMed, PsycINFO, and CINAHL through August 2011. Information on study methods and results was abstracted independently by the authors using a standardized form.

Results: From the 840 reviewed studies, only 11 (1.3%) investigated the therapeutic relationship. The majority of the reviewed studies were focused on the therapeutic alliance—a central element of the therapeutic relationship. Although the results do not allow firm conclusions, they indicate that e-therapy seems to be at least equivalent to face-to-face therapy in terms of therapeutic alliance, and that there is a relationship between the therapeutic alliance and e-therapy outcome.

Conclusions: Overall, the current literature on the role of therapeutic relationship in e-therapy is scant, and much more research is needed to understand the therapeutic relationship in online environments.

(J Med Internet Res 2012;14(4):e110) doi:10.2196/jmir.2084

KEYWORDS

e-Therapy; therapeutic relationship; therapeutic alliance; common factors in psychotherapy

Introduction

As of 2011, 78.3% of the US population had Internet access [1]. The widespread use of the Internet has affected mental health care delivery, with a rapid expansion of e-therapy [2]. E-therapy is defined as “a licensed mental health care professional providing mental health services via e-mail, video conferencing, virtual reality technology, chat technology, or any combination of these” [3].

Although there is growing evidence that e-therapy is effective for a variety of conditions [2,4-6], researchers have expressed concerns about e-therapy use [7,8]. One of the primary concerns about e-therapy is related to the perceived difficulty of developing an effective therapeutic relationship in the absence of nonverbal cues [6].

Extensive literature on face-to-face psychotherapy indicates that the therapeutic relationship accounts for more variability in psychotherapy outcomes than do specific therapy ingredients [9-11]. Given the crucial role of the therapeutic relationship in face-to-face interventions, it is important to assess the role of therapeutic relationship in e-therapy as well.

Although e-therapy research began over 15 years ago, there has been no attempt to review the findings pertaining to the status of the therapeutic relationship in online interventions. Heeding the guidelines published by the American Psychological Association (Division 29), which state that descriptions of effective psychotherapies that do not mention the therapeutic relationship are “seriously incomplete and potentially misleading on both clinical and empirical grounds” [12], it is imperative to investigate systematically the status of the therapeutic relationship in e-therapy. This paper represents the first attempt to summarize and review the existing findings. More specifically, the review examined (1) how the therapeutic relationship is being assessed in e-therapy, (2) patients’ satisfaction with the therapeutic relationship in e-therapy, (3) differences in the therapeutic relationship between e-therapy and face-to-face therapy, (4) factors that may influence the therapeutic relationship in e-therapy, and (5) the relationship between the therapeutic relationship and treatment outcome in e-therapy.

Methods

Search Strategy

We searched 3 electronic databases (PubMed, PsycINFO, and CINAHL) from their respective inceptions through August 2, 2011. For PubMed, the search terms were (counseling[MeSH] OR psychotherapy[MeSH]) AND Internet[MAJR]). The search was limited by language (the paper had to be in English), by methodology (the study had to be a clinical trial; randomized controlled trial; clinical trial, phase 1; clinical trial, phase 2; clinical trial, phase 3; clinical trial, phase 4; comparative study; controlled clinical trial; or a technical report), and by sample (human subjects). This search, with these limits, and taking only the items with an abstract, yielded a total of 468 abstracts.

For PsycINFO, the major search terms were ([exp counseling OR exp psychotherapy] AND exp Internet). The search was limited by language (the paper had to be in English), by methodology (the study had to be an empirical study, experimental replication, follow-up study, longitudinal study, prospective study, retrospective study, quantitative study, or treatment outcome/randomized clinical trial), by publication type (the study had to be a journal article published in a peer-reviewed journal), and by sample (the study had to be conducted on humans). This search, with these limits, and taking only the items with an abstract, yielded a total of 188 abstracts.

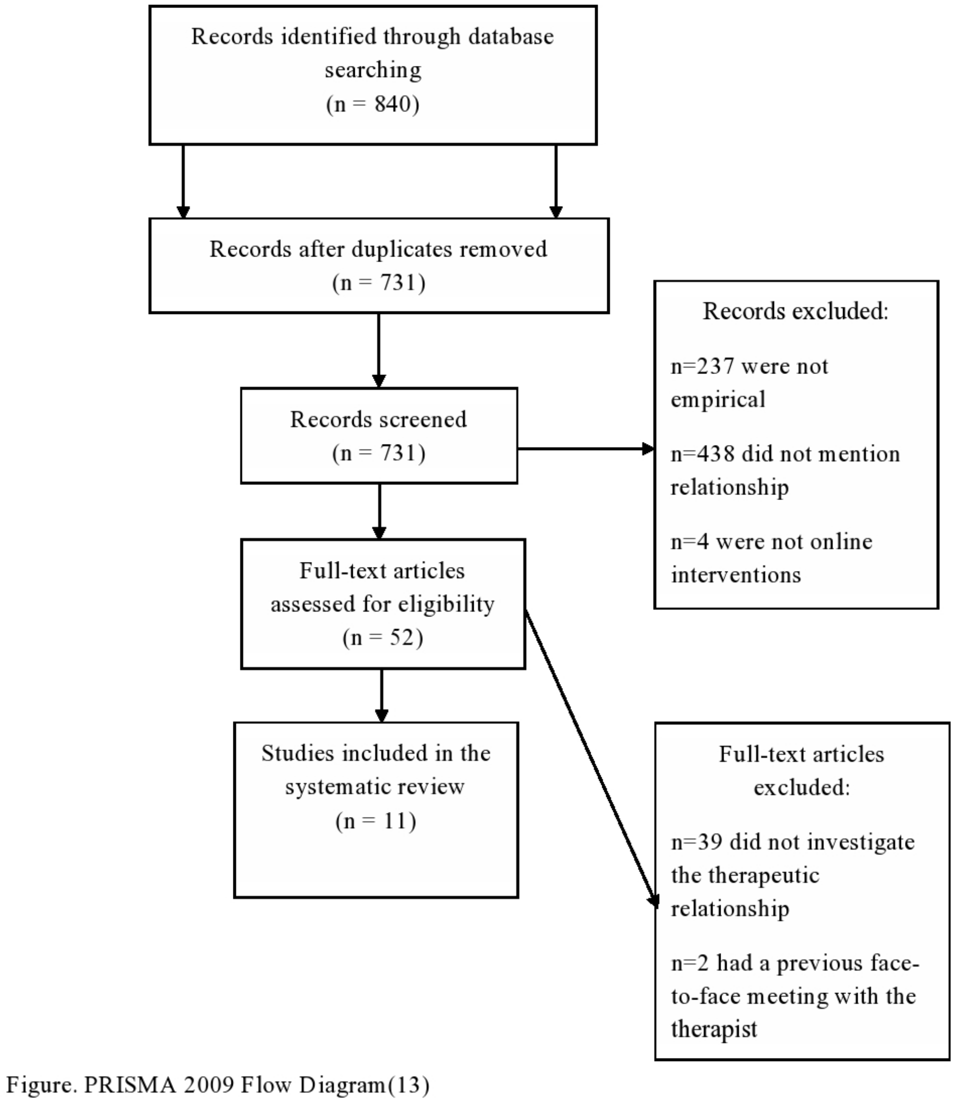
For CINAHL, the major search terms were ([MH psychotherapy OR MH counseling] AND MH Internet). The search was limited by language (the paper had to be in English) and by publication type (the study had to be a peer-reviewed research article). This search, with these limits, and taking only the items with an abstract, yielded a total of 184 abstracts.

Selection Strategy

We carefully screened the abstracts of all articles identified by the electronic searches (840 in total) to determine whether the abstracts met the following inclusion criteria: (1) described an intervention study that empirically assessed the effects of e-therapy on an outcome (excluding qualitative studies, survey studies, reviews, meta-analyses, etc), and (2) reported data relevant to the therapeutic relationship. Specifically, abstract text had to use the word *relationship* or *alliance* to be included in the review. Interventions had to be consistent with the above definition of e-therapy. This excluded interventions that were described as self-help, peer-led groups, online communities, or volunteer-led online support. If a given study had multiple intervention groups, at least one intervention group had to meet the e-therapy definition. There were no inclusion or exclusion criteria regarding the focus of the treatment.

Based on these criteria, the number of eligible abstracts was reduced from 840 abstracts to 56 abstracts. Figure 1 details reasons for exclusion [13]. The 56 manuscripts were obtained and read in full, independently, by two of the authors (MS and SJM). They completed a standardized form assessing the above-listed criteria. Any lack of consensus was discussed with JBS and GHM until consensus was reached.

Figure 1. PRISMA 2009 Flow Diagram.



Data Abstraction and Study Characteristics

We included 11 of the 56 studies in the review (Table 1). Each of the 11 papers was abstracted independently by MS and SJM. The data abstraction form included (1) authors and publication date, (2) the study sample (eg, demographic characteristics), (3) the intervention and the interventionists, (4) study design elements, (5) the therapeutic relationship element studied and how it was measured (eg, scale), and (6) the findings pertaining to the therapeutic relationship. Any discrepancies were discussed among the authors (MS and SJM) with reference to the original manuscript until consensus was reached.

The quality of each study was evaluated independently by MS and SJM according to the following eight validity criteria, which were adapted from the Consolidated Standards of Reporting Trials (CONSORT) guidelines [14,15] and Delphi criteria list [16]: randomization; allocation concealment; blinding of outcome assessments; comparability of groups at baseline; withdrawals; handling of dropouts in analyses; use of intention-to-treat analysis; and multiple follow-up assessments. Scores were given, with 1 point allocated for each criterion satisfied (range 0-8 points). The interrater reliability between them was .84, indicating strong agreement [17]. Any discrepancies were discussed (with JBS and GHM) with reference to the original manuscript until consensus was reached.

Although some quantitative data were available, there were insufficient data for formal comprehensive meta-analyses. Therefore, we report effect sizes where possible and informative.

Results

Among the included studies, investigating the therapeutic relationship was a primary objective for 6 [4,6,18-21], whereas for the other 5 studies, the assessment of the therapeutic relationship was a secondary outcome.

Study and Participant Characteristics

Table 1 summarizes design characteristics and quality scores. The quality scores for the studies ranged from 0 to 7 out of a maximum of 8 points. Because blinding of participants to the type of intervention is often practically impossible in psychosocial interventions, as participants must actively engage in them, no study could receive a perfect score of 8. A total of 6 studies were described as randomized; 2 studies used only pre-post comparisons to analyze data pertaining to the therapeutic relationship. The other studies had a nonequivalent group design: 2 studies compared e-therapy data with data from previously published studies and 1 study used naturalistic independent samples of participants provided by a youth counseling service. The main limitations for the studies were not comparing groups at baseline [4,20,22,23]; not reporting the use of intention-to-treat analyses or handling of missing data [4,6,18,21]; and not using follow-up assessments [4,6,18,20,22,24].

The main therapeutic approach used in the analyzed studies was cognitive behavioral therapy (CBT) ($k = 9$). E-therapists were psychologists and psychotherapists ($k = 6$), psychology students ($k = 4$), and counselors ($k = 1$). Overall, the dose of e-therapy ranged from 1 session to 11 weeks, with a mean of 7.75 (SD 2.37) weeks. Communication between therapist and patient was conducted via asynchronous email and website postings ($k = 8$), synchronous website text exchange ($k = 1$), synchronous chat ($k = 1$), or a combination of asynchronous email and synchronous chat ($k = 1$).

Table 1. Study design characteristics[a].

Authors	Sample size	E-therapy communication modality	Therapists	Treatment length	Study design	Study quality[b]
Cook and Doyle [6]	15	Email and chat, both asynchronous and synchronous communication[c]	1 PhD, 3 masters level, and 1 masters student	1–5 sessions	Nonequivalent groups design; compared e-therapy versus previous normative data from face-to-face counseling	1
King et al [18]	186	Website postings, synchronous communication	"Trained counselors"	1 session, with a typical session duration of 50–80 minutes	Nonequivalent group design; e-therapy versus telephone counseling	2
Kiropoulos et al [24]	86	Website and email, asynchronous communication	9 registered psychologists, 1 probationary psychologist	8 weeks, weekly assignments	RCT[d] design; e-CBT[e] versus face-to-face CBT	6
Klein et al [22]	16	Website and email, asynchronous communication	6 registered psychologists, 1 probationary registered psychologist	10 weeks, weekly assignments	RCT design; pre- to posttreatment comparisons	3
Klein et al [23]	22	Website and email, asynchronous communication	6 registered psychologists, 1 probationary registered psychologist	10 weeks, weekly assignments	RCT design; pre- to posttreatment comparisons	4
Knaevelsrud and Maercker [20]	48	Email, asynchronous communication	2 psychologists	5 weeks, 2 weekly 45-minute writing assignments	RCT; e-CBT versus waiting list	3
Knaevelsrud and Maercker [19]	96	Email, asynchronous e-communication	2 clinical psychologists at the doctoral level	5 weeks, 2 weekly 45-minute writing assignments	RCT; e-CBT versus waiting list	6
Reynolds et al [4]	17	Email, asynchronous communication	16 psychotherapists (62.5% qualified to work in the United States)	Not reported	Nonequivalent group design; e-therapy versus data from prior face-to-face studies	0
Ruwaard et al [25]	239	Email, asynchronous communication	25 doctoral and 1 postgraduate student in clinical psychology	7 weeks, 5 hours of therapist time	RCT; e-CBT versus waiting list	6
Ruwaard et al [26]	54	Website, asynchronous communication	18 graduate-level clinical psychologists and 6 therapists	11 weeks, 22–44 hours of patient time and 7–14 hours of therapist time	RCT; e-CBT versus waiting list	6
Trautmann and Kroner-Herwig [21]	18	Chat, synchronous communication	3 clinical psychology graduate students	6 weeks, a weekly chat with the therapist	RCT; e-CBT versus e-psychoeducation	4

[a] The table presents the information about the studies' characteristics; not all of the studies provided a detailed description of the methods.

[b] Score for number of validity criteria satisfied (range 1–8).

[c] Synchronous communication between therapist and patient takes places in real time, in a same-time/different-place mode (eg, chat); asynchronous communication takes place over a period of time through a different-time/different-place mode (eg, email).

[d] Randomized controlled trial.

[e] Cognitive behavioral therapy.

Table 2 summarizes participants' characteristics by study. The participants were receiving e-therapy for a variety of problems, including mental health diagnosis (eg, posttraumatic stress disorder, k = 4; depression, k = 1; and panic disorder and agoraphobia, k = 1), psychological distress related to medical problems (eg, headaches, k = 1), work-related distress (k = 1), general distress (k = 1), and other self-reported presenting problems (eg, symptoms of depression, symptoms of anxiety, stress, relationship issues, or childhood abuse; k = 2). Participants were both adolescents (k = 2) and adults (k = 9). A majority of the adult patients were women (at least 60% across the studies) with a high level of education (at least 44% across the studies completed college).

Table 2. Participants' characteristics[a].

Authors	Presenting problem	Age (years), mean (SD)	Gender	Education	Race/ ethnicity
Cook and Doyle [6]	Relationship issues and depression	41.40 (15.99)	93%women	All participants completed at least high school	"Primarily white"
King et al [18]	Distress	14.25 (no SD provided)	80.1% women	Not reported	Not reported
Kiropoulos et al [24]	Panic disorder and agoraphobia	38.96 (11.13)	72.1% women	Mean education level 12.53 (SD 6.14) years	Not reported
Klein et al [22]	Posttraumatic stress disorder	48 (no SD provided)	81.2% women	Mean education level 13 years (no SD provided)	Not reported
Klein et al [23]	Posttraumatic stress disorder	45.8 (no SD provided)	77.2% women	Mean education level 13.3 (SD 3.5) years	Not reported
Knaevelsrud and Maercker [20]	Posttraumatic stress disorder	35 (no SD provided)	92% women	55% had a university degree	Not reported
Knaevelsrud and Maercker [19]	Posttraumatic stress disorder	35 (no SD provided)	90% women	44% had a university degree	Not reported
Reynolds et al [4]	Depression, stress, anxiety, and childhood abuse	Median 39	71% women	94.1% completed high school	82% Caucasian
Ruwaard et al [25]	Work-related stress	44 (8)	60% women	84% had a university degree	Not reported
Ruwaard et al [26]	Depression	21 (10)	69% women	65% had a university degree	Not reported
Trautmann and Kroner-Herwig [21]	Headache	13.4 (2.6)	Not reported	Not reported	Not reported

[a] The table presents the information about the patients' characteristics that the studies provided; not all the studies provided the full range of demographic information.

Assessment of the Therapeutic Relationship in E-Therapy

Table 3 presents, by study, the characteristics of the measures used to assess the therapeutic relationship.

A total of 3 of the studies [22-24] used the Therapist/Therapeutic Alliance Questionnaire, a modified version of the Helping Alliance Questionnaire [29]. The scale requires the participants to estimate the degree to which the therapeutic alliance with their therapist was helpful.

Cook and Doyle [6] used the Working Alliance Inventory [27]. The scale is based on Bordin's concept of therapeutic alliance: therapist-patient agreement on therapeutic goals; therapist-patient agreement on therapeutic tasks, and the quality of the emotional bond between the therapist and the patient [35,36]. Two other studies [19,20] used the short version of the Working Alliance Inventory [30].

King et al [18] used the Therapeutic Alliance Scale [28]. The scale evaluates the overall therapeutic alliance with 3 subcomponents: mutual liking between therapist and patient, collaboration between therapist and patient, and resistance (ie, resistance to the treatment program).

Reynolds et al [4] used the Agnew Relationship Measure-Short Form [31]. The scale evaluates the overall therapeutic alliance with 3 subcomponents: bond and partnership, confidence (defined as the confident collaboration between patient and therapist), and openness (defined as "the felt freedom to disclose and reveal personal material without fear of censure or embarrassment") [32].

Trautmann and Kroner-Herwig [21] used an Internet-based patient-therapist alliance/assistance questionnaire [33], which was adapted for use with children and adolescents and to the conditions of e-therapy (eg, "My therapist's explanations helped and supported me;" "This week I learned something that can help me cope with my headache"). The questionnaire was developed for the purpose of this study, and no other details were provided about this measure.

The other 2 studies included in this review [25,26] investigated patients' satisfaction with the therapeutic relationship. Both used a scale with 4 items, which assessed the extent to which the participants regarded their relationship with the therapist as being pleasant and personal, whether they perceived the relationship as growing during the treatment, and whether they missed face-to-face contact. No other details were provided about this measure.

Table 3. Therapeutic relationship measures and findings.

Authors	Therapeutic relationship measure	Moment of assessment[a]	Measured constructs	Item description and properties	Instrument description and psychometric properties	Patient ratings
Cook and Doyle [6]	Working Alliance Inventory [27]	Third session	Overall therapeutic alliance and 3 subcomponents: (1) agreement on goals, (2) agreement on tasks, (3) bond	36 items scored on a 7-point Likert scale ranging from 1 (*never*) to 7 (*always*). Subscales scores can range from 12 to 84, and total scores can range from 36 to 252. Higher scores reflect more positive ratings of therapeutic alliance.	Good construct validity and high internal consistency on the composite score (.93) as well as for the subscales (.85-.88) [27]	Overall therapeutic alliance, mean 215.07; agreement on task, mean 70.33; agreement on goals, mean 72.27; bond, mean 72.47
King et al [18]	Therapeutic Alliance Scale [28]	Posttreatment	Overall therapeutic alliance and 3 subcomponents: (1) mutual liking, (2) resistance, (3) collaboration	30 items, scored on a 3-point Likert scale (*disagree, somewhat agree, agree*), with scores ranging from 30 to 90, where higher scores indicate more positive perceptions of alliance.	Good internal consistency for the subscales (.83-.90) [28] and for entire measure (.92) [18]	Overall alliance, mean 74.0 (SD 10.4); resistance, mean 24.6 (SD 4.6); mutual liking, mean 26.2 (SD 3.6); collaboration, mean 23.1 (SD 4.8)
Kiropoulos et al [24]	Therapist Alliance Questionnaire [29]	Posttreatment	Overall therapeutic alliance (ie, the degree to which patients experience their therapeutic relationship as being helpful)	17 items rated on a Likert scale ranging from 1 to 7 and summed to produce a total score, which can range from 17 to 102. Higher scores reflect more positive ratings of the therapeutic alliance.	No data were found about the psychometric properties	Mean 83.13 (SD 11.20)
Klein et al [22]	Therapeutic Alliance Questionnaire [29]	Posttreatment	Overall therapeutic alliance (ie, the degree to which patients experience their therapeutic relationship as being helpful)	17 items rated on a Likert scale ranging from 1 to 7 and summed to produce a total score, which can range from 17 to 102. Higher scores reflect more positive ratings of the therapeutic alliance.	No data were found about the psychometric properties	Mean 86.25 (SD 16.23)
Klein et al [23]	Therapeutic Alliance Questionnaire [29]	Posttreatment	Overall therapeutic alliance (ie, the degree to which patients experience their therapeutic relationship as being helpful)	17 items rated on a Likert scale ranging from 1 to 7 and summed to produce a total score, which can range from 17 to 102. Higher scores reflect more positive ratings of the therapeutic alliance.	No data were found about the psychometric properties	Mean 89.18 (SD 15.13)
Knaevelsrud and Maercker [20]	Working Alliance Inventory-short version [30]		Overall therapeutic alliance and 3 subcomponents: (1) agreement on goals, (2) agreement on tasks, (3) bond	12 items scored on a 7-point Likert scale with scores ranging from 1 to 7. Higher scores reflect more positive ratings of therapeutic alliance.	Good internal consistency for the subscales (.90-.92) and for the composite score (.98) [30]	Overall alliance, mean 5.8 (SD 0.62); agreement on goals, mean 5.8 (SD 0.77); agreement on tasks, mean 5.7 (SD 0.80); bond, mean 6.2 (SD 0.75)
Knaevelsrud and Maercker [19]	Working Alliance Inventory-short version [30]	Fourth session	Overall therapeutic alliance and 3 subcomponents: (1) agreement on goals, (2) agreement on tasks, (3) bond	12 items scored on a 7-point Likert scale with scores ranging from 1 to 7. Higher scores reflect more positive ratings of therapeutic alliance.	Good internal consistency for the subscales (.90-.92) and for the composite score (.98) [30]	Overall alliance, mean 5.8 (SD 0.72); agreement on goals, mean 5.8 (SD 0.77); agreement on tasks, mean 5.7 (SD 0.83); bond, mean 6.2 (SD 0.69)

Authors	Therapeutic relationship measure	Moment of assessment[a]	Measured constructs	Item description and properties	Instrument description and psychometric properties	Patient ratings
Reynolds et al [4]	Agnew Relationship Measure-Short Form [31]	Fourth session	Overall therapeutic alliance and 3 subcomponents: (1) bond and partnership, (2) confidence, (4) openness	12 items, each rated on a 7-point Likert scale, with higher scores indicating more positive perceptions of alliance.	Good construct validity and high internal consistency, ranging from .83 to .89 [32]	Bond and partnership, mean 5.97 (SD 1.26); confidence, mean 6.19 (SD 1.24); openness, mean 5.27 (SD 1.42)
Ruwaard et al [25]	Treatment satisfaction items [25]	Posttreatment	Aspects of patients' perceived relationship with their therapists	4 items	The scale was developed for the purpose of the study; no data provided about the psychometric properties	Participants rated the relationship as pleasant (88%) and personal (75%); they perceived the relationship to grow during the treatment (57%); 68% said that they did not miss face-to-face contact
Ruwaard et al [26]	Treatment satisfaction items [25]	Posttreatment	Aspects of patients' perceived relationship with their therapists	4 items	The scale was developed for the purpose of the study; no data provided about the psychometric properties	Participants rated the relationship as pleasant (88%) and personal (75%); they perceived the relationship to grow during the treatment (57%); 89% said that they did not miss face-to-face contact
Trautmann and Kroner-Herwig [21]	Internet-based questionnaire on patient-trainer alliance/assistance scale [33]	Posttreatment	Patient–trainer alliance	The scale ranges from 0 to 3, with higher scores indicating more positive perceptions of alliance.	The scale was developed for the purpose of the study; no data provided about the psychometric properties	E-CBT[b], median 2.8, range 2–3; e-psychoeducation, median 2.7, range 2–3

[a] If multiple assessment points were used, we present the data for the earliest point of assessment, since previous studies showed that the level of alliance, regardless of the length of therapy, is established within the first sessions, recommending that alliance be assessed at the beginning of therapy [34].

[b] Cognitive behavioral therapy.

Patients' Satisfaction With the Therapeutic Relationship in E-Therapy

Ruwaard and colleagues [25] investigated the impact of e-therapy on work-related stress on a sample of 239 participants. Results indicated that participants rated the e-therapy relationship as being pleasant (210/239, 88%) and personal (179/239, 75%); 136 of 239 (57%) perceived the relationship to grow during the treatment, and 163 of 239 (68%) said that they did not miss face-to-face contact.

A second study conducted by Ruwaard and colleagues [26] investigated the impact of e-therapy on depression on a sample of 54 participants. The study yielded similar results to those previously described, in that the majority of the participants rated the e-therapy relationship as being pleasant (47/54, 87%) and personal (42/54, 78%), perceived the relationship to be growing during the treatment (42/54, 78%), and reported that they did not miss face-to-face contact (48/54, 89%). The results of the studies indicated that patients offered high ratings for therapeutic relationship in e-therapy (see Table 3).

Differences in the Therapeutic Relationship Between E-Therapy and Face-to-Face Therapy

A total of 3 studies investigated differences in the therapeutic alliance between e-therapy and face-to-face therapy [4, 6, 24]. Results were mixed, with 2 studies showing no significant differences in therapeutic alliance (eg, overall alliance and various subscales) between e-therapy and face-to-face therapy, and 1 study showing higher scores for therapeutic alliance in e-therapy than in face-to-face therapy.

Kiropoulos and colleagues [24] investigated whether the therapeutic alliance in e-therapy is different from face-to-face therapy in a study comparing e-CBT with face-to-face CBT for panic disorder and agoraphobia. Results indicated that there was no significant difference between groups for therapeutic alliance score ($t_{47} = –1.02$, $P = .31$, $d = 0.29$; according to Cohen [37], an effect size of 0.2 to 0.3 represents a small effect, around 0.5 represents a medium effect, and 0.8 or greater represents a large effect).

Reynolds et al [4] compared e-therapy data with data from previously published face-to-face studies. Results indicated that patients gave high ratings for therapeutic alliance in e-therapy, with the means for the subscales of bond and partnership between therapist and patient (mean 5.97, SD 1.26) and confident collaboration between therapist and patient (mean 6.19, SD 1.24) within the range of reported means for previous face-to-face therapy studies. The mean for openness (mean 5.27, SD 1.42) in e-therapy was below the range of means from the prior face-to-face studies. However, it is important to note that no test for statistical significance was performed.

Cook and Doyle [6] investigated whether the therapeutic alliance in e-therapy is different from face-to-face therapy in a sample of 15 participants. Results indicated that the overall working alliance scores ($t_{14} = 3.03$, $P < .001$, $d = 0.60$) and the agreement between therapist and patient on the therapy goals subscale scores ($t_{14} = 2.30$, $P = .03$, $d = 0.79$) were significantly higher in e-therapy than in face-to-face interventions, with medium to large effect sizes. The agreement between therapist and patient on tasks ($t_{14} = 1.26$, $P = .22$, $d = 0.22$) and the bond between therapist and patient were rated higher as well ($t_{14} = 1.62$, $P = .12$, $d = 0.33$), although the difference did not reach statistical significance and the effect sizes were small.

Factors That May Influence the Therapeutic Relationship in E-Therapy

A total of 2 studies investigated factors that might influence the therapeutic relationship in e-therapy [6, 20]. Knaevelsrud and Maercker [20] reported an inverse relationship between pretreatment symptom severity and therapeutic alliance ratings, such that patients who experienced more severe anxiety symptoms at the beginning of treatment tended to give lower ratings for the bond between therapist and patient subscale ($r = –.34$, $P < .05$, $d = 0.72$). There was an overall tendency for an inverse relationship between pretreatment anxiety and depression symptoms, and agreement on goals and task subscales ratings, but the correlations did not reach statistical significance and the effect sizes were small to moderate (all $P > .05$, all $d < 0.40$).

Cook and Doyle [6] investigated the impact of communication modality on the therapeutic relationship. Their results did not reach statistical significance. However, they reported that participants who used chat as the primary mode of communication (eg, as opposed to email) had consistently higher means for the therapeutic alliance than did participants who used email (overall alliance, $t_{13} = 1.54$, $P = .10$, $d = 1.13$; agreement on task, $t_{13} = 0.89$, $P = .37$, $d = 0.54$; agreement on goals, $t_{13} = 1.54$, $P = .12$, $d = 1.09$; bond, $t_{13} = 1.92$, $P = .07$, $d = 1.19$), obtaining medium to large effect sizes. Participants who used more than one modality of communication (eg, email plus chat) had higher ratings for the therapeutic alliance than did participants who used only one modality of communication (overall alliance, $t_{13} = 1.87$, $P = .08$, $d = 1.02$; agreement on tasks, $t_{13} = 1.67$, $P = .11$, $d = 0.91$; agreement on goals, $t_{13} = 1.40$, $P = .18$, $d = 0.76$; bond, $t_{13} = 1.67$, $P = .11$, $d = 0.91$). However, it should be noted that these results were based on comparisons made on very small samples of participants (eg, participants who used chat as a primary communication mode, n = 3, versus participants who used email as a primary communication mode, n = 12).

Is the Quality of the Therapeutic Relationship Linked to Treatment Outcome in E-Therapy?

A total of 3 studies investigated the impact of the therapeutic alliance on treatment outcome [18-20]. Knaevelsrud and Maercker [20] investigated the relationship between working alliance and the outcome of e-therapy for patients with posttraumatic stress disorder. Results showed that the composite score for therapeutic alliance correlated positively with residual gain scores for anxiety ($r = .33$, $P < .05$, $d = 0.69$), which indicates that patients who rated the alliance as better had greater reduction of their anxiety scores at posttreatment.

Knaevelsrud and Maercker [19], in a later study investigating the impact of e-therapy on posttraumatic stress disorder, found that overall patient-rated working alliance at posttreatment predicted 15% of the variance in the scores for posttraumatic stress symptoms (adjusted $R^2 = .148$, $F_{2,39} = 8.15$, $P < .001$), obtaining a large effect size.

King and colleagues [18] investigated the impact of online versus telephone counseling for adolescents. Their results revealed a modest trend toward a relationship between the collaboration subscale scores and posttreatment distress (beta = 0.25, $t = 1.83$, $P = .07$, $d = 0.14$) and a significant effect of the resistance subscale on posttreatment distress (beta = 1.21, $t = 2.40$, $P < .05$, $d = 0.19$).

Discussion

To our knowledge, this study is the first to summarize and review the findings on the role of the therapeutic relationship in e-therapy. The most striking finding was the limited number of studies investigating the therapeutic relationship. Of the 840 reviewed studies, only 11 (1.3%) addressed and investigated the issue of the therapeutic relationship, and of these, only 6 investigated the therapeutic relationship as a primary objective. In other words, the results indicate that, although the therapeutic relationship is considered to be an important common factor operating across all psychotherapies [34,38], the study of the therapeutic relationship appears to have been largely ignored in the e-therapy literature.

The reviewed studies have the merit of providing a first glimpse into the role of the therapeutic relationship in e-therapy. However, due to the small number of studies and to their methodological limits, the findings must be interpreted with caution.

Study and Participant Characteristics

The methodological limits of the studies included the lack of suitable control groups (nonrandom allocation or nonequivalent group design), lack of pretest information, poor reporting and handling of dropouts in the analyses, and more generally an often incomplete presentation of results (eg, not reporting standard deviations, not reporting effect sizes). As research moves forward, it is important for future studies to adhere to the standards of conducting and reporting psychosocial interventions [13]. Improved reporting will lead to the enrichment of systematic reviews and allow for better-informed treatment decision making among practitioners. Another issue that might limit enthusiasm for the findings is that the majority of studies were affected by a selection bias. The recruitment was performed through webpages or email announcements, which already rely on a certain familiarity with the use of the Internet. This is a particularly important methodological limitation, since previous studies indicated that, the more familiar participants are with Internet-based contact, the more positively they judge Internet-based contact to be [39]. Future studies should clarify the role of Internet familiarity in the therapeutic relationship in e-therapy.

As for the studies' characteristics, it is interesting to note that the main therapeutic approach was CBT, which easily lends itself to standardized instructions and short-term, manualized approaches. Because almost all studies included in this review used a CBT approach, it is difficult to infer the status of therapeutic relationship in online interventions that use therapeutic approaches that are less structured.

The overwhelming majority of participants were women, consistent with previous research that has found that more women than men use the Internet for mental health information and services [40]. Participants tended to be highly educated. Not all the studies provided information about other important demographics, such as race and ethnicity, but it is interesting to note that in the 2 studies that reported this information, the participants were primarily white. These results are consistent with previous findings, indicating that health information seeking over the Internet is more prevalent among white, educated women, and that mental health information seeking in particular tends to have the same type of consumers [40]. To determine the appropriateness of e-therapy and to investigate the status of the therapeutic relationship in online environments across ethnic groups, future research should include more diverse samples of patients.

Assessment of the Therapeutic Relationship in E-Therapy

The majority of the studies focused on a specific element of the therapeutic relationship, namely the therapeutic alliance. The studies used a variety of measures to assess the therapeutic alliance, defining the concept by the instrument used to measure it. In that sense, as Norcross [34] suggested, "instrumentation defines the construct." In addition, some of the studies used measures that had been created on an ad hoc basis. If progress is to be made in this field, future studies should reach toward a consensus by using validated measures based on supported conceptualizations of what therapeutic alliance is [34,41].

As for the timing of the assessment, the majority of the studies investigated the therapeutic alliance at the end of therapy. Previous studies indicate that the level of alliance, regardless of the length of therapy, is established within the first sessions [34]. Meta-analytic studies also revealed that early alliance is more predictive of outcome than is alliance assessed later in therapy [42]. Accordingly, it is recommended that future e-therapy studies assess alliance at the beginning of the treatment.

Differences in the Therapeutic Relationship Between E-Therapy and Face-to-Face Therapy

A surprising finding, given the previous concerns related to the lack of nonverbal cues in e-therapy, is that e-therapy seems to be at least equivalent to face-to-face therapy in terms of the therapeutic relationship (more specifically therapeutic alliance). Although very promising and clearly worthy of attention, this line of research is in its infancy, and further research is needed to draw firm conclusions.

Factors That May Influence the Therapeutic Relationship in E-Therapy

Although the results do not allow firm conclusions to be drawn, it seems that investigating factors such as communication modality and pretreatment symptom severity as moderators of the therapeutic relationship might be a fruitful direction of research. In addition, all of the studies included in this review used text-based communication methods; thus, it would be important to investigate the status of the therapeutic relationship when the communication modality includes video conferencing (eg, through Skype), where the verbal cues are not missing and the communication is synchronous.

Is the Quality of the Therapeutic Relationship Linked to Treatment Outcome in E-Therapy?

The 3 studies investigating the impact of the therapeutic alliance on treatment outcome indicate that these two factors have a positive relationship. This avenue of research should be further pursued, as it offers a hint that the beneficial effects of this therapeutic relationship element are not restricted to face-to-face therapies.

Limitations, Conclusions, and Future Directions

The present review has limitations. First, it was based on searches in three databases—PubMed, PsycINFO, and CINAHL—and was limited to published papers in English. It is possible that additional relevant papers exist outside of the present sample of papers. Second, the reviewed abstracts were required to report the assessment of the therapeutic relationship. It is possible that papers exist for studies in which investigating the relationship was not a main goal, and thus their abstracts might not refer to it. Future work may include more languages, include unpublished manuscripts, and use a wider variety of search terms to confirm the generalizability of the present conclusions. Additionally, once the literature grows large enough, a formal meta-analysis should be conducted to estimate the overall effect size for both the impact of the relationship on psychotherapy outcome and differences in the relationship between face-to-face therapy and e-therapy. Future meta-analyses would also have the potential to explore moderators of relationship effects and would be an important step forward for the field.

Overall, this review summarizes research to date on the therapeutic relationship in e-therapy. If relationship is considered a common factor in successful psychotherapy, it should become commonly studied in e-therapy as well. Looking to the future, we hope that the present findings will spur investigation into the role of the therapeutic relationship in e-therapy.

Acknowledgments

Preparation of this paper was supported by the National Cancer Institute (CA131473, CA081137, CA129094, CA159530). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Miniwatts Marketing Group. 2012. Internet World Stats URL: http://www.internetworldstats.com/ [accessed 2012-02-15] [WebCite Cache ID 65TVJUUCS]
2. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. J Technol Hum Serv 2008;26(2):109-160. [doi: 10.1080/15228830802094429]
3. Manhal-Baugus M. E-therapy: practical, ethical, and legal issues. Cyberpsychol Behav 2001 Oct;4(5):551-563. [Medline: 11725648]
4. Reynolds DJ, Stiles WB, Grohol JM. An investigation of session impact and alliance in internet based psychotherapy: preliminary results. Couns Psychother Res 2006;6(3):164-168. [doi: 10.1080/14733140600853617]
5. Mallen MJ, Vogel DL, Rochlen AB, Day SX. Online counseling: reviewing the literature from a counseling psychology framework. Couns Psychol 2005;33:819-871. [doi: 10.1177/0011000005278624]
6. Cook JE, Doyle C. Working alliance in online therapy as compared to face-to-face therapy: preliminary results. Cyberpsychol Behav 2002 Apr;5(2):95-105. [Medline: 12025884]
7. Rochlen AB, Zack JS, Speyer C. Online therapy: review of relevant definitions, debates, and current empirical support. J Clin Psychol 2004 Mar;60(3):269-283. [doi: 10.1002/jclp.10263] [Medline: 14981791]
8. Wells M, Mitchell KJ, Finkelhor D, Becker-Blease KA. Online mental health treatment: concerns and considerations. Cyberpsychol Behav 2007 Jun;10(3):453-459. [doi: 10.1089/cpb.2006.9933] [Medline: 17594270]
9. Lambert MJ, Barley DE. Research summary on the therapeutic relationship and psychotherapy outcome. Psychotherapy 2001;38(4):357-361. [doi: 10.1037/0033-3204.38.4.357]
10. Norcross JC. Empirically supported therapy relationships. In: Norcross JC, editor. Psychotherapy Relationships That Work: Therapist Contributions and Responsiveness to Patients. New York, NY: Oxford University Press; 2002:3-16.
11. Wampold BE. The Great Psychotherapy Debate: Models, Methods, and Findings. Mahwah, NJ: L Erlbaum Associates; 2001.
12. Ackerman SJ, Benjamin LS, Beutler LE, Gelso CJ, Goldfried MR, Hill C, et al. Empirically supported therapy relationships: conclusions and recommendations of the Division 29 Task Force. Psychotherapy 2001;38(4):495-497. [doi: 10.1037/0033-3204.38.4.495]
13. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [Medline: 19622511]
14. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001 Apr 17;134(8):663-694. [Medline: 11304107]
15. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001 Apr 14;357(9263):1191-1194. [Medline: 11323066]
16. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol 1998 Dec;51(12):1235-1241. [Medline: 10086815]
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
18. King R, Bambling M, Reid W, Thomas I. Telephone and online counselling for young people: a naturalistic comparison of session outcome, session impact and therapeutic alliance. Couns Psychother Res 2006;6(3):175-181. [doi: 10.1080/14733140600874084]
19. 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]
20. Knaevelsrud C, Maercker A. Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients? J Med Internet Res 2006;8(4):e31 [FREE Full text] [doi: 10.2196/jmir.8.4.e31] [Medline: 17213049]
21. Trautmann E, Kroner-Herwig B. Internet-based self-help training for children and adolescents with recurrent headache: a pilot study. Behav Cogn Psychother 2008;36:241-245. [doi: 10.1017/S1352465808004219]

22. Klein B, Mitchell J, Gilson K, Shandley K, Austin D, Kiropoulos L, et al. A therapist-assisted Internet-based CBT intervention for posttraumatic stress disorder: preliminary results. Cogn Behav Ther 2009 Jun;38(2):121-131. [doi: 10.1080/16506070902803483] [Medline: 20183691]
23. Klein B, Mitchell J, Abbott J, Shandley K, Austin D, Gilson K, et al. A therapist-assisted cognitive behavior therapy internet intervention for posttraumatic stress disorder: pre-, post- and 3-month follow-up results from an open trial. J Anxiety Disord 2010 Aug;24(6):635-644. [doi: 10.1016/j.janxdis.2010.04.005] [Medline: 20447802]
24. Kiropoulos LA, Klein B, Austin DW, Gilson K, Pier C, Mitchell J, et al. Is internet-based CBT for panic disorder and agoraphobia as effective as face-to-face CBT? J Anxiety Disord 2008 Dec;22(8):1273-1284. [doi: 10.1016/j.janxdis.2008.01.008] [Medline: 18289829]
25. Ruwaard J, Lange A, Bouwman M, Broeksteeg J, Schrieken B. E-mailed standardized cognitive behavioural treatment of work-related stress: a randomized controlled trial. Cogn Behav Ther 2007;36(3):179-192. [doi: 10.1080/16506070701381863] [Medline: 17852171]
26. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009 Dec;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]
27. Horvath AO, Greenberg LS. Development of the Working Alliance Inventory. In: Greenberg LS, Pinsoff WM, editors. The Psychotherapeutic Process: A Research Handbook. New York, NY: Guilford Press; 1986:529-556.
28. Bickman L, Vides de Andrade AR, Lambert EW, Doucette A, Sapyta J, Boyd AS, et al. Youth therapeutic alliance in intensive treatment settings. J Behav Health Serv Res 2004;31(2):134-148. [Medline: 15255222]
29. Luborsky L, McLellan AT, Woody GE, O'Brien CP, Auerbach A. Therapist success and its determinants. Arch Gen Psychiatry 1985 Jun;42(6):602-611. [Medline: 4004503]
30. Tracey TJ, Kokotovic AM. Factor structure of the working alliance inventory. Psychol Assess 1989;1(3):207-210. [doi: 10.1037/1040-3590.1.3.207]
31. Stiles WB, Hardy GE, Cahill J. The short ARM (a short form of the Agnew Relationship Measure). 2003 Presented at: Annual meeting of the North American Society for Psychotherapy Research; 2003; Newport, RI, USA.
32. Cahill J, Stiles WB, Barkham M, Hardy GE, Stone G, Agnew-Davies R, et al. Two short forms of the Agnew Relationship Measure: the ARM-5 and ARM-12. Psychother Res 2012;22(3):241-255. [doi: 10.1080/10503307.2011.643253] [Medline: 22191469]
33. Krampen G, Wald B. Kurzinstrumente fur die Prozessevaluation und adaptive Indikation in der Allgemeinen und Differentiellen Psychotherapie und Beratung. Diagnostica 2001;47:43-50. [doi: 10.1026//0012-1924.47.1.43]
34. Norcross JC. In: Norcross JC, editor. Psychotherapy Relationships That Work: Evidence-Based Responsiveness. 2nd edition. New York, NY: Oxford University Press; 2011.
35. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy 1976;16(3):252-260. [doi: 10.1037/h0085885]
36. Bordin ES. Theoryresearch on the therapeutic working alliance: new directions. In: Horvath AO, Greenberg LS, editors. The Working Alliance: Therory, Research, and Practice. New York, NY: Wiley; 1994:13-37.
37. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: L Erlbaum Associates; 1988.
38. Castonguay LG, Constantino MJ, Holtforth MG. The working alliance: Where are we and where should we go? Psychotherapy (Chic) 2006;43(3):271-279. [doi: 10.1037/0033-3204.43.3.271] [Medline: 22122096]
39. Mallen MJ, Day SX, Green MA. Online versus face-to-face conversations: an examination of relational and discourse variables. Psychotherapy 2003;40(1/2):155-163. [doi: 10.1037/0033-3204.40.1/2.15]
40. 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]
41. Hatcher RL, Barends A, Hansell J, Gutfreund MJ. Patients' and therapists' shared and unique views of the therapeutic alliance: an investigation using confirmatory factor analysis in a nested design. J Consult Clin Psychol 1995 Aug;63(4):636-643. [Medline: 7673541]
42. Eaton TT, Abeles N, Gutfreund MJ. Therapeutic alliance and outcome: impact of treatment length and pretreatment symptomatology. Psychotherapy 1988;25(4):536-542. [doi: 10.1037/h0085379]

Abbreviations

CBT: cognitive behavioral therapy

Edited by G Eysenbach; submitted 19.02.12; peer-reviewed by B Wampold; accepted 24.05.12; published 02.08.12.

Please cite as:
Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH
The Therapeutic Relationship in E-Therapy for Mental Health: A Systematic Review
J Med Internet Res 2012;14(4):e110
URL: http://www.jmir.org/2012/4/e110/
doi:10.2196/jmir.2084
PMID:22858538

Viewpoint

Munchausen by Internet: Current Research and Future Directions

Andy Pulman[1], MA; Jacqui Taylor[2], BSC, MSc, PhD

[1]School of Health & Social Care, Bournemouth University, Bournemouth, United Kingdom

[2]School of Design, Engineering & Computing, Psychology Research Centre, Bournemouth University, Bournemouth, United Kingdom

Corresponding Author:
Andy Pulman, MA
School of Health & Social Care
Bournemouth University
Royal London House
Christchurch Road
Bournemouth, BH1 3LT
United Kingdom
Phone: 44 1202 962749
Fax: 44 1202 962736
Email: apulman@bournemouth.ac.uk

Abstract

Background: The Internet has revolutionized the health world, enabling self-diagnosis and online support to take place irrespective of time or location. Alongside the positive aspects for an individual's health from making use of the Internet, debate has intensified on how the increasing use of Web technology might have a negative impact on patients, caregivers, and practitioners. One such negative health-related behavior is Munchausen by Internet.

Objective: Munchausen by Internet occurs when medically well individuals fake recognized illnesses in virtual environments, such as online support groups. This paper focuses on the aspect of Munchausen by Internet in which individuals actively seek to disrupt groups for their own satisfaction, which has not yet been associated with the wider phenomena of Internet trolls (users who post with the intention of annoying someone or disrupting an online environment).

Methods: A wide-ranging review was conducted to investigate the causes and impacts of online identity deception and Munchausen by Internet drawing on academic research and case studies reported online and in the media.

Results: The limited research relating to motivation, opportunity, detection, effects, and consequences of Munchausen by Internet is highlighted and it is formally linked to aspects of trolling. Case studies are used to illustrate the phenomenon. What is particularly worrying is the ease with which the deception can be carried out online, the difficulty in detection, and the damaging impact and potential danger to isolated victims.

Conclusions: We suggest ways to deal with Munchausen by Internet and provide advice for health group facilitators. We also propose that Munchausen by Internet and Munchausen by Internet trolling should be formally acknowledged in a revised version of the Diagnostic and Statistical Manual DSM-5. This will assist in effectively identifying and minimizing the growth of this behavior as more people seek reassurance and support about their health in the online environment. We also suggest directions for future research.

(J Med Internet Res 2012;14(4):e115) doi:10.2196/jmir.2011

KEYWORDS

Munchausen by Internet; Internet trolls; identity deception; malingering; factitious disorder; deviance; social networking sites; health support groups

Introduction

Lying to receive medical attention is not a new concept. Although it was not until the 1800s that factitious disorder was formally recognized, evidence of malingering dates back as far as Roman times. Munchausen syndrome was first described in 1951 [1], followed in 1977 by Munchausen syndrome by proxy [2]. More recently, the increasing use of the Internet to provide support for illnesses and other medical issues has introduced the concept of health-related online identity deception—Munchausen by Internet (identified in 2000) [3].

This paper reviews the research about the origins and evidence relating to these disorders. The limited research relating to motivation, opportunity, detection, effects, and consequences is highlighted with case studies. We conclude with practical and theoretical suggestions. We suggest practical ways for the health care community to deal with Munchausen by Internet and provide advice for health support group facilitators. We propose that Munchausen by Internet and Munchausen by Internet trolling be formally acknowledged and suggest future research directions.

Types of Deception Regarding Health

Malingering

Malingering is defined as a deliberate behavior for a known external purpose [4]. Resnick [5] described three types of malingering: (1) *pure malingering* in which the individual falsifies all symptoms; (2) *partial malingering* in which an individual has symptoms but exaggerates the impact they have on daily life; and (3) *false imputation* in which the individual has valid symptoms but is dishonest as to the source of the problems. Other forms of malingering are *simulation* in which a person emulates symptoms of a specific disability and *dissimulation* in which the patient denies the existence of problems that would account for the symptoms (eg, drug abuse).

Historically, the Roman physician, Galen, presented the earliest evidence of malingering in the second century AD. One patient pretended to suffer from colic in order to avoid a public meeting and another faked a knee injury in order to remain home while his master took a long journey [6]. Causes of malingering vary. Although the malingering individual is seeking tangible gains, such as time away from work or avoiding an activity, the underlying motivation differs [3].

Malingering was widespread throughout Russia during the early 1950s because people sought to escape sanctions or coercion [7]. Russian physicians were limited by the state to only four medical dispensations. Patients were identified as: (1) needing medical care, (2) thinking they needed medical care, (3) faking, or (4) making direct pleas for medical dispensation. Low social trust is exhibited when certain ethnic groups have low trust in health care as an institution because of historical oppression and deception of their cultural group [8]. The early dependence upon doctors by poor Russian laborers has been said to have altered the doctor-patient relationship to one of mutual mistrust and deception [7]. If an individual patient trusts their doctor because the doctor has been assessed as trustworthy over time, this is an example of interpersonal trust. Recent health care literature has explored the role of patients' trust in patient behaviors such as adhering to medical advice, malpractice litigation [9], and seeking health care services [10]. Some researchers believe that recent changes in health care practices are undermining the trust relationship between patients and physicians [11], with particular emphasis on the fact that technology is replacing the human element in medical practice thereby reducing patient's trust in physicians [9]. Trust has been defined as a feeling (often based on inconclusive evidence) of certainty that a person or a thing will not fail [12]. There are several trust relationships. The three trust relationships relevant to this discussion are: (1) *interpersonal trust* defined as a human's trust with another human whether face-to-face or through a device; (2) *social trust* defined as a human's trust with a system or institution; and (3) *trust in automation* defined as a human's trust with technology or a device.

There are no quantifiable numbers on how many people might misuse the Internet to abuse trust relationships, but many researchers have shown how the Internet lends itself to deception [3]. People might malinger online for external incentives, such as money or gifts, but fraud is usually the principle motivation [4]. Within the medical arena, the Internet offers anonymous access to vast amounts of information on illnesses and support groups for rare diseases, opening up opportunities for people with the urge to pretend they are sick and abuse trust [13].

Factitious Disorder and Munchausen Syndrome

Factitious disorder (FD) is an umbrella category covering a group of mental disturbances in which patients intentionally feign physical or mental illness without obvious benefit. Gavin [14] first described "factitious disease" in 1843. He described soldiers and seamen who mimicked illness to incite compassion or attention. The exact incidence of FD is unknown, but studies suggest that approximately 9% of hospitalized patients on specialty services in tertiary care have FD [15]. Modern study began in 1951 with an article by Asher [1], who coined the term "Munchausen syndrome" to describe a subtype of FD. The name came from a German baron who liked to embellish stories of his military exploits to impress listeners. Subsequently, lying and Munchausen were connected in German culture [16]. A series of patients whose medical histories consisted mainly of falsehoods and who visited and deceived hospitals and doctors have been described in the medical literature [1].

People with Munchausen syndrome go to incredible lengths to appear sick or to make themselves sick. For example, a young woman surreptitiously ingested laxatives to cause diarrhea, hypokalemia, and weight loss, and another young woman feigned cancer with the help of fabricated medical records in order to receive chemotherapy [17]. Those with Munchausen syndrome who have medical training are especially convincing. Others read up on diseases so they can mimic symptoms accurately [18]. Some become so proficient that they can fool doctors into ordering needless tests and even operations in some cases. Munchausen syndrome tends to be chronic and people with it usually become habitual deceivers [19]. Savino and Fordtran [17] proposed that it was likely that many cases of factitious cancer were never discovered and predicted that with advances in computer technology the quality of forged medical records would improve in the future. This is shown by the recent case of an individual using medical images obtained from the Internet to support claims of ankle dislocation [20].

Munchausen Syndrome by Proxy and Fabricated or Induced Illness

In 1977, Munchausen syndrome by proxy (MSP)—also known as factitious disorder by proxy—was first described by Meadow [2]. One mother had poisoned her toddler with excessive quantities of salt and another mother had introduced her own blood into her baby's urine sample. MSP refers to a parent or other adult caretaker who repeatedly seeks medical attention for their children, whose symptoms they have faked or induced, sometimes causing real harm to the child, and/or subjecting them to unnecessary investigations and interventions. Many hypotheses have been proposed to explain MSP. Some have noted that patients with the condition often present traumatic events—particularly abuse and deprivation and numerous hospitalizations in childhood—and as adults may have lacked support from family and friends. Others consider that it allows patients to feel in control as they never felt in childhood [21]. Savino and Fordtran suggested that it might represent the patient's attempt to cope with emotional distress [17].

In the United Kingdom, MSP is now termed "fabricated or induced illness" (FII) [22], although MSP is still widely used in other countries. The British Paediatric Surveillance Unit (BPSU) carried out a study of FII cases and identified 97 cases in the United Kingdom over a two-year period. This equated to 1 child in every 5000 being affected by FII, but it is likely that this figure underestimates the true scale of the problem. Another study estimates that the number of children affected by FII could be as high as 1 in 1100 [22].

Munchausen by Internet and Munchausen by Proxy by Internet

In 1983, the case of "Joan and Alex" shocked users of a CB radio channel of the national network, CompuServe, when a male psychologist (Alex) masqueraded as a disabled woman (Joan) in an attempt to use the trust and intimacy attained in the online interactions for his own social gain [23]. In 2000, Munchausen by Internet was identified by Feldman [3] to describe an individual seeking attention by playing out a series of dramatic near-fatal illnesses and recoveries that increasingly strain credulity. A Google search for the term yields more than 7000 search results. There is also an extensive Wikipedia entry [24] that has been revised 285 times since its creation in 2007—an average of 61.5 edits per year—suggesting that the term has now comfortably entered the online vernacular of Internet users. Munchausen by Internet can have devastating effects within online discussion groups, destroying trust when the hoax is exposed [3]. The virtual communities that were created to give support, as well as general non-medical communities, often express genuine sympathy and grief for the purported victims. However, when fabrications are suspected or confirmed the ensuing discussion can create schisms, destroying some communities and altering the trusting nature of members in others.

As yet, Munchausen by Internet has not been officially recognized by the American Psychiatric Association in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). The DSM-5 is currently in review, but makes no mention of Munchausen by Internet although FD is listed in two proposed subtypes: (1) factitious disorder imposed on self and (2) factitious disorder imposed on another. However, these entries make no reference to the Internet in the diagnoses [25]. A generic search on the DSM-5 review website for the term "Internet" locates one relevant mention in illness anxiety disorder under somatic symptom disorders, but this does not mention FD. One of the diagnostic criteria is that the person "performs related excessive behaviors (eg, checking one's body for signs of illness, repeatedly seeking information and reassurance from the Internet or other sources)..." [26]. The DSM-5 review proposes minor modifications to factitious disorders under the umbrella of somatic symptom disorders in their own chapter. The most important modification is the elimination of the distinction between factitious disorders involving predominantly physical versus psychological symptoms. Also, factitious disorder by proxy is now termed "factitious disorder imposed on another" [25], with the perpetrator receiving the diagnosis, not the victim.

New cases of Munchausen by Internet are identified regularly. Recent examples demonstrate the versatility of using online formats for FD, including the adoption of multiple personae and the substantial time and effort expended to contrive complex and dramatic fake identities and conditions [27]. One case documented a brother-sister dyad who created an elaborate narrative to lure a woman into providing time and attention under false pretences. Elizabeth, the victim, had multiple sclerosis and was seeking online support through a social networking website hoping to connect with others coping with chronic illness [27]. In another case, after being diagnosed with a chronic illness, Helen sought to better understand her rare condition by turning to Internet resources for more information [27]. Her research led her to discover an online support community and she joined a support group designed specifically for people with her disease. Helen created a number of fake personae: "Isabelle" (Helen's good friend), "Justin" (her boyfriend), and Justin's father and sister to carry on the story after "Justin" passed away. Helen also fabricated two other storylines including one that featured an ailing mother who had tragically lost two children to illness. Another case concerned a 44-year-old woman who said she had been diagnosed with chronic myeloid leukemia. Had the documents presented by her not aroused suspicion, she would have undergone a bone marrow aspirate and chemotherapy [28].

Munchausen by Internet has also expanded to MSP on the Internet [29]. In 2009, Emily McDonald was arrested for injuring her daughter, Dakota, who had been in and out of hospital since her premature birth. When Dakota did not recover in hospital and her blood cultures showed odd results, staff became suspicious and set up a camera in her hospital room. On video, they caught McDonald putting fecal matter into Dakota's feeding tube. McDonald's case appeared to be MSP; however, she was also posting about her daughter and her illness on her own blog (no longer publically accessible). Although not diagnosed with MSP, she admitted to second-degree injury to a child and was sentenced to 20 years in jail [30].

Munchausen by Internet

Negative Impact from Health-Related Online Support

Debate has intensified on how the increasing use of Web technology might have a negative impact on patients, caregivers, and practitioners. For example, there can be a high ratio of false or irrelevant information compared to useful information on the Web. Eysenbach et al [31] systematically reviewed studies of health website evaluations and found that the most frequently used quality criteria included accuracy, completeness, and technical criteria related to site design (eg, visual appeal, layout, and readability). In their review, the authors noted that in 70% of the studies they had examined, the quality of health-related Web content was found to be low according to the authors of the reviewed studies. The incidence of false data in online self-help groups is unknown, although assumed to be high because of the absence of group rules and guidelines and few controls to prevent people from posting erroneous or off-topic information. Joinson [32] noted that the format of a childcare email list, which seemingly encouraged venting and the name calling of parents and children amidst unconditional support of other caregivers, created an environment which led to the acceptance of practices that were not child-centered and were potentially damaging to the children. This was because the legitimization of negative attitudes and approaches could have led to the continuation of these behaviors. The online "pro-anorexia" underground is a movement that supports those with anorexia and adopts an anti-recovery perspective on the disease [33]. While encouraging a non-healthy diet to sustain an anorexic lifestyle, the movement also recommends the radical use of weight-loss pharmaceuticals—conventionally used to treat obesity—to pursue and maintain low body weight. There are similar movements in other online disease communities, such as supporters of chronic fatigue syndrome who advise abundant rest and avoidance of activity for sufferers, which is in direct opposition to medical advice [34]. Finally, the Internet may also play a major role in the development and spread of beliefs that are unsupported by scientific evidence. For example, the spread of information about Morgellons disease on the Internet has led to several cases of delusional parasitosis [35]. In response, Vila-Rodriguez and MacEwan [36] recommended in a letter to the American Journal of Psychiatry that an awareness of the capacity of the Internet to enable and spread shared delusional ideation was essential to current medical practice.

Review of Recent Munchausen by Internet Literature

We conducted a review of Munchausen by Internet literature over the previous two years by using both academic and social media sources. Searches using the term "Munchausen by Internet" were conducted on both PubMed and Google Scholar. In 2012, there were 8 published articles listed on Google Scholar, but none were relevant because they either contained citations which referred to old research or had no specific link between Munchausen syndrome and the Internet. The Technorati search engine was used in addition to search the blogosphere for recent blog postings or blogs on the terms "Munchausen by Internet" and "Munchausen" generated during the first half of 2012, but no relevant blogs or posts were identified.

Motivation

Because many instances of Munchausen by Internet take place in a group situation, social psychology offers a number of theories that can be applied to explain this type of online behavior. Drawing on disinhibition theory, Suler [37] highlighted two features of the Internet that made deception easier: (1) asynchronicity allows a dynamic approach to identity presentation and enables quick changes between identities and styles, and (2) the lack of feedback and the anonymity or unfamiliarity of the audience can reduce concern for others' views. Similarly, Taylor and MacDonald [38] applied the theories of de-individuation and social identity to explain more uninhibited behavior and more self-disclosure in some online settings. Drawing on motivations for deception and group effects, Mealy et al [39] found that lies motivated by a desire to benefit others were considered to be more acceptable than lies that primarily benefited the self. Additionally, lying to the *out-group* (the social group to which someone does not identify) was perceived as being more acceptable than lying to the *in-group* (the social group to which someone psychologically identifies themselves with as a member).

Other areas of psychology offer theories of relevance, for example those relating to self-presentation and identity. Research has shown a link between low self-esteem and the need for popularity to the way individuals manipulate relationships and the way they perceive others' online relationships [40]. Walther [41] noted there is a propensity for disinhibition, projection, and transference, wherein there are no visual or auditory inputs that can place the text in its proper context or assist the correct interpretation of that text. People fill in the missing pieces in the picture of others they meet online, not fully aware that the picture they are forming is based partly on their own unconscious desires regarding who they want that person to be and how they want them to act. This occurs at the same time as the person is taking advantage of the anonymity inherent in text-only communications to present their best possible face. A feedback loop can arise as these selective presentations are responded to in-kind, creating a hyperpersonal aspect to Internet communications. The Hyperpersonal Model [42] proposes that the Internet affects three parts of the communication process. These can be applied to understand the way in which Munchausen by Internet users manipulate the Internet to manage impressions and also to explain how others make interpersonal impressions based on that information. The three parts of the communication process are:

1. Receivers who have an idealized perception of the message sender because subtle context cues take on a stronger value in online communication. An absence of face-to-face cues means that receivers may be acutely sensitive to any subtle social or personality cues, so partners build impressions of one another based upon minimal cues.

2. Message senders have a greater opportunity to optimize their self-presentation to others and themselves.

3. The asynchronous channel allows more time for senders and receivers to consider the messages they send and receive, so that only using text can create an idealized picture.

It is possible that some Munchausen by Internet sufferers could be driven by the simple enjoyment of online deception as highlighted by a study of Web users that found that most online deceivers felt a sense of enjoyment while engaging in online deception [43]. In 1999, a columnist was introduced in the UK newspaper, *The Observer*, following his aborted suicide attempt and detailing the columnist's last few months before he killed himself [44]. After protest, it was revealed that the column was a spoof. Chris Morris, an innovative but controversial British broadcaster, wrote the columns that highlighted the inherent cruel, dark comedy of fooling gullible members of the public with stories of fake medical conditions. This willingness of broadsheet media to embrace and portray a borderline style of comedy can have unpleasant outcomes within an Internet environment. One clue might lie in the power granted by online communities to quantify the sympathy for an illness or the shock of a death through comments boxes or replies to a journal thread [13]. During a lengthy battle against terminal illness, blog writers can attract support from thousands of friends who follow them through treatments and who become emotionally involved when they die. In more than one example, bereaved online friends have created tribute websites where they have posted poetry and photographs in memorial books. This can feed the desire of a narcissist and as they create an imaginary online long-term condition leading to a fake death, can provide a lonely individual with attention that they may never have previously known. This view is supported by Feldman [3], who linked the engaging and intense nature of these deceptions to sadism. This motivation can be seen to be at both intrapersonal (sadistic) and interpersonal (attention) levels.

Stokes [45] argued that online social networks offer more methods to manage the impressions of others than are available through structured websites. He referred to a study that found Facebook users' identities were not the identities that they had established in the offline world, nor were they close to the identities that they had constructed in other anonymous online environments. They were the hoped-for possible identities users wanted to have in the offline world, but had not yet been able to establish.

There is little research regarding the psychology of Munchausen by Internet, but there are indications that some online self-presentation may be motivated by narcissistic or sadistic tendencies, as is the case with FD. In one analysis of patients with FD, it was found that 9 patients (50%) had borderline personality disorder, 6 (33%) had narcissistic personality disorder, and only 3 patients (17%) did not demonstrate coexisting self-pathology [15]. A study of Internet dating sites by Ellison et al [46] found that people acted differently in social networking environments depending on whether or not they were interacting anonymously. This finding has important implications for understanding identity in the online world because it indicates that online self-presentation varies according to the nature of the online setting. Oltmanns [47] described narcissism as a pervasive pattern of grandiosity, a need for admiration, and an exaggerated sense of self-importance. Mehdizadeh [48] associated the term with positive self-views of agentic traits, including intelligence, physical attractiveness, and power. Central to most theoretical models of narcissism, the use of social relationships is employed to regulate narcissistic esteem. However, narcissists do not focus on interpersonal intimacy, warmth, or other positive aspects of relational outcomes. Instead, they use relationships to appear popular and successful, and they seek attractive, high-status individuals as romantic partners [49]. Despite their tendency to seek out many superficial, empty relationships, narcissists rarely pursue these commitments for long periods of time. Relationships are pursued solely when an opportunity for public glory presents itself [49].

Netiquette and Trolls

The aspect of Munchausen by Internet where individuals actively seek to disrupt and cause problems for their own satisfaction or enjoyment (sadism versus narcissism) has until now not been consciously associated with the wider phenomena of Internet "trolls." We believe that this connection should be formally acknowledged to assist in controlling, effectively identifying, and minimizing the growth of this behavior as more people seek reassurance and support about their health in an online environment. Netiquette (short for "network etiquette") is the dos and don'ts of online communication covering common courtesy online and the informal "rules of the road" of cyberspace [50]. A *troll* is someone who posts or sends messages online with the intention of annoying someone or disrupting a discussion or environment [51]. The practice of trolling has been compared to the fishing term in which a line is set in the water and the bait is dragged slowly back and forth in the hope of getting a bite [52].

Donath [52] outlined the ambiguity of identity in a disembodied virtual community and provided a concise overview of identity deception games, which trade on the confusion between physical and epistemic communities. Trolling has been portrayed disdainfully in mainstream media outlets, often referencing the willingness of some Internet trolls to go to extreme lengths in their attempts to elicit reactions. In 2010, the Australian government became involved after trolls defaced the Facebook tribute pages of two murdered children. The Australian Minister of Communications decried the attacks as evidence of the need for greater Internet regulation [53]. In the wake of these events, Facebook responded by strongly urging administrators to be aware of ways to ban users and remove inappropriate content from Facebook pages. It is recommended that ignoring a troll is almost always the best approach, because if nobody responds the troll will eventually get bored and go away [51]. Experienced participants in online forums know that responding tends to encourage trolls to continue disruptive postings—hence the oft-seen warning, "do not feed the troll" (DNFTT). However, experts tend to inhabit the tougher, streetwise environs of Internet technical or film forums rather than in the supportive and empathetic environment of an online health support group, which can cause more of a shock when Munchausen by Internet trolls are unmasked.

Opportunity

Recupero [54] highlighted instances of psychiatric patients engaged in impression management to influence the outcomes of psychiatric interviews [55]. Impression management also plays an important role in online conversations. Barak [56] believes that many Internet users prefer others to perceive and interact with an online persona and that material chosen to post online can help to deepen a preferred social impression. Conversely, profiles posted on social networking sites (eg, Facebook) may contain information contradicting the evaluee's intended impression. Photographs, perhaps artificially composed, and other material can be posted and tagged online showing a person's name or identity without their knowledge or permission regardless of whether the person is familiar with or naïve about Internet influence. Similarly, records of negative behavior can remain online for years becoming part of an individual's digital footprint [57].

Computers allow people with sufficient technical skills to access medical records and use them to falsify medical histories [20], while the open trusting environments of communication forums—established for the sole purpose of giving support to members facing significant health or psychological problems—are easily infiltrated because of the social nature of the groups. This is an endemic problem with online communication as the Internet helps to break down the physical barriers that assist in preventing the spread of lies [58]. The proliferation of newsgroups and chat rooms offers a limitless audience for fake narratives with people able to move from one support group to another [18]. Some, pretending to be ill, have joined more than one, and some might sign on to a single group multiple times by using different names and acting out different roles. A *sock puppet* is an online identity used for purposes of deception. The term, a reference to the manipulation of a hand puppet made from a sock, originally referred to a false identity assumed by a member of an Internet community who spoke to, or about, himself while pretending to be another person, but it now includes other uses of misleading online identities, such as those created to praise, defend, or support a third party [59]. The development of FD in online groups and forums is made easier by the anonymous and malleable nature of online identity along with easy access to the Internet, which allows sock puppets to thrive without any negative consequences to themselves. For example, Andrea, a 40-year-old single mother, began posting on an ovarian cancer forum that she had concerns over her worsening abdominal pain. Shortly after joining the forum, Andrea announced that she had been diagnosed with ovarian cancer. After being confronted by suspicious group members, Andrea confessed to her deception and the use of sock puppets. In addition to playing herself during her illness and online death, she had posed as her daughter "Brittney" and her daughter's boyfriend "Chris" [27].

Detection

Although some Munchausen by Internet perpetrators display a remarkable degree of research and endurance, able to stretch the fiction over many months, even the most dedicated can slip up eventually. Outwardly, there might not be any clues to suggest that they are anything other than normal support group members, but slight details can introduce contradictions and, although anyone caught up in the deception may be willing to forgive a slight oversight, some are more methodical. They are prepared to investigate so that few contradictions will escape over time. Savino and Fordtran [17] suggested that diagnosis of factitious cancer is usually made by detection of inconsistencies in medical history, by evidence of fabrication of medical records, by lies patients tell about their health insurance, or by doctors who begin to doubt the patient's story. A story of prolonged survival with a usually lethal cancer has helped reveal factitious cancer in some cases. Based on his experience with Munchausen by Internet, Feldman [3] listed some methods of detection:

1. Posts consistently duplicating material in other posts, books, or health-related websites.

2. Characteristics of the supposed illness emerging as caricatures.

3. Near-fatal bouts of illness alternating with miraculous recoveries.

4. Fantastical claims, contradicted by subsequent posts, or flatly disproved.

5. Continual dramatic events in the person's life, especially when other group members have become the focus of attention.

6. Feigned blitheness about crises that will predictably attract immediate attention.

7. Others apparently posting on behalf of the individual having identical patterns of writing.

Griffiths et al [20] recommended that all clinicians question histories that did not match examination findings, ensuring that all radiographs were adequately labeled with patient-specific information, and being aware of radiographic inconsistencies.

There are several strategies commonly employed in confronting FD. In one case series, patients were carefully confronted with the factitious nature of their illness. Although 13 of the 33 (approximately 39%) admitted feigning illness, most of the patients' illnesses improved following this strategic confrontation [15]. In most cases, group members' discovery of Munchausen by Internet can lead to a similar strategic confrontation [3,27] with the typical response being a protest of innocence and an allegation of mistreatment by the group, followed by disappearance [3]. Due to the elusive nature of online identities, most wronged group members are unable to pursue the fakers. Many either lock their journals so that only their friends can access them, or else purge them entirely and deny the fraud. Suspicious group members are sometimes able to take screenshots as evidence of the fraud, but many perpetrators slip away, either sufficiently chastened to stop the deceptions or to simply reappear in another online group.

Effects and Consequences

The Munchausen by Internet troll can be costly in several ways. A troll can disrupt the discussion on a newsgroup, disseminate bad advice, and damage the feeling of trust in a Web community. Furthermore, in a group that has become sensitive to trolling—where the rate of deception is high—some honest but naïve question can be quickly rejected as trolling. This can be off-putting to a new user who is immediately bombarded with accusations when venturing a first post. Even if an accusation is unfounded, being branded as a troll can be damaging to an online reputation. Herring et al [60] discussed the difficulty inherent in monitoring trolling and maintaining freedom of speech in online communities, concluding that inevitably harassment was more likely to occur in environments where lack of censure was a key factor. In wider discussion forums, the broadly accepted ethic of free speech may lead to tolerance of trolling behavior, further complicating the members' efforts to maintain an open-yet-supportive discussion area, especially for sensitive topics. Reactions from Internet forums have been critical of media portrayals of trolls, stating that trolling is nothing new and has become part of accepted Internet culture. While not condoning the viciousness of troll attacks, forum discussions regularly express concern that mainstream media coverage of trolling ultimately results in more trolling because widespread attention represents the reaction that trolls seek. However, in the insular, empathic environment of a support group, the reaction can be more severe. Grady [18] detailed the case of a 15-year-old girl who communicated with members of a virtual support group for parents of babies who were critically ill. Kim claimed to have a baby requiring treatment and as she detailed the timeline of her baby's treatment, other members of the group became personally involved and were devastated when she said her baby had died. Subsequently, she appeared online saying she was pregnant again and that she feared the second baby would be born prematurely. After the early birth of her second baby, the same cycle was reenacted, but this time a group member (a psychologist and the mother of twins born prematurely) gradually became suspicious. She confronted Kim, who subsequently posted a confession and apology. Kim withdrew from the group and was taken off the list by its owner, but her behavior had a negative effect on a group who had been trusting and close-knit until then [18]. Some parents expressed feelings of betrayal, and many stopped posting messages. People in the group agreed to provide information so that a coordinator could verify they were really parents of premature babies. Some new participants were put off by the atmosphere of suspicion, but the group gradually bounced back. However, those who encountered Kim would obviously never view new postings in quite the same way again.

Application of the term “troll” is highly subjective. Some readers may characterize a post as trolling, whereas others might regard the same post as a legitimate discussion contribution, even if controversial. Sometimes the term is used as a strategy to discredit an opposing position by attacking its proponent. Calling someone a troll makes assumptions about a writer's motives. Regardless of the circumstances, controversial posts may attract a particularly strong response from those unfamiliar with the robust dialogue found in some online, rather than physical, communities. The popularity of Facebook means strangers can often build and maintain relationships entirely on the Internet. This anonymity permits any number of lies to be accepted as truth, and inventive deceptions, with whatever motive, can be carried all the way to an online grave. However, there is also the potential danger that overzealous group members might make the erroneous assumption that every death encountered online is fake. Most social network sites now allow relatives of deceased users to choose to keep profiles online as a memorial, allowing users to post tributes and messages, sometimes speaking of the dead in the third person, sometimes in the second person. In effect, a profile site is converted into a tribute site, a space of commemoration—a sort of open-ended electronic wake [45]. In the case of a real death, it can be just as harrowing for the bereaved to read comments claiming that the death has been faked.

Conclusions

More research is required to be able to provide evidence-based advice to victims of suspected Munchausen by Internet trolls and for facilitators of discussion groups to effectively manage interactions. As this is one area of the literature that does not yet yield much information, one methodology that could possibly be adopted for further study is based on qualitative content analysis (QCA), which is gaining much support in studies of social interactions in online support communities [61]. QCA provides a way to study manifest and latent content within a body of text. Analysis of what the text says describes the visible components, referred to as the *manifest content*. Analysis of what the text talks about involves interpreting the underlying meaning of the text, referred to as the *latent content*. Therefore, manifest content might highlight descriptions of illnesses, while latent content could include descriptions of feelings of sufferers and the context of an illness. The manifest content is usually presented in categories, whereas the latent content is expressed as themes. QCA differs from pure qualitative research as it allows the researcher to emphasize differences between and similarities within codes and categories. Therefore, it could be used to differentiate types of Munchausen by Internet and its different motives.

When Munchausen by Internet seems likely, it might be practical to have some established group members gently question any dubious post owner privately. Although the typical response is vehement denial regardless of the strength of the evidence, the author typically will disappear from the group. In some Munchausen by Internet cases, much like FD cases, individuals can be both perpetrators and patients. For example, the previously described case of Helen who was diagnosed with a chronic illness and went on to create a number of fake personae [27]. Savino and Fordtran [17] suggested some useful steps for confronting patients suspected of FD:

1. Let the patient know what you suspect but without outright accusation.

2. Support the suspicion with facts.

3. Provide empathetic and face-saving comments.

4. Avoid probing to uncover the patient's underlying feelings and motivations so as to minimize disruption of emotional defences that are essential to her function.

5. Assure the patient that only those who need to know will be informed of the suspicion of factitious disease.

6. Make sure the staff demonstrate continued acceptance of the patient as a person worthy of their help.

7. Encourage psychiatric help, but if the patient resists do not force the issue.

However, once a perpetrator has been confronted, remaining members of the online group may need psychological help at an individual or group level depending on the extent of the deception and the health topic concerned. For example, individuals may need help in processing their feelings. As a group, help could be directed toward dealing with conflict and blame, and moving forward to refocus the group on its original goals with the aim of protecting and encouraging the original sense of trust. As Whitworth and de Moor [62] suggest, laws in a physical community are expressed in terms of physical actions and concrete objects that govern what people do, not what they think or feel. Historical law assumes a physical world constrained by time and space, but virtual environments have significantly different functionality. This means the virtual world is a functionally different world; it may not be appropriate, or even possible, to transfer laws from the physical to the electronic world. Therefore, laws must be re-invented by re-applying legitimacy concepts to virtual contexts such as formulating direct policy to protect health information users in the new world of Internet-based health searching and support. Legal sanctions have yielded some results in controlling the "acting out" of Munchausen's syndrome, but the literature suggests that such measures are ineffective and can sometimes even reinforce bad behavior [63]. It is suggested that the best results within a physician-patient relationship can be achieved by approaching the dilemma from diverse angles [28].

Historically, the consequences for perpetrators have tended to be minor because few can be pursued or punished unless the wronged individuals are able to prove that the perpetrators have committed an illegal act. However, a 2005 legal case concerning self-help members pursuing an online campaign against a Munchausen by Internet member who challenged posts as defamatory suggests that wronged individuals are able to respond without fear of successful legal reprisal. They might also be able to win a precedent-setting civil case [64]. Indeed, there is a strong case for considering the sadistic misuse of health-related forums as a form of cybercrime, rather than as an everyday negative risk of using the Internet that must be tolerated and accepted. Consider if a malicious user deliberately (or accidentally) gave out medical information that resulted in a worsening of health or had fatal consequences. Internet protocol (IP) addresses of Munchausen by Internet trolls could be identified and Internet service providers (ISPs) could be enlisted to help identify and "out" frequent perpetrators such as been seen in recent online copyright disputes. Social network providers, such as Facebook, should tighten up their own procedures or, as an alternative, group users might want to consider relocating to more private group-based Internet communities, such as private Wikis [51]. Although these do not have the same large population of users, they might increase the security and lessen the chance of encountering a Munchausen by Internet troll online.

Enhanced self-regulation is the most positive action to reduce group risk. It might also be advisable for a health support group to identify a gatekeeper. Although adding extra layers of security and formality before a user could post might be viewed as onerous, the long-term benefits might be worth the additional effort. Facilitators could also clearly state to all members that although most people participating in support groups are honest, all members should balance their empathy with some degree of circumspection. Group members should be especially careful about basing any of their own health care decisions on uncorroborated information supplied in groups, just as they should with any other source on the Internet [65].

We have reviewed potential motivations and consequences for Munchausen by Internet behaviors, but it is clear that further research is necessary to investigate the psychology and methods of coping with Munchausen by Internet. However, there is a clear, compelling need to recognize that in addition to Munchausen by Internet being classed as a condition in its own right, there is a subset of people currently tagged as Munchausen by Internet sufferers who are actually Munchausen by Internet trolls purposefully harming well-intentioned support groups and abusing members for their own pleasure or enjoyment. We propose that Munchausen by Internet and Munchausen by Internet trolling be formally acknowledged in a revised version of DSM-5 (within the factitious disorder revisions), and that this sphere of behavior needs wider consideration and action, either by group users or by the creators of the host software. As Berners-Lee [66] said, "Technologists cannot simply leave the social and ethical questions to other people, because the technology directly affects these matters."

Acknowledgments

Publication was made possible by funding from the Bournemouth University Open-Access Publishing Fund (OAPF).

Authors' Contributions

All of the authors are responsible and qualified for the reported research. They have all participated in the concept and design, analysis, and interpretation of data, and drafting and revision of the manuscript and approve the manuscript as submitted.

Conflicts of Interest

None declared.

References

1. Asher R. Munchausen's syndrome. Lancet 1951 Feb 10;1(6650):339-341. [Medline: 14805062]
2. Meadow R. Munchausen syndrome by proxy. The hinterland of child abuse. Lancet 1977 Aug 13;2(8033):343-345. [Medline: 69945]
3. Feldman MD. Munchausen by Internet: detecting factitious illness and crisis on the Internet. South Med J 2000 Jul;93(7):669-672. [Medline: 10923952]
4. PsychNet-UK. Malingering URL: http://www.psychnet-uk.com/x_new_site/DSM_IV/malingering.html [accessed 2011-11-15] [WebCite Cache ID 63DXBCUTe]
5. Resnick PJ. Malingering of posttraumatic disorders. In: Rogers R, editor. Clinical Assessment of Malingering and Deception, Third Edition. New York: The Guilford Press; 1997:130-152.

6. Lurid FB. Galen on malingering, centaurs, diabetes, and other subjects more or less related. In: Proceedings of the Charaka Club, Volume X. Baltimore: Williams & Wilkins Co; 1941:52-55.
7. Field MG. Structured strain in the role of the Soviet physician. The American Journal of Sociology 1953;58(5):493-502.
8. Doescher MP, Saver BG, Franks P, Fiscella K. Racial and ethnic disparities in perceptions of physician style and trust. Arch Fam Med 2000 Dec;9(10):1156-1163. [Medline: 11115223]
9. Boehm FH. Building trust. Family Practice News 2003;33(15)(12(11)).
10. Piette JD, Heisler M, Krein S, Kerr EA. The role of patient-physician trust in moderating medication nonadherence due to cost pressures. Arch Intern Med 2005;165(15):1749-1755. [doi: 10.1001/archinte.165.15.1749] [Medline: 16087823]
11. Pearson SD, Raeke LH. Patients' trust in physicians: many theories, few measures, and little data. J Gen Intern Med 2000 Jul;15(7):509-513. [Medline: 10940139]
12. Montague EN, Winchester WW, Kleiner BM. Trust in medical technology by patients and health care providers in obstetric work systems. Behav Inf Technol 2010 Sep;29(5):541-554. [doi: 10.1080/01449291003752914] [Medline: 20802836]
13. Swains H. Wired UK. 2009 May 25. Q&A: Munchausen by internet URL: http://www.wired.co.uk/news/archive/2009-03/24/qa-munchausen-by-internet [accessed 2011-11-15] [WebCite Cache ID 63DQh4Nsq]
14. Gavin H. On feigned and factitious diseases, chiefly of soldiers and seamen; on the means used to simulate or produce them, and on the best modes of discovering imposters: Being the Prize Essay in the Class of Military Surgery, in the University of Edinburgh, 1835-6, with Additions. London: John Churchill; 1843.
15. Alinejad NA, Oettel DJ. Factitious disorder as repeated diabetic ketoacidosis: a case report. Innov Clin Neurosci 2011 Feb;8(2):41-47. [Medline: 21468297]
16. Lipsitt DR. Introduction. In: The Spectrum of Factitious Disorders. Washington, DC: American Psychiatric Press; 1996:xix-xxviii.
17. Savino AC, Fordtran JS. Factitious disease: clinical lessons from case studies at Baylor University Medical Center. Proc (Bayl Univ Med Cent) 2006 Jul;19(3):195-208. [Medline: 17252033]
18. Grady D. NY Times. 1998 Apr 23. Faking pain and suffering in internet support groups URL: http://www.nytimes.com/1998/04/23/technology/faking-pain-and-suffering-in-internet-support-groups.html [accessed 2011-11-15] [WebCite Cache ID 63DPqDD3i]
19. Doherty AM, Sheehan JD. Munchausen's syndrome--more common than we realize? Ir Med J 2010 Jun;103(6):179-181. [Medline: 20669602]
20. Griffiths EJ, Kampa R, Pearce C, Sakellariou A, Solan MC. Munchausen's syndrome by Google. Ann R Coll Surg Engl 2009 Mar;91(2):159-160. [doi: 10.1308/003588409X391938] [Medline: 19317939]
21. PsychNet-UK. Factitious disorder by proxy URL: http://www.psychnet-uk.com/x_new_site/DSM_IV/factitious_disorder_by_proxy.html [accessed 2012-08-03] [WebCite Cache ID 69du1qGmR]
22. NHS Choices. Fabricated or induced illness URL: http://www.nhs.uk/conditions/fabricated-or-induced-illness/Pages/Introduction.aspx [accessed 2011-12-23] [WebCite Cache ID 648yKJ5k8]
23. Turkle S. Life on the Screen: Identity in the Age of the Internet. New York: Simon & Schuster; 1995.
24. Wikipedia. Münchausen by Internet URL: http://en.wikipedia.org/wiki/Cybermunch [accessed 2012-01-24] [WebCite Cache ID 64va6BCyD]
25. American Psychiatric Association. American Psychiatric Association DSM-5 Development. Proposed Revisions/Other Disorders/V 06 Factitious Disorder/Factitious Disorder URL: http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=386 [accessed 2012-02-01] [WebCite Cache ID 6586wzAFq]
26. American Psychiatric Association. American Psychiatric Association DSM-5 Development. Proposed Revisions/Somatic Symptom Disorders/J 02 Illness Anxiety Disorder URL: http://www.dsm5.org/proposedrevision/pages/proposedrevision.aspx?rid=10 [accessed 2012-02-01] [WebCite Cache ID 6588Rwt9X]
27. Cunningham JM, Feldman MD. Munchausen by Internet: current perspectives and three new cases. Psychosomatics 2011 Apr;52(2):185-189. [doi: 10.1016/j.psym.2010.11.005] [Medline: 21397112]
28. Caocci G, Pisu S, La Nasa G. A simulated case of chronic myeloid leukemia: the growing risk of Munchausen's syndrome by internet. Leuk Lymphoma 2008 Sep;49(9):1826-1828. [doi: 10.1080/10428190802179889] [Medline: 18608864]
29. McCulloch V, Feldman MD. Munchausen by proxy by Internet. Child Abuse Negl 2011 Nov;35(11):965-966. [doi: 10.1016/j.chiabu.2011.05.020] [Medline: 22074757]
30. Daily Mail Reporter. Mail Online. 2011 Apr 05. Mother caught smearing faeces on her sick three-year-old daughter's IV is jailed for 20 years URL: http://www.dailymail.co.uk/news/article-1373500/Emily-Beth-McDonald-jailed-20-years-smearing-faeces-sick-daughters-IV.html [accessed 2011-12-23] [WebCite Cache ID 648zY8KcG]
31. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700. [Medline: 12020305]
32. Joinson A. Understanding the Psychology of Internet Behaviour: Virtual Worlds, Real Lives. New York: Palgrave Macmillan; 2003.
33. Fox N, Ward K, O'Rourke A. Pro-anorexia, weight-loss drugs and the internet: an "anti-recovery" explanatory model of anorexia. Sociol Health Illn 2005 Nov;27(7):944-971. [doi: 10.1111/j.1467-9566.2005.00465.x] [Medline: 16313524]

34. Wright B, Partridge I, Williams C. Management of chronic fatigue syndrome in children. Advances in Psychiatric Treatment 2000;6:145-152. [doi: 10.1192/apt.6.2.145]
35. Lustig A, Mackay S, Strauss J. Morgellons disease as internet meme. Psychosomatics 2009;50(1):90. [doi: 10.1176/appi.psy.50.1.90] [Medline: 19213978]
36. Vila-Rodriguez F, Macewan BG. Delusional parasitosis facilitated by web-based dissemination. Am J Psychiatry 2008 Dec;165(12):1612. [doi: 10.1176/appi.ajp.2008.08081283] [Medline: 19047336]
37. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004 Jun;7(3):321-326. [doi: 10.1089/1094931041291295] [Medline: 15257832]
38. Taylor J, MacDonald J. The effects of asynchronous computer-mediated group interaction on group processes. Social Science Computer Review 2002;20(3):260-274. [doi: 10.1177/089443930202000304]
39. Mealy M, Stephan W, Urrutia IC. The acceptability of lies: a comparison of Ecuadorians and Euro-Americans. Int J Intercultural Relations 2007;31(6):689-702.
40. Utz S, Beukeboom CJ. The role of social network sites in romantic relationships. Journal of Computer-Mediated Communication 2011;16(4):511-527. [doi: 10.1111/j.1083-6101.2011.01552.x]
41. Walther JB. Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction. Communication Research 1996;23(1):3-43. [doi: 10.1177/009365096023001001]
42. Walther JB. Selective self-presentation in computer-mediated communication: hyperpersonal dimensions of technology, language, and cognition. Computers in Human Behavior 2007;23(5):2538-2557. [doi: 10.1016/j.chb.2006.05.002]
43. Caspi A, Gorsky P. Online deception: prevalence, motivation, and emotion. Cyberpsychol Behav 2006 Feb;9(1):54-59. [doi: 10.1089/cpb.2006.9.54] [Medline: 16497118]
44. Geefe R. The Observer. 1999 May 30. Time to go URL: http://www.guardian.co.uk/theobserver/1999/may/30/featuresreview.review1 [accessed 2011-11-15] [WebCite Cache ID 63DQUzje9]
45. Stokes P. Ghosts in the machine: do the dead live on in Facebook? Philosophy & Technology 2011;25:363-317. [doi: 10.1007/s13347-011-0050-7]
46. Ellison N, Heino R, Gibbs J. Managing impressions online: self-presentation processes in the online dating environment. Journal of Computer-Mediated Communication 2006;11:415-441. [doi: 10.1111/j.1083-6101.2006.00020.x]
48. Mehdizadeh S. Self-presentation 2.0: narcissism and self-esteem on Facebook. Cyberpsychol Behav Soc Netw 2010 Aug;13(4):357-364. [doi: 10.1089/cyber.2009.0257] [Medline: 20712493]
47. Oltmanns FT, Emery ER, Taylor S. Abnormal Psychology. Upper Saddle River, NJ: Pearson/Prentice Hall; 2007.
49. Campbell WK. Narcissism and romantic attraction. Journal of Personality and Social Psychology 1999;77:1254-1270. [doi: 10.1037/0022-3514.77.6.1254]
50. Shea V. Netiquette. San Francisco: Albion Books; 1994.
51. Pulman A. Blogs, Wikis, Podcasts and More (Pocket Study Skills). Basingstoke: Palgrave Macmillan; 2009.
52. Donath JS. Identity and deception in the virtual community. In: Communities in Cyberspace. London: Routledge; 1999:29-59.
53. Bachl M, Paget H. 9 News. 2010 Apr 01. Internet without laws a 'recipe for anarchy' URL: http://news.ninemsn.com.au/national/1034842/internet-without-laws-a-recipe-for-anarchy [accessed 2011-11-15] [WebCite Cache ID 63DVLwJ4R]
54. Recupero PR. The mental status examination in the age of the internet. J Am Acad Psychiatry Law 2010;38(1):15-26. [Medline: 20305070]
55. Sherman M, Trief P, Sprafkin R. Impression management in the psychiatric interview: quality, style, and individual differences. J Consult Clin Psychol 1975 Dec;43(6):867-871. [Medline: 1194481]
56. Barak A. Phantom emotions: psychological determinants of emotional experiences on the Internet. In: The Oxford Handbook of Internet Psychology. Oxford: Oxford University Press; 2007:303-329.
57. Madden M, Fox S, Smith A, Vitak J. Digital footprints: online identity management and search in the age of transparency. Washington, DC: Pew Internet & American Life Project; Dec 16, 2007.
58. Johnson B. The Guardian. 2001 May 28. The short life of Kaycee Nicole URL: http://www.guardian.co.uk/technology/2001/may/28/internetnews.mondaymediasection/print [accessed 2011-11-15] [WebCite Cache ID 63DQA0soE]
59. Stone B, Richtel M. New York Times. 2007 Jul 16. The hand that controls the sock puppet could get slapped URL: http://www.nytimes.com/2007/07/16/technology/16blog.html?ex=1342238400&en=9a3424961f9d2163&ei=5088&partner=rssnyt&emc=rss [accessed 2011-12-23] [WebCite Cache ID 6490ktBUe]
60. Herring S, Job-Slider K, Scheckler R, Barb S. Searching for safety online: managing Trolling in a feminist forum. The Information Society 2002;18(5):371-384. [doi: 10.1080/01972240290108186]
61. Pfeil U, Zaphiris P. Applying qualitative content analysis to study online support communities. Universal Access in the Information Society 2009;9:1-16. [doi: 10.1007/s10209-009-0154-3]
62. Whitworth B, de Moor A. Legitimate by design: towards trusted virtual community environments. Behaviour & Information Technology 2003;22(1):31-51. [doi: 10.1080/01449290301783]
63. Elmore JL. Munchausen syndrome: an endless search for self, managed by house arrest and mandated treatment. Ann Emerg Med 2005 May;45(5):561-563. [doi: 10.1016/j.annemergmed.2004.11.034] [Medline: 15855960]

64. Feldman MD, Peychers ME. Legal issues surrounding the exposure of "Munchausen by Internet". Psychosomatics 2007;48(5):451-452. [doi: 10.1176/appi.psy.48.5.451-a] [Medline: 17878508]
65. White R, Horvitz E. Microsoft Research. 2008. Cyberchondria: studies of the escalation of medical concerns in web search URL: http://research.microsoft.com/apps/pubs/default.aspx?id=76529 [accessed 2011-11-15] [WebCite Cache ID 63DTmOnPf]
66. Berners-Lee T, Fischetti M. Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor. San Francisco: HarperSanFrancisco; 1999.

Abbreviations

BPSU: British Paediatric Surveillance Unit
DNFTT: Do not feed the troll
FD: factitious disorder
FII: fabricated or induced illness
IP: Internet protocol
ISP: Internet service provider
MSP: Munchausen syndrome by proxy
QCA: qualitative content analysis

Edited by G Eysenbach; submitted 22.11.11; peer-reviewed by PM Whitty, T Buchanen, M Feldman; comments to author 16.12.11; revised version received 06.03.12; accepted 15.06.12; published 22.08.12.

Please cite as:
Pulman A, Taylor J
Munchausen by Internet: Current Research and Future Directions
J Med Internet Res 2012;14(4):e115
URL: http://www.jmir.org/2012/4/e115/
doi:10.2196/jmir.2011
PMID:22914203

Original Paper

Online Advertising to Reach and Recruit Latino Smokers to an Internet Cessation Program: Impact and Costs

Amanda L Graham[1,2], PhD; Ye Fang[1], MS; Jose L Moreno[1], MA; Shawn L Streiff[3]; Jorge Villegas[4], PhD; Ricardo F Muñoz[5], PhD; Kenneth P Tercyak[2], PhD; Jeanne S Mandelblatt[2], PhD; Donna M Vallone[6], PhD, MPH

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

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

[3]Marketing Department, American Legacy Foundation, Washington, DC, United States

[4]Business Administration, University of Illinois at Springfield, Springfield, IL, United States

[5]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

[6]Research and Evaluation Department, American Legacy Foundation, Washington, DC, United States

Corresponding Author:
Amanda L Graham, PhD
Schroeder Institute for Tobacco Research & Policy Studies
American Legacy Foundation
1724 Massachusetts Ave NW
Washington, DC, 20036
United States
Phone: 1 202 454 5938
Fax: 1 202 454 5785
Email: agraham@legacyforhealth.org

Abstract

Background: Tobacco cessation among Latinos is a public health priority in the United States, particularly given the relatively high growth of this population segment. Although a substantial percentage of American Latinos use the Internet, they have not engaged in Web-based cessation programs as readily as other racial/ethnic subgroups. A lack of culturally specific advertising efforts may partly explain this disparity.

Objective: Phase I of this study focused on the development of four Spanish-language online banner advertisements to promote a free Spanish-language smoking cessation website (es.BecomeAnEX.org). Phase II examined the relative effectiveness of the four banner ads in reaching and recruiting Latino smokers to the cessation website.

Methods: In Phase I, 200 Spanish-speaking Latino smokers completed an online survey to indicate their preference for Spanish-language banner ads that incorporated either the cultural value of family (*familismo*) or fatalism (*fatalismo*). Ads included variations on message framing (gain vs loss) and depth of cultural targeting (surface vs deep). In Phase II, a Latin square design evaluated the effectiveness of the four preferred ads from Phase I. Ads were systematically rotated across four popular Latino websites (MySpace Latino, MSN Latino, MiGente, and Yahoo! en Español) over four months from August to November 2009. Tracking software recorded ad clicks and registrants on the cessation website. Negative binomial regression and general linear modeling examined the main and interacting effects of message framing and depth of cultural targeting for four outcomes: number of clicks, click-through rate, number of registrants, and cost per registrant.

Results: In Phase I, smokers preferred the four ads featuring *familismo*. In Phase II, 24,829,007 impressions were placed, yielding 24,822 clicks, an overall click-through rate of 0.10%, and 500 registrants (2.77% conversion rate). Advertising costs totaled US $104,669.49, resulting in an overall cost per click of US $4.22 and cost per registrant of US $209.34. Website placement predicted all four outcomes (all P values < .01). Yahoo! en Español yielded the highest click-through rate (0.167%) and number of registrants (n = 267). The message framing and cultural targeting interaction was not significant. Contrary to hypotheses, loss-framed ads yielded a higher click-through rate than gain-framed ads (point estimate = 1.08, 95% CI 1.03 1.14, $P = 0.004$), and surface-targeted ads outperformed deep-targeted ads for clicks (point estimate = 1.20, 95% CI 1.13 1.28, $P < .001$), click-through rate (point estimate = 1.22, 95% CI 1.16 1.29, $P < .001$), and number of registrants (point estimate = 2.73, 95% CI 2.14 3.48, $P < .001$).

Conclusions: Online advertising can be an effective and cost-efficient strategy to reach and engage Spanish-speaking Latino smokers in an evidence-based Internet cessation program. Cultural targeting and smoking-relevant images may be important factors for banner ad design. Online advertising holds potential for Web-based cessation program implementation and research.

(J Med Internet Res 2012;14(4):e116) doi:10.2196/jmir.2162

KEYWORDS

Internet; smoking cessation; recruitment; Hispanic/Latino; advertising

Introduction

Cigarette smoking is a major cause of disease and death in Latinos [1]. Approximately 12.5% of Latinos currently smoke [2] translating into roughly 6.3 million Latino smokers. With a projected growth of the Latino population to 133 million by 2050 (30.2% of the US population) [3,4] and the potential for smoking prevalence to increase as Latinos become more acculturated [5,6], there may be more than 16 million Latino smokers by 2050 if smoking rates remain unchecked. Tobacco cessation among Latinos is a public health priority [7,8].

Web-based cessation programs are a promising approach to reach Latino smokers with evidence-based cessation treatment. More than 60% of Latinos (32 million) are online—projected to increase to 70% by 2014 [9]—and almost half of online Latinos (45%) have used the Internet to search for health information [10]. A growing body of evidence supports the reach, efficacy, and cost-efficiency of Web-based cessation programs [11-19], but additional research is needed to understand their impact among subgroups of online smokers, including racial/ethnic minorities. In general, there are few studies of smoking cessation interventions developed specifically for Latinos [20,21], and only two studies of the effectiveness of Web-based interventions for Latinos [12,22]. One of the challenges in conducting Web-based cessation research with Latinos in the United States is that they have not engaged in Internet cessation programs as readily as other racial/ethnic subgroups [23] and recruitment of Latinos into Internet cessation trials has been challenging [11,12,22,24].

Online advertising is widely acknowledged as an increasingly important method to reach Latinos who use the Internet [25-28]. Also known as “display ads,” banner advertisements appear as graphical ads embedded into a webpage, typically including a combination of static/animated images, text, and/or video designed to convey a marketing message and/or cause the user to take an action [29]. Unlike “offline” ads (eg, billboards, newspapers, and flyers), banner ads can immediately link smokers to Web-based cessation programs, thus capitalizing on the motivation to quit when it occurs which can be critical to engaging consumers with cessation treatment [23,30]. Moreover, banner ads can target consumers by strategic placement on selected websites with synergistic content, known demographic profiles, or past online behavior. Therefore, online advertising may represent a “participant-friendly” and cost-efficient solution to reach and recruit online Latino smokers to Web-based cessation programs [31-34].

Reaching online Latino smokers requires not only efficient and effective advertising channels, but also an understanding of the target audiences’ preferred content and context of messages [32,33,35]. Approximately half of online Latinos use the Internet either primarily in Spanish (19%) or in both English and Spanish (28%), with the growth of both segments outpacing English language usage as more Spanish content becomes available online [36]. Therefore, we tested ads in Spanish to reach this large and growing segment of online Latinos. With regard to message content, Prospect Theory [37,38] suggests that people respond differently to factually equivalent messages depending on whether they are framed to emphasize benefits through gain-framed messages (eg, “if you quit smoking, you will live longer”) or costs through loss-framed messages (eg, “if you do not quit smoking, you will die sooner”). Applied to health behaviors, research suggests that when behaviors have a relatively certain outcome, gain-framed messages are more persuasive; if behaviors result in an uncertain outcome, loss-framed messages are more effective [39]. Because quitting smoking will almost certainly prevent disease, Prospect Theory predicts that gain-framed messages will be more persuasive than loss-framed messages. Recent evidence suggests that highlighting the benefits of quitting (through gain-framed messages) is more effective to encourage preventive behaviors such as smoking cessation [40-42]. However, there have been few studies that have examined the impact of message framing in health interventions designed specifically for racial/ethnic minorities [43,44].

With regard to message context, research with Latinos indicates that addressing cultural elements is critical in developing smoking cessation interventions [5,45-48]. The process of designing messages around group-level characteristics has been referred to as both cultural tailoring and cultural targeting [49-50]. We use the term “cultural targeting” in this study to refer to the delineation of a particular population segment. Cultural targeting can be done at a surface level (eg, attending to the visual characteristics or language of intervention materials) or at a deep level (eg, incorporating specific Latino cultural values) [51]. In general, few studies have investigated how surface- and deep-targeted messaging functions in specific subgroups of racial/ethnic minorities [52].

The purpose of this study was: (1) to investigate whether online advertising is an effective strategy to promote engagement with a Spanish-language cessation website among online Latino smokers, and (2) to identify the optimal message characteristics for this audience. To address these questions, we conducted a two-phased, mixed-methods study. In Phase I, formative research with online Latino smokers guided the development of four Spanish-language banner ad prototypes (deep-targeted/gain-framed, deep-targeted/loss-framed, surface-targeted/gain-framed, and surface-targeted/loss-framed).

In Phase II, we tested the effectiveness of these four ads in reaching and recruiting Latino smokers to a free Spanish-language smoking cessation website. The primary outcomes in Phase II were: (1) the absolute number of clicks on an ad, (2) the click-through rate to the smoking cessation website, (3) the number of ad responders who went on to register on the cessation website (“registrants”), and (4) the cost per registrant. We hypothesized there would be an interaction between message framing and cultural targeting, such that the deep-targeted/gain-framed ad would outperform all others on all metrics. Based on Prospect Theory, we also hypothesized that gain-framed messages would outperform loss-framed messages on all metrics holding constant cultural targeting. Finally, we hypothesized that deep-targeted ads would outperform surface-targeted ads on all metrics holding constant the message frame.

Methods

Phase I: Advertising Development and Pre-testing

A multicultural, multidisciplinary, and bilingual expert panel with expertise in behavioral science, marketing, communication technology, and advertising identified two dominant cultural values to consider for cultural targeting of messages to Latino smokers: *familismo* and *fatalismo*. *Familismo* is central to the Latino health experience; family and close friends are often consulted for general health issues before a medical professional, and family is often a strong motivator of health behavior. *Fatalismo* refers to the belief “*si Dios quiere*” (“what God wants”) which often guides Latino perceptions about their control over health and wellness. Our formative Phase I focused on developing banner ad prototypes for each cultural value that crossed message frame (gain vs loss) with level of cultural targeting (surface vs deep) resulting in a total of eight Spanish-language ads (see Figure 1). All ads were written in Spanish, developed in Adobe Flash with animation and motion by a professional creative developer, and included a call to action (ie, “click here for more information”).

To examine preferences for ad characteristics, 200 American Spanish-speaking smokers were recruited from two Spanish-language websites: MSN Latino [53] and MySpace Latino [54]. A study invitation banner ad transferred potential participants to an online screening form where eligible individuals (ie, current smokers, aged 18 or older, and Hispanic/Latino ethnicity of any race) provided online consent and completed a survey. Survey items included demographics, smoking history, Internet use, general receptivity to online ads, and the language subscale of the Short Acculturation Scale for Hispanics (SASH) [55]. Next, each participant sequentially viewed two of the eight ads. The two ads differed on cultural value (*familismo* vs *fatalismo*), but were identical in message frame and depth of cultural targeting (eg, gain-framed/deep-targeted) for direct comparison of message characteristics. Ad presentation was counterbalanced so that an equal number of participants saw each of the 8 ads first. Participants were asked to indicate their receptivity to the ads using a 10-point Likert scale (1 = not at all, 10 = very much). Participants also responded to 12 statements that addressed the strength of cultural targeting (eg, this ad is for someone like me), ad credibility (eg, this ad is believable), perceived efficacy of the ad (eg, this ad has me thinking about quitting smoking), and intention to respond (eg, I would click on this ad if I saw it) using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Finally, both ads were presented in tandem, and participants were asked which one they liked the most and why. An equal number of participants viewed each ad pair. The survey was administered in Spanish.

Figure 1. Gain-framed/deep-targeted ad for familismo.

Data Analysis

Descriptive statistics were used to examine the demographic, smoking history, Internet use, online advertising receptivity, acculturation measures, and ratings of individual ad characteristics. Analysis of variance (ANOVA) procedures, Chi-square tests (χ^2), and *t* tests were used to evaluate differences between the two ad types (*familismo* vs *fatalismo*).

Phase II: Evaluation of Online Banner Ad Effectiveness

Phase II examined the effectiveness of the four final ads (see Multimedia Appendix) in promoting the free Spanish-language version of BecomeAnEX (www.becomeanex.org) [56]. All ads were tested in Spanish (the English language version of each ad is shown in the Appendix for reference only). BecomeAnEX is a free, branded smoking cessation website designed to engage smokers through videos, interactive content, a personalized quit plan, and an online community of current and former smokers. Using a Latin square design (Table 1), the four ads (deep-targeted/gain-framed, deep-targeted/loss-framed, surface-targeted/gain-framed, and surface-targeted/loss-framed) were systematically rotated across four Spanish-language websites over four time periods of approximately one month each between August and November 2009. This design controlled for variability in ad performance associated with website placement and time.

The four websites chosen for this study were MySpace Latino [54], MSN Latino [53], MiGente [57], and Yahoo! en Español [58]. Our media partner, PHD Media LLC, examined various metrics of reach and performance [59] using data from comScore [60]—a widely used Internet analytics company—to identify sites with the best ability to reach our target audience of American adult Latinos. We selected a mix of portals with vast foreign-language segments and broad reach to Latino users as well as social network sites that specifically targeted Latinos. Banner ads ran on the home page and in various sections of each site (eg, email, weather, news, women's health, and men's health) to optimize their reach.

Banner ads were assigned a click tag that linked an individual's use of the BecomeAnEX website with the specific banner ad that prompted the visit. Each time an individual visited a website where one of the banner ads was placed, an impression (a unit of analysis analogous to a page view) was recorded by a tracking software program. If the user clicked on the banner ad, a click-through was also recorded. Click-through rate is calculated by dividing the total number of clicks by the total number of impressions, expressed as a percentage. This is a standardized metric that allows for comparison of ad performance across websites and varying costs by essentially controlling for the number of impressions [61]. All banner ads linked directly to the Spanish BecomeAnEX home page. When a user arrived at the BecomeAnEX website from a banner ad, a "home page visit" was recorded and the click tag associated with the banner ad was recorded along with standard website utilization metrics (eg, whether the individual registered to become a member, number of page views, and number of minutes online). Advertising media contracts with the four websites were designed to ensure a specific number of impressions of each banner ad over each of the four time blocks. Advertising costs by website for each time block are shown in Table 1. We also report the cost per click for each ad—calculated as the total cost divided by the number of clicks in a given ad campaign—to enable comparison to other published studies. Cost per registrant is calculated as the total cost divided by total number of registrants in a given ad campaign.

Table 1. Overview of Latin square design and overall banner ad performance by website placement and time.

Website	Metric	Month 1	Month 2	Month 3	Month 4
MySpace Latino	Ad type	Surface gain	Deep loss	Deep gain	Surface loss
	Cost (US$)	$5100.00	$5100.00	$5100.00	$5100.00
	Impressions	1,517,216	1,520,018	1,519,437	1,520,487
	Clicks	1146	988	1080	1278
	Click-through rate[a]	0.076%	0.065%	0.071%	0.084%
	Registrants	19	4	4	7
	Cost per click[b] (US$)	$4.45	$5.16	$4.72	$3.99
	Cost per registrant[c] (US$)	$268.42	$1275.00	$1275.00	$728.57
Yahoo	Ad type	Surface loss	Surface gain	Deep loss	Deep gain
	Cost (US$)	$8040.00	$8016.52	$7997.56	$7972.60
	Impressions	1,342,521	1,336,274	1,332,926	1,328,766
	Clicks	2737	2388	2052	1741
	Click-through rate	0.204%	0.179%	0.154%	0.131%
	Registrants	109	101	34	23
	Cost per click (US$)	$2.94	$3.36	$3.90	$4.58
	Cost per registrant (US$)	$73.76	$79.37	$235.22	$346.63
MSN Latino	Ad type	Deep gain	Surface loss	Surface gain	Deep loss
	Cost (US$)	$6827.00	$6791.74	$6852.86	$6771.21
	Impressions	1,498,058	1,490,065	1,509,246	1,489,548
	Clicks	1776	2289	2396	2185
	Click-through rate	0.119%	0.154%	0.159%	0.147%
	Registrants	37	71	74	11
	Cost per click (US$)	$3.84	$2.97	$2.86	$3.10
	Cost per registrant (US$)	$184.51	$95.66	$92.61	$615.56
Mi Gente	Ad type	Deep loss	Deep gain	Surface loss	Surface gain
	Cost (US$)	$6250.00	$6250.00	$6250.00	$6250.00
	Impressions	1,745,792	2,116,132	1,781,070	1,781,451
	Clicks	653	673	703	737
	Click-through rate	0.037%	0.032%	0.039%	0.041%
	Registrants	0	1	3	2
	Cost per click (US$)	$9.57	$9.29	$8.89	$8.48
	Cost per registrant (US$)	N/A[d]	$6250.00	$2083.33	$3125.00

[a] the number clicks divided by the number of impressions

[b] cost divided by the number of clicks in a given ad campaign (not an outcome metric in this study, but included in the table for comparison to other published studies)

[c] cost divided by total number of registrants in a given ad campaign

[d] cost per registrant cannot be calculated since there were 0 registrants during this segment of the study

Data Analysis

First, descriptive statistics were calculated for each of the four outcome variables (ie, absolute number of clicks, click-through rate, number of registrants, and cost per registrant) by website placement, time, and ad type. Next, a series of regression models explored the interaction of message framing and cultural targeting, and the main effects of website placement, time, message framing, and cultural targeting. Three of the outcome variables are count data (eg, clicks and registrants) or are based on count data (eg, click-through rate) and typically would be analyzed using Poisson regression. However, since the variance of these outcomes demonstrated over-dispersion, we determined that Poisson regression models would not be a good fit. Instead, we used negative binomial regression models to examine clicks, registrants, and click-through rate, and general linear regression models to examine cost per registrant. We evaluated the presence of an interaction between message framing and cultural targeting by including an interaction term in the model. For those models without evidence of an interaction, we evaluated separately the main effects of message framing and cultural targeting for loss-framed versus gain-framed ads and surface-targeted versus deep-targeted ads, respectively. All statistical analyses were performed using SAS Version 9.2 for Windows [62].

Results

Phase I: Prototype Development and Pre-testing

A total of 5,698,776 impressions to the Phase I recruitment banner ad were generated from MSN Latino and MySpace Latino between June 1-24, 2009. A total of 8231 individuals clicked on the study invitation ad. Of those who reached the survey site, approximately half (3712/7436, 49.92%) completed eligibility screening and 25.51% (947/3712) met the eligibility criteria. Of those eligible, 67.2% (636/947) provided informed consent, and of those who consented, 32.1% (204/636) completed the survey.

A total of 204 Latino current smokers completed the survey: 56.4% (115/204) were male, average age was 36.7 years (SD 11.7, range 18-70), 53.9% (110/204) were white, 3.4% (7/204) black, 4.9% (10/204) American Indian/Alaskan native, 1.5% (3/204) native Hawaiian/other Pacific Islander, 0.5% (1/204) Asian, and 35.8% (73/204) choose the “other” category. Participants used the “other” category to describe multiracial background and country of origin. The mean score on the SASH was 1.63 (SD 0.56, range 1.0-3.5) indicating a low level of acculturation. Respondents reported smoking an average of 10.1 cigarettes per day (SD 9.7). Over half (56.9%, 116/204) reported smoking every day (43%, 88/204 smoke some days). The average number of smoking days per month was 22.9 (SD 9.7). Approximately half (48.5%, 99/204) had their first cigarette > 60 minutes after waking suggesting a low level of nicotine dependence, and 84.3% (172/204) expressed a desire to quit smoking within the next 6 months. Respondents were largely Internet-savvy and receptive to online advertising: 72.5% (148/204) reported using the Internet several times a day and about half (49.5%, 101/204) reported using the Internet for more than 5 years. Two-thirds (66.2%, 135/204) of respondents viewed Web advertising favorably (ie, like it “somewhat” or “a lot”).

All eight ads were reviewed positively, with mean ratings of “liking” ranging from 7.24 (SD 2.8) to 8.17 (SD 2.6). For the 12 items related to ad characteristics, there were no differences between ratings of *familismo* versus *fatalismo* ads: all eight ads were rated positively across all 12 dimensions both with regard to mean score and proportion of positive responses. When presented with both ads simultaneously, most of the participants (55.9%, 114/204) preferred the *familismo* ad, and gave reasons that related directly to its cultural relevance and the value of *familismo* (eg, “Both talk about quitting smoking, but this one tells me to think about my family;” “This ad definitely has a more profound message and most of all talks about family—it is more profound and I identify with it;” or “Because it speaks of future consequences that could happen; both [ads] are important but I believe the family one leaves more of a mark”). Given these results, we selected the Latino cultural value of *familismo*—illustrated by a young boy playing soccer with his parents—to incorporate into the deep-targeted ads evaluated in Phase II.

Phase II: Evaluation of Online Banner Ad Effectiveness

Across the Phase II study period (August 10 to November 29, 2009), a total of 24,829,007 impressions were placed across the four websites for the four Spanish-language banner ads. There were 24,822 clicks on the banner ads, yielding an overall click-through rate of 0.10%. A total of 500 unique individuals registered on BecomeAnEX, yielding a conversion rate of 2.77%. The cost of the online advertising efforts totaled US $104,669.49, resulting in an overall cost per click of US $4.22 and an overall cost per registrant of US $209.34.

As shown in Table 2, website placement was a significant predictor of clicks (point estimate = 0.51, 95% CI 0.46 0.56, $P < .001$), click-through rate (point estimate = 0.45, 95% CI 0.41 0.48, $P < .001$), number of registrants (point estimate = 0.14, 95% CI 0.10 0.20, $P < .001$), and cost per registrant ($P = 0.01$). The MiGente website yielded the lowest number of clicks (n = 2766), click-through rate (0.037%), and number of registrants (n = 6), and highest cost per registrant (US $4166.67). Yahoo! and MSN Latino were the two best performing websites, with Yahoo! yielding a higher click-through rate (0.167% vs 0.144%; $\chi^2_1 = 9.8$, $P = .002$) and higher number of registrants (267 vs 193; $\chi^2_1 = 8.0$, $P = .005$).

Our main hypothesis that there would be a significant interaction between message framing and cultural targeting was not supported by any of the four outcomes (all P values > .05). Descriptive data for the message framing and cultural targeting interaction are shown in Table 3. Our second hypothesis that gain-framed ads would outperform loss-framed ads holding constant level of cultural targeting was not supported for absolute number of clicks, the number of registrants, or the cost per registrant. There was a statistically significant difference in click-through rate, with loss-framed ads yielding a higher click-through rate than gain-framed ads (0.105% vs 0.095%, point estimate = 1.08, 95% CI 1.03 1.14, P =.004). Our third hypothesis was that deep-targeted ads would outperform surface-targeted ads holding message frame constant; however, the results showed that for all outcomes, surface-targeted ads performed better than deep-targeted ads. Surface-targeted ads outperformed deep targeted ads on clicks (13,674 vs 11,148; point estimate = 1.20, 95% CI 1.13 1.28, *P* < .001), click-through rate (0.111% vs 0.089%; point estimate = 1.22, 95% CI 1.16 1.29, *P* < .001), registrants (386 vs 114; point estimate = 2.73, 95% CI 2.14 3.48, *P* < .001), and cost per registrant (US $135.75 vs US $458.49, *P* < .001).

Table 2. Banner advertising results by website placement.

Metric	Website				*P* value
	Yahoo! en Español	MSN Latino	MySpace Latino	MiGente	
Clicks	8918	8646	4492	2766	< .001
Click-through rate	0.167%	0.144%	0.074%	0.037%	< .001
Registrants	267	193	34	6	< .001
Cost per registrant (US$)	$119.95	$141.15	$600.00	$4166.67	.01

Table 3. Banner advertising results by message framing and cultural targeting.

Message Frame	Metric	Level of cultural targeting		
		Surface	Deep	Total
Loss	Clicks	7007	5878	12,885
	Click-through rate	0.114%	0.097%	0.105%
	Registrants	190	49	239
	Cost per registrant (US$)	$137.80	$533.04	$218.83
Gain	Clicks	6667	5270	11,937
	Click-through rate	0.109%	0.082%	0.095%
	Registrants	196	65	261
	Cost per registrant (US$)	$133.77	$402.30	$200.65
Totals	Clicks	13,674	11,148	24,822
	Click-through Rate	0.111%	0.089%	0.100%
	Registrants	386	114	500
	Cost per registrant (US$)	$135.75	$458.49	$209.34

Discussion

This is one of the first studies to examine the impact and costs of online Spanish-language banner advertising to reach and recruit Latino smokers to a Web-based cessation intervention. Our strategy was specifically designed to test the efficacy of various messaging elements and website placement to generate online smoker response and engagement. Overall, the results demonstrated that online advertising can effectively reach Latino smokers: during the 4-month study period, 24,822 individuals responded to banner advertising yielding a click-through rate of 0.10%; of these, 500 registered on the BecomeAnEX Spanish-language smoking cessation website.

The overall click-through rate in this study (0.10%) compares favorably to industry averages [63] and other research studies. Klausner et al [64] used online banner ads to promote a San Francisco Department of Public Health website for syphilis with an overall click-through rate in the 2-month study of 0.1% with different ads yielding click-through rates of 0.05% to 0.14%. Bull et al [65] evaluated the effectiveness of online banner ads for recruitment to human immunodeficiency virus (HIV) and sexually transmitted disease (STD) prevention research trials and reported a click-through rate of 0.05%. A recent study examining the effectiveness of online study recruitment via Facebook advertising yielded a click-through rate of 0.05% [66]. To some, these rates may raise questions about the generalizability of banner advertising. However, it is important to remember that of the tens of thousands of impressions placed on any given website, only a small percentage of those who view a banner ad may be members of the target audience or meet eligibility criteria. Given the current and growing magnitude of the Latino online audience, even reaching a small percentage can translate into relatively large numbers of smokers. This recruitment yield may, in fact, be comparable to or even higher than a newspaper or radio advertisement that yields several hundred responses when thousands of individuals may have been exposed to the ad. The primary difference is that the denominator can be determined with more certainty in online advertising compared to traditional mass media [30].

Results also indicate that online banner ads can be cost efficient, particularly when compared to other traditional program recruitment methods for racial/ethnic minority participants. While there was significant variation in cost per registrant across website placement and message effects, the most efficient and cost-effective advertising approach (surface-targeted ads placed on Yahoo! en Español) yielded 210 registrants in 2 months at a cost of US $73-$79 per registrant. These results compare favorably to other efforts to recruit minority participants using traditional approaches (ie, proactive face-to-face methods or reactive “offline” media-based methods) that reported lower and/or slower recruitment yield at higher cost [67-69]. The capacity of online advertising to reach and recruit 500 Latino smokers to a cessation website during a 4-month period has important implications for cessation-related eHealth research. Registered users of cessation websites report high levels of motivation to quit [13,24,70-74] and may represent the strongest pool available to efficiently recruit Latino smokers into clinical trials.

Contrary to the published literature and our hypotheses, results indicated that surface targeting was a more effective approach than deep targeting. However, these results should be considered within the context of several study parameters. The use of family imagery to reflect deep-level targeting and tobacco product imagery to reflect surface-level targeting was intentional because we were interested in examining two different types of imagery. It is possible that surface-targeted ads may have yielded better results because smokers are more likely to attend and respond to images featuring tobacco products given their more obvious relevance. Thus, we cannot conclude from this study that deep-level targeting is not effective. Future research should consider an additive approach in which deep cultural targeting graphics include culturally relevant elements as well as tobacco products/images to catch the attention of smokers.

Also contrary to our hypotheses, loss-framed messages yielded a higher click-through rate than gain-framed messages. This finding is intriguing, especially in light of the non-significant results for the other outcome metrics for message framing. On the one hand, it suggests that the behavior of clicking on an ad is indeed sensitive to message frame whereas the more distal behavior of registering on a cessation website is not. However, a number of variables may explain why this may be true. For example, previous advertising research has shown that consumers’ level of involvement or interest with a product or issue exerts a strong influence on the effect of message framing [75]. Consumers with high involvement pay more attention to the issue of losses, whereas consumers with low levels of involvement with a product or issue pay more attention to gains in messaging [76]. Thus, response variation across websites may reflect varying levels of engagement whereby visitors who are specifically looking to obtain general news information, check email, or connect in an online social network were more likely to attend to an ad with negatively framed content. Future research will be important to examine factors such as consumer involvement, ad placement, and message processing to better understand how to engage smokers at the point of an initial recruitment message.

Data from this study also suggest that social media websites may not be as fruitful for online recruitment as more general websites. Click-through rates from the MiGente and MySpace websites were significantly lower than MSN Latino and Yahoo! en Español. Bull and colleagues [77] also found lower-than-average click-through rates for MiGente (0.06%). These lower click-through rates may reflect differences in engagement patterns on social networking sites as compared to websites related to news and information, particularly for Latinos. Social networking sites, which seek to maximize interpersonal connections, may elicit high levels of engagement and possibly lower the probability of attending to an advertisement or of having a positive attitude toward online advertising on the site. An exception to this hypothesis is Facebook which has been an effective context for research recruitment [66,78] perhaps due to its particular advertising guidelines. Future research is needed to understand the opportunities and challenges of using online social networks to reach and recruit smokers to cessation interventions [79].

The results should be considered within the context of several limitations. First, campaign costs reflect a variety of factors and may be difficult to replicate. Specific information related to costs and advertising yield is provided to enable comparisons. Second, the extent to which our results are generalizable to other websites is unclear. The similarity in results obtained for Yahoo! en Español and MSN Latino lends support to the notion that other portals may perform similarly, but this remains to be tested. Relatedly, these findings are limited to Latinos using a Spanish-language website. The effectiveness of online ads in reaching English-preferring or bilingual Latinos should be addressed in future research. Third, it is possible that repeat visitors to any of the four websites where we placed banner ads could have seen more than one study-related ad, and this repeat exposure may have primed a subsequent click. Fourth, although not specifically a limitation, it is important to note that each of the four outcomes may be influenced by various factors, including but not limited to the variables examined in this study (ie, message framing and cultural targeting). For example, click-through rate is largely determined by the content of an ad, but it may also be a function of the degree to which users of certain websites attend to advertising; the number of registrants involves not only clicking on an ad but also then perusing a website and deciding to provide personal information to register. Lastly, the reach of this method is obviously limited to those Latino smokers who are online. Alternative strategies are needed to reach “offline” Latino smokers with cessation services.

Despite these limitations and considerations, this study has a number of strengths. Evaluating the use of online advertising to promote evidence-based online cessation treatment among Spanish-speaking Latino smokers is unique within the field of tobacco control. We assembled a multicultural, multidisciplinary, and bilingual team with expertise that spans behavioral science, marketing, communication technology, and advertising to examine how to optimize our ability to reach and recruit online Latino smokers to an evidence-based cessation program. The Latin square quasi-experimental design was extremely efficient in allowing us to control for two potentially confounding factors: the volume of website traffic and time.

We considered other research designs and approaches including laboratory experiments to examine cognitive processing factors such as message recall, message recognition, and psychological variables that mediate or moderate response to ads. However, message testing under highly controlled conditions at the expense of external validity would not yield the kind of information needed to successfully promote registration to a Web-based cessation website that could be used to inform recruitment of Latino smokers in subsequent cessation trials. We encourage the use of quasi-experimental methods as appropriate and hope that the results presented herein provide a useful benchmark in future eHealth research. Our approach may also help other public health researchers consider how to use the Internet for disease prevention.

In summary, identifying the effectiveness of various advertising parameters to promote an evidence-based online cessation program among Spanish-speaking Latino smokers holds potential for online program implementation and intervention research. Although a large number of Latino smokers use the Internet and are interested in using it for smoking cessation, the participation rate of Latinos in Web-based cessation interventions has been relatively low. Digital metrics such as click-through rate and registration rate related to various online message characteristics provide a deeper understanding of factors related to actual attention and appeal of online programs, particularly for specific subgroups [80]. Future research is needed to build on the methods used in this preliminary investigation to better understand how to increase consumer demand for tobacco cessation services, especially among racial/ethnic minorities [81].

Acknowledgments

Primary funding for this work was from the National Cancer Institute at the National Institutes of Health (R21 CA133319 to ALG). This work was also supported in part by the National Cancer Institute of the National Institutes of Health (U01 CA114593). The funding agency had no involvement in the conduct of the study or preparation of this manuscript. The authors would like to acknowledge the contributions of Jeffrey Costantino (Marketing Department, Legacy) in managing the creative development work done by GSD&M and the media planning and buying done by PHD Media LLC. We also thank Raymond Niaura, PhD, for his comments and assistance in conducting statistical analyses.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Final banner ads in English (N=4) and Spanish (N=4).

[PDF File (Adobe PDF File), 310KB - jmir_v14i4e116_app1.pdf]

References

1. Tobacco use among U.S. racial/ethnic minority groups--African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, Hispanics. A Report of the Surgeon General. Executive summary. MMWR Recomm Rep 1998 Oct 9;47(RR-18):v-xv, 1 [FREE Full text] [Medline: 9784089]
2. Centers for Disease Control and Prevention (CDC). Vital signs: current cigarette smoking among adults aged ≥18 years--United States, 2005-2010. MMWR Morb Mortal Wkly Rep 2011 Sep 9;60(35):1207-1212 [FREE Full text] [Medline: 21900875]
3. Ennis SR, Rios-Vargas M, Albert NG. The Hispanic Population: Census Briefs. 2010. 2011 URL: http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf [accessed 2012-05-01] [WebCite Cache ID 67LE3E6I4]
4. United States Census Bureau. US Department of Commerce. 2008 Aug 14. An older and more diverse nation by midcentury URL: http://www.census.gov/newsroom/releases/archives/population/cb08-123.html [accessed 2012-05-01] [WebCite Cache ID 67LEGLUuq]
5. Pérez-Stable EJ, Ramirez A, Villareal R, Talavera GA, Trapido E, Suarez L, et al. Cigarette smoking behavior among US Latino men and women from different countries of origin. Am J Public Health 2001 Sep;91(9):1424-1430. [Medline: 11527775]
6. Unger JB, Cruz TB, Rohrbach LA, Ribisl KM, Baezconde-Garbanati L, Chen X, et al. English language use as a risk factor for smoking initiation among Hispanic and Asian American adolescents: evidence for mediation by tobacco-related beliefs and social norms. Health Psychol 2000 Sep;19(5):403-410. [Medline: 11007148]
7. US Department of Health and Human Services. Tobacco Use Among US Racial/Ethnic Minority Groups - African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: A Report of the Surgeon General. Atlanta, Georgia: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1998. URL: http:/

/www.cdc.gov/tobacco/data_statistics/sgr/1998/complete_report/pdfs/complete_report.pdf [accessed 2012-05-02] [WebCite Cache ID 67MpENZSf]

8. Fiore MC, Jaén CR, Baker TB, et al. Treating Tobacco Use and Dependence: 2008 Update. Clinical Practice Guideline. Rockville, MD: US Department of Health and Human Services. Public Health Service; May 2008.
9. Interactive Advertising Bureau Hispanic Research Working Group. US Latinos Online: A driving force. 2011 May. URL: http://www.iab.net/media/file/USLatinosOnline_RevisedMay2011FINAL.pdf [accessed 2012-05-01] [WebCite Cache ID 67LNTwAGM]
10. Fox S, Jones S. The social life of health information: Americans' pursuit of health takes places within a widening network of both online and offline sources. Washington, DC: Pew Internet & American Life Project; 2009 Jun. URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2012-05-01] [WebCite Cache ID 67LNe4tne]
11. Graham AL, Cobb NK, Papandonatos GD, Moreno JL, Kang H, Tinkelman DG, et al. A randomized trial of Internet and telephone treatment for smoking cessation. Arch Intern Med 2011 Jan 10;171(1):46-53. [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]
12. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine Tob Res 2009 Sep;11(9):1025-1034. [doi: 10.1093/ntr/ntp090] [Medline: 19640833]
13. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216. [doi: 10.1080/14622200500055319] [Medline: 16036277]
14. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
15. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]
16. Barrera AZ, Pérez-Stable EJ, Delucchi KL, Muñoz RF. Global reach of an Internet smoking cessation intervention among Spanish- and English-speaking smokers from 157 countries. Int J Environ Res Public Health 2009 Mar;6(3):927-940. [doi: 10.3390/ijerph6030927] [Medline: 19440423]
17. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. J Med Internet Res 2008;10(5):e45 [FREE Full text] [doi: 10.2196/jmir.1008] [Medline: 19033146]
18. An LC, Schillo BA, Saul JE, Wendling AH, Klatt CM, Berg CJ, et al. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. J Med Internet Res 2008;10(5):e55 [FREE Full text] [doi: 10.2196/jmir.1018] [Medline: 19103587]
19. Leykin Y, Aguilera A, Torres LD, Pérez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an International Smoking Cessation Study. J Med Internet Res 2012;14(1):e5 [FREE Full text] [doi: 10.2196/jmir.1829] [Medline: 22314016]
20. Webb MS, Rodríguez-Esquivel D, Baker EA. Smoking cessation interventions among Hispanics in the United States: A systematic review and mini meta-analysis. Am J Health Promot 2010;25(2):109-118. [doi: 10.4278/ajhp.090123-LIT-25] [Medline: 21039291]
21. Lawrence D, Graber JE, Mills SL, Meissner HI, Warnecke R. Smoking cessation interventions in U.S. racial/ethnic minority populations: an assessment of the literature. Prev Med 2003 Feb;36(2):204-216. [Medline: 12590996]
22. 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]
23. McCausland KL, Curry LE, Mushro A, Carothers S, Xiao H, Vallone DM. Promoting a Web-Based Smoking Cessation Intervention: Implications for Practice. Cases in Public Health Communication & Marketing 2011;5(5 Proc):3-26.
24. Graham AL, Bock BC, Cobb NK, Niaura R, Abrams DB. Characteristics of smokers reached and recruited to an internet smoking cessation trial: a case of denominators. Nicotine Tob Res 2006 Dec;8 Suppl 1:S43-S48. [Medline: 17491170]
25. Cox J. Getting Connected.: Hispanic Business Magazine; 2004 Dec. URL: http://www.hispanicbusiness.com/news/newsbyid.asp?id=19387&page=1&cat=&more= [accessed 2012-05-01] [WebCite Cache ID 67LNkq3va]
26. La Ferle C. Lee W-N. Can English Language Media Connect with Ethnic Audiences? Ethnic Minorities' Media Use and Representation Perceptions. Journal of Advertising Research 2005 Mar;45(1):140-153. [doi: 10.1017/S002184990505004X]
27. Lopez M. iMedia Connection. 2005. Reaching Hispanics online URL: http://www.imediaconnection.com/content/6693.asp [accessed 2012-05-24] [WebCite Cache ID 67u9XUENr]
28. Interactive Advertising Bureau. 2004. Reaching Hispanics through online marketing URL: http://www.iab.net/media/file/resources_docs_Hispanic_Presentation_Final.pdf [accessed 2012-05-01] [WebCite Cache ID 67LO0TeqJ]
29. Interactive Advertising Bureau. 2012 Feb 26. IAB display advertising guidelines: the new 2012 portfolio URL: http://www.iab.net/guidelines/508676/508767/displayguidelines [accessed 2012-05-24] [WebCite Cache ID 67u14Z4SB]

30. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. J Med Internet Res 2008;10(5):e50 [FREE Full text] [doi: 10.2196/jmir.1001] [Medline: 19073542]
31. Swanson GM, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. J Natl Cancer Inst 1995 Dec 6;87(23):1747-1759. [Medline: 7473831]
32. Bolen S, Tilburt J, Baffi C, Gary TL, Powe N, Howerton M, et al. Defining "success" in recruitment of underrepresented populations to cancer clinical trials: moving toward a more consistent approach. Cancer 2006 Mar 15;106(6):1197-1204 [FREE Full text] [doi: 10.1002/cncr.21745] [Medline: 16453333]
33. Ford JG, Howerton MW, Bolen S, Gary TL, Lai GY, Tilburt J, et al. Knowledge and access to information on recruitment of underrepresented populations to cancer clinical trials. Evid Rep Technol Assess (Summ) 2005 Jun(122):1-11 [FREE Full text] [Medline: 15989377]
34. Lai GY, Gary TL, Tilburt J, Bolen S, Baffi C, Wilson RF, et al. Effectiveness of strategies to recruit underrepresented populations into cancer clinical trials. Clin Trials 2006;3(2):133-141. [Medline: 16773955]
35. Maibach E, Parrott RL. Designing health messages: approaches from communication theory and public health practice. Thousand Oaks, Calif: Sage Publications; 1995.
36. CapturaGroup. Hispanic Online Marketing. 2012 Mar 07. Are online Hispanics part of your 2012 plans? URL: http://www.hispaniconlinemarketing.com/2012/03/are-online-hispanics-part-of-your-2012-plans/ [accessed 2012-07-06] [WebCite Cache ID 68xanCYz9]
37. Kahneman D, Tversky A. Prospect Theory: An Analysis of Decision under Risk. Econometrica 1979 Mar;47(2):263-291.
38. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [Medline: 7455683]
39. Rothman AJ, Salovey P. Shaping perceptions to motivate healthy behavior: the role of message framing. Psychol Bull 1997 Jan;121(1):3-19. [Medline: 9000890]
40. Gallagher KM, Updegraff JA. Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. Ann Behav Med 2012 Feb;43(1):101-116. [doi: 10.1007/s12160-011-9308-7] [Medline: 21993844]
41. Toll BA, Martino S, Latimer A, Salovey P, O'Malley S, Carlin-Menter S, et al. Randomized trial: Quitline specialist training in gain-framed vs standard-care messages for smoking cessation. J Natl Cancer Inst 2010 Jan 20;102(2):96-106 [FREE Full text] [doi: 10.1093/jnci/djp468] [Medline: 20056957]
42. Salovey P. National Registry of Evidence-based Programs and Practices (NREPP): SAMHSA's National Registry of Evidence-based Programs and Practices - Message Framing. 2002 Dec. Legacy program summary URL: http://www.nrepp.samhsa.gov/ViewLegacy.aspx?id=82 [accessed 2012-05-24] [WebCite Cache ID 67uAjBidq]
43. Apanovitch AM, McCarthy D, Salovey P. Using message framing to motivate HIV testing among low-income, ethnic minority women. Health Psychol 2003 Jan;22(1):60-67. [Medline: 12558203]
44. Schneider TR, Salovey P, Apanovitch AM, Pizarro J, McCarthy D, Zullo J, et al. The effects of message framing and ethnic targeting on mammography use among low-income women. Health Psychol 2001 Jul;20(4):256-266. [Medline: 11515737]
45. Marin G, Marin BV. Research with Hispanic populations. Newbury Park: Sage Publications; 1991.
46. Marin BV, Perez-Stable EJ, Marin G, Sabogal F, Otero-Sabogal R. Attitudes and behaviors of Hispanic smokers: implications for cessation interventions. Health Educ Q 1990;17(3):287-297. [Medline: 2228631]
47. Huerta EE, Macario E. Communicating health risk to ethnic groups: reaching Hispanics as a case study. J Natl Cancer Inst Monogr 1999(25):23-26. [Medline: 10854453]
48. Marín G. Defining culturally appropriate community interventions: Hispanics as a case study. Journal of Community Psychology 1993;21(2):149-161. [doi: 10.1002/1520-6629(199304)21:2<149::AID-JCOP2290210207>3.0.CO;2-Y]
49. Pasick RJ. Response to Kreuter and Skinner: consensus regarding use of the terms targeting and tailoring. Health Educ Res 2001 Aug;16(4):503-508 [FREE Full text] [Medline: 11525396]
50. Kreuter MW, Skinner CS. Tailoring: what's in a name? Health Educ Res 2000 Feb;15(1):1-4 [FREE Full text] [Medline: 10788196]
51. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
52. Resnicow K, Braithwaite R, Dilorio C, Glanz K. Applying theory to culturally diverse and unique populations. In: Glanz K, Rimer BK, Lewis FM, editors. Health behavior and health education: theory, research, and practice, 3rd Edition. San Francisco: Jossey-Bass; 2002:485-509.
53. MSN Latino. URL: http://latino.msn.com/ [accessed 2012-07-24] [WebCite Cache ID 69OdZPZaw]
54. MySpace Latino. URL: http://latino.myspace.com/ [accessed 2012-07-24] [WebCite Cache ID 69OddPvNX]
55. Marín G, Sabogal F, Marín B, Otero-Sabogal F, Pérez-Stable E. Development of a short acculturation scale for Hispanics. Hispanic Journal of Behavioral Sciences 1987;9:183. [doi: 10.1177/07399863870092005]
56. BecomeAnEX en Español. URL: http://es.becomeanex.org/ [accessed 2012-07-24] [WebCite Cache ID 69Odvs44M]
57. MiGente. URL: http://www.migente.com/ [accessed 2012-07-24] [WebCite Cache ID 69OdgcO5f]
58. Yahoo! en Español. URL: http://espanol.yahoo.com/ [accessed 2012-08-19] [WebCite Cache ID 6A26xa9zW]

59. Hackett J. comScore. 2011. What's my comp index? URL: http://blog.comscore.com/2011/03/whats_my_comp_index.html [accessed 2012-07-16] [WebCite Cache ID 69ChsvbCH]
60. comScore. URL: http://www.comscore.com/ [accessed 2012-07-24] [WebCite Cache ID 69Odj94o7]
61. Novak TP, Hoffman DL. New metrics for new media: toward the development of web measurement standards. World Wide Web Journal 1997;2(1):213-246.
62. SAS Institute Inc. SAS/STAT Users Guide, Version 9.2. Cary, NC: SAS Institute, Inc; 2008.
63. Wasserman T. Mashable Business. 2011 Jan 24. Behind the race to build a better banner ad URL: http://mashable.com/2011/01/24/better-banner-ad/ [accessed 2012-07-06] [WebCite Cache ID 68xXbrnZp]
64. Klausner JD, Levine DK, Kent CK. Internet-based site-specific interventions for syphilis prevention among gay and bisexual men. AIDS Care 2004 Nov;16(8):964-970. [doi: 10.1080/09540120412331292471] [Medline: 15511728]
65. 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]
66. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
67. Harris KJ, Ahluwalia JS, Catley D, Okuyemi KS, Mayo MS, Resnicow K. Successful recruitment of minorities into clinical trials: The Kick It at Swope project. Nicotine Tob Res 2003 Aug;5(4):575-584. [Medline: 12959796]
68. Keyzer JF, Melnikow J, Kuppermann M, Birch S, Kuenneth C, Nuovo J, et al. Recruitment strategies for minority participation: challenges and cost lessons from the POWER interview. Ethn Dis 2005;15(3):395-406. [Medline: 16108298]
69. Graham AL, Lopez-Class M, Mueller NT, Mota G, Mandelblatt J. Efficiency and cost-effectiveness of recruitment methods for male Latino smokers. Health Educ Behav 2011 Jun;38(3):293-300. [doi: 10.1177/1090198110372879] [Medline: 21460176]
70. Cobb NK, Graham AL. Characterizing Internet searchers of smoking cessation information. J Med Internet Res 2006;8(3):e17 [FREE Full text] [doi: 10.2196/jmir.8.3.e17] [Medline: 17032633]
71. Graham AL, Cobb NK, Raymond L, Sill S, Young J. Effectiveness of an internet-based worksite smoking cessation intervention at 12 months. J Occup Environ Med 2007 Aug;49(8):821-828. [doi: 10.1097/JOM.0b013e3180d09e6f] [Medline: 17693778]
72. Etter JF. Internet-based smoking cessation programs. Int J Med Inform 2006 Jan;75(1):110-116. [doi: 10.1016/j.ijmedinf.2005.06.014] [Medline: 16125450]
73. Etter JF, Perneger TV. A comparison of cigarette smokers recruited through the Internet or by mail. Int J Epidemiol 2001 Jun;30(3):521-525 [FREE Full text] [Medline: 11416075]
74. Stoddard JL, Augustson EM. Smokers who use internet and smokers who don't: data from the Health Information and National Trends Survey (HINTS). Nicotine Tob Res 2006 Dec;8 Suppl 1:S77-S85. [Medline: 17491174]
75. Jung WS, Villegas J. The effects of message framing, involvement, and nicotine dependence on anti-smoking public service announcements. Health Mark Q 2011;28(3):219-231. [doi: 10.1080/07359683.2011.595641] [Medline: 21815740]
76. Yoo CY. Interplay of message framing, keyword insertion and levels of product involvement in click-through of keyword search ads. International Journal of Advertising 2011;30(3):399-424.
77. Bull SS, Vallejos D, Levine D, Ortiz C. Improving recruitment and retention for an online randomized controlled trial: experience from the Youthnet study. AIDS Care 2008 Sep;20(8):887-893. [doi: 10.1080/09540120701771697] [Medline: 18777217]
78. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
79. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop Participants. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13(4):e119 [FREE Full text] [doi: 10.2196/jmir.1911] [Medline: 22182518]
80. 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]
81. National Tobacco Cessation Collaborative. Innovations in Building Consumer Demand for Tobacco Cessation Products and Services. 2007 May. URL: http://www.rwjf.org/files/publications/other/InnovationsinBuildingConsumerDemand.pdf [accessed 2012-05-24] [WebCite Cache ID 67uBjNIDu]

Abbreviations

ANOVA: Analysis of variance
SASH: Short Acculturation Scale for Hispanics

Edited by G Eysenbach; submitted 25.05.12; peer-reviewed by D Ramo, S Bull; comments to author 13.06.12; revised version received 17.07.12; accepted 23.07.12; published 27.08.12.

Please cite as:
Graham AL, Fang Y, Moreno JL, Streiff SL, Villegas J, Muñoz RF, Tercyak KP, Mandelblatt JS, Vallone DM
Online Advertising to Reach and Recruit Latino Smokers to an Internet Cessation Program: Impact and Costs
J Med Internet Res 2012;14(4):e116
URL: http://www.jmir.org/2012/4/e116/
doi:10.2196/jmir.2162
PMID:22954502

Original Paper

Socioeconomic and Geographic Disparities in Health Information Seeking and Internet Use in Puerto Rico

Lila J Finney Rutten[1], MPH, PhD; Bradford W Hesse[2], PhD; Richard P Moser[3], PhD; Ana Patricia Ortiz Martinez[4], PhD; Julie Kornfeld[5], PhD; Robin C Vanderpool[6], DrPH; Margaret Byrne[7], PhD; Guillermo Tortolero Luna[4], MD, PhD

[1]Clinical Monitoring Research Program, SAIC-Frederick, Inc, National Cancer Institute at Frederick, Frederick, MD, United States
[2]Behavioral Research Program, Health Communications and Informatics Research Branch, National Cancer Institute, Bethesda, MD, United States
[3]Behavioral Research Program, Science of Research and Technology Branch, National Cancer Institute, Bethesda, MD, United States
[4]Comprehensive Cancer Center, Medical Sciences Campus, University of Puerto Rico, San Juan, Puerto Rico
[5]University of Miami, Miller School of Medicine, Department of Epidemiology and Public Health, Miami, FL, United States
[6]College of Public Health, University of Kentucky, Lexington, KY, United States
[7]University of Miami, Miami, FL, United States

Corresponding Author:
Lila J Finney Rutten, MPH, PhD
Clinical Monitoring Research Program
SAIC-Frederick, Inc
National Cancer Institute at Frederick
5705 Industry Lane
Frederick, MD, 21702
United States
Phone: 1 301 947 4912
Fax: 1 301 480 2198
Email: finneyl@mail.nih.gov

Abstract

Background: Geographically isolated Hispanic populations, such as those living in Puerto Rico, may face unique barriers to health information access. However, little is known about health information access and health information-seeking behaviors of this population.

Objective: To examine differences in health and cancer information seeking among survey respondents who ever used the Internet and those who did not, and to explore sociodemographic and geographic trends.

Methods: Data for our analyses were from a special implementation of the Health Information National Trends Survey conducted in Puerto Rico in 2009. We collected data through random digit dialing, computer-assisted telephone interviews (N = 639). The sample was drawn from the eight geographic regions of the Puerto Rico Department of Health. To account for complex survey design and perform weighted analyses to obtain population estimates, we analyzed the data using SUDAAN. Frequencies, cross-tabulation with chi-square, and logistic regression analyses were conducted. Geographic information system maps were developed to examine geographic distributions of Internet use and information seeking.

Results: Of 639 participants, 142 (weighted percentage 32.7%) indicated that they had ever gone online to access the Internet or World Wide Web; this proportion was substantially lower than that of US mainland Hispanics who reported using the Internet (49%). While 101 of 142 (weighted percentage 59.6%) respondents who used the Web had ever sought health information, only 118 of 497 (weighted percentage 20.0%) of those who did not use the Web had sought health information. The pattern was similar for cancer information: 76 of 142 respondents (weighted percentage 47.2%) who used the Web had ever sought cancer information compared with 105 of 497 (weighted percentage 18.8%) of those who had not used the Web. These results were slightly lower but generally consistent with US mainland Hispanics' health (50.9%) and cancer (26.4%) information seeking. Results of separate logistic regression models controlling for sociodemographic characteristics demonstrated that, compared with individuals who did not seek health or cancer information, those who did were over 5 times as likely to have used the Internet (odds ratio 5.11, $P < .001$). Those who sought cancer information were over twice as likely to have used the Internet (odds ratio 2.5, $P < .05$). The frequency of Internet use and health and cancer information seeking was higher in the San Juan metro region than in more rural areas.

Conclusions: Our results contribute to the evidence base for health and cancer communication planning for Puerto Rico, and suggest that health education and outreach efforts should explore the use of available and trusted methods of dissemination such as radio and television, as well as community-based health care providers and organizations, to supplement and encourage use of the Internet as a source of health information.

(J Med Internet Res 2012;14(4):e104) doi:10.2196/jmir.2007

KEYWORDS

Health information seeking; cancer information seeking; Internet use; disparities; special populations; geographic trends

Introduction

The Internet has become a valuable tool in supporting consumers' health information and health care needs [1-4] and is used for health information seeking by millions of people worldwide [3]. As access to the Internet has increased, and as the availability of health information has expanded, online health information seeking has increased [3]. National Healthy People 2020 objectives for health communication and health information technology in the United States are directed toward the overarching goal to "use health communication strategies and health information technology to improve population health outcomes and health care quality, and to achieve health equity" [5]. The Healthy People 2020 objectives for health communication and health information technology draw on an expanding body of empirical evidence that documents an association between greater access to health information through the Internet and improved health knowledge, attitudes, and behavior [5-9].

Despite the building of momentum and support for improving access to health information, it remains to be seen whether the benefits of online access will be equitably enjoyed by consumers, particularly among traditionally underserved and disadvantaged populations in the United States [3]. Previous studies with national survey data have documented profound subpopulation disparities in access to health information, wherein persons with lower socioeconomic status or lower levels of education appear to face greater barriers to obtaining health information than those with higher socioeconomic status and education [10-14].

Disparities in information access in the United States may be particularly significant for Hispanic populations, who often face additional barriers in terms of language, culture, and media use [10-12,15,16]. Despite the proliferation of the Internet, mobile technology, and wireless devices, research demonstrates the persistence of a digital divide between Hispanic populations in the mainland United States and other groups, wherein less than one-third of Spanish-speaking Hispanics use the Internet, compared with 71% in the non-Hispanic population [13]. Geographically isolated Hispanic populations, such as those living in Puerto Rico, may face additional or unique barriers to health information access [17]. However, little is known about health information access and health information-seeking behaviors of this population.

Health information seeking has been shown to increase individuals' involvement in health-related decision making and improve their satisfaction with their health choices [10]. Information has been demonstrated to improve individuals' coping abilities, reduce anxiety, alleviate mood disturbances, and improve communication with family members [10]. In Puerto Rico, cancer is the second-leading cause of death, and significant disparities in cancer incidence, mortality, and survival for several cancer types have been documented in this US territory [18-24], some of which may be partly influenced by health information access. For example, breast cancer mortality rates for certain cancers vary by region and by socioeconomic status in Puerto Rico. This disparity may be due to both lack of information and lack of access to information about state-of-the-art treatment options.

Puerto Rico is an island of approximately 3.7 million people with approximately one-third of the population concentrated in the San Juan metropolitan area and its adjacent municipalities [25]. Over 44% of Puerto Ricans live in poverty; this is nearly double the proportion of Puerto Rican-origin Hispanics living in poverty in the United States (24%) [25].

Prior research has documented socioeconomic disparities across the municipalities, including disparities in socioeconomic indicators such as the proportion of residents living below the poverty level and the proportion of residents with lower levels of education [24]. Poverty rates range from approximately 35% in affluent areas near San Juan to 77.8% in rural areas in the geographical center of the island [26]. Specifically, greater socioeconomic deprivation has been observed in the centrally located municipalities of the island relative to the municipalities in the capital city, San Juan, metropolitan area [24].Such socioeconomic disparities have been shown to coincide with cancer morbidity and mortality [24,27], as well as access to and use of health information [10-14]. Previous research in Puerto Rico has documented socioeconomic disparities in incidence and mortality of selected cancers by an area-based socioeconomic position index [24].

Identifying and addressing barriers to information access and resultant knowledge gaps is critical to prevent and rectify the existing excess burden of disease shouldered by vulnerable populations [10,11]. Although recent data showed that the Internet is the most frequently reported source of information among adults in Puerto Rico [17], factors associated with Internet use, and lack of use, have not been well documented in this population. Thus, our investigation aimed to assess use of the Internet, health information seeking, and cancer information seeking, and to explore differences in these behaviors between respondents who use the Internet and those who do not in a representative sample of adults in Puerto Rico. Our analyses also explored geographic variability in Internet use and health information seeking.

Methods

The US National Cancer Institute developed the Health Information National Trends Survey (HINTS) as a surveillance mechanism for tracking population trends in cancer-relevant behavior, knowledge, and attitudes in the rapidly evolving health communication and informatics environment to inform effective health communication strategies across populations. In 2009, the University of Puerto Rico's Comprehensive Cancer Center and the Department of Health in Puerto Rico–Puerto Rico Behavioral Risk Factor Surveillance System (BRFSS), in a collaborative effort supported by the National Cancer Institute, implemented a fully translated Spanish version of HINTS using existing BRFSS infrastructure in Puerto Rico [17]. The HINTS 2007 US mainland Spanish-language instrument, which contained response options and directions in English, was carefully reviewed, fully translated, and edited by the HINTS Puerto Rico team to ensure appropriateness for the island population. Greater details about the translation process have been reported elsewhere [17].

Data for this study were collected from April 27 through June 28, 2009 through random digit dial, computer-assisted telephone interview (N = 639). We used a stratified sampling frame to represent the eight geographic regions of the Puerto Rico Department of Health. Interviews were conducted primarily in Spanish by experienced bilingual Puerto Rican interviewers. The unweighted response rate for the screener and extended interview was 76.3% (837/1097).

Measures

Sociodemographic Characteristics

Sociodemographic characteristics included the following: age (18–34, 35–49, 50–64, or 65+ years); gender; education (less than high school, high school graduate, some college or technical school, or college graduate and beyond); employment status (employed or not employed); annual household income (<US $15,000, $15,000–24,999, $25,000–34,999, or $35,000+); and marital status (married or living as married, or not married).

Access to Health Care and Health Status

To assess access to health care, respondents were asked if they had a regular source of health care and whether they had any kind of health insurance coverage. Respondents were also asked to rate their own health status on a scale ranging from poor to excellent.

Use of the Internet

Use of the Internet was assessed with the question "Do you ever go online to access the Internet or World Wide Web, or to send and receive email?" Responses were coded as yes or no.

Information Seeking

Respondents were asked the following questions about general health and cancer-specific information seeking: "Have you ever looked for information about health or medical topics from any source?" and "Have you ever looked for cancer information from any source?"

Geographic Location

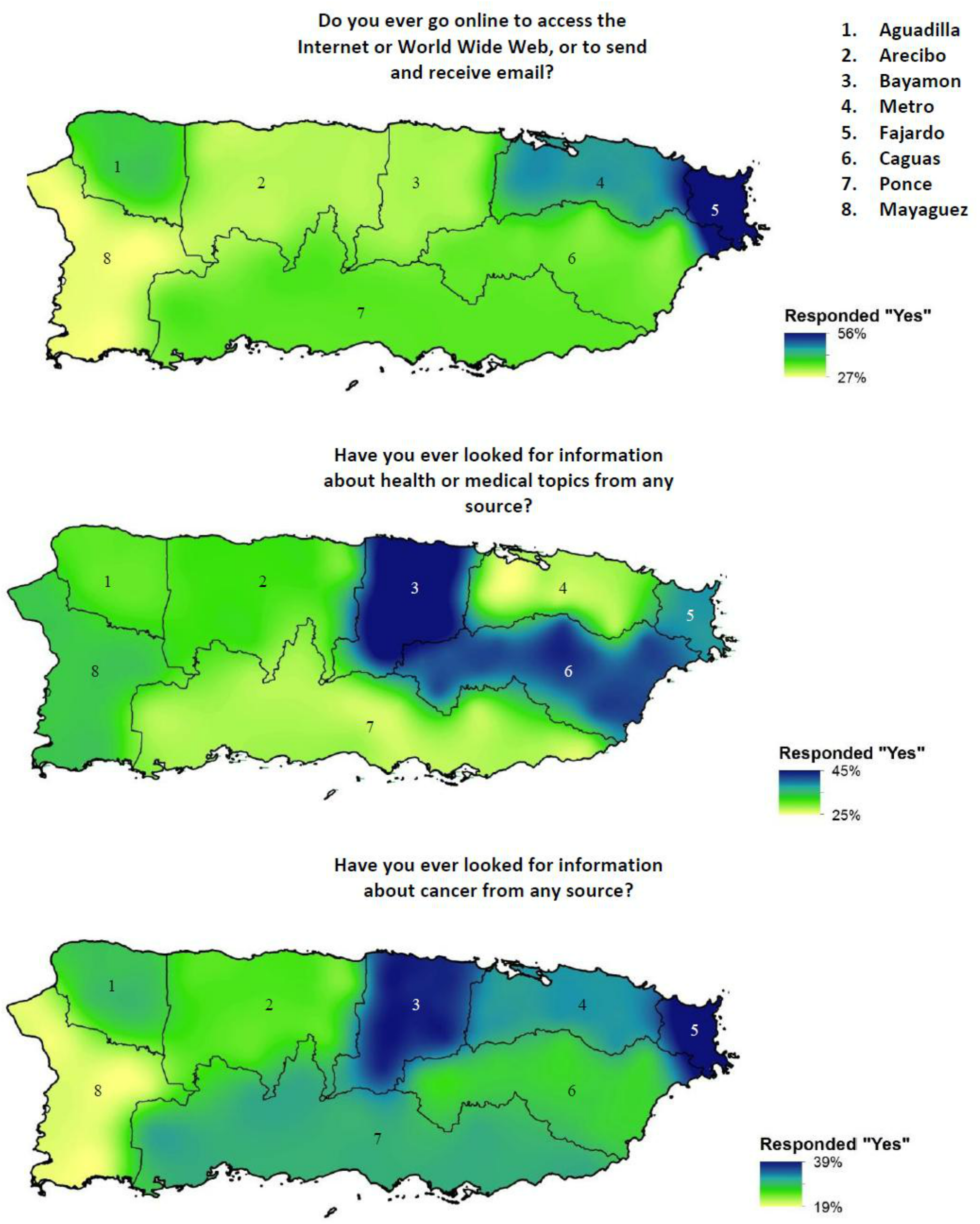
Geographic information was obtained through the sampling process and each respondent was placed into one of eight geographic regions in Puerto Rico: (1) Aguadilla, (2) Arecibo, (3) Bayamon, (4) Metro, (5) Fajardo, (6) Caguas, (7) Ponce, and (8) Mayaguez (see Figure 1).

Data Analysis

To account for the complex sampling design and to calculate accurate standard errors, we used SUDAAN version 9.0.1 [28] for our analyses. All data were weighted according to key sociodemographic estimates from the US Current Population Survey to be representative of the population in Puerto Rico [29]. Frequencies were calculated for all measures. Cross-tabulation tables with chi-square tests of significance were conducted to identify significant bivariate associations of Internet use with information-seeking behaviors, sociodemographic characteristics, and health care access. Two multivariable logistic regression analyses were conducted including variables that were significantly associated with Internet use in the bivariate analyses to test for significant independent associations of health and cancer information seeking with Internet use while controlling for sociodemographic variables. Income was not included in the multivariable models because only 61.0% (390 of the total sample of 639) reported their income; inclusion of the income variable would have substantially reduced the sample size for the models.

We generated geographic information system maps to examine the geographic distribution of use of the Internet, cancer information seeking, and health information seeking on the island using a spline method of interpolation [30]. This interpolation method estimates values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points. We selected the spline method of interpolation to ensure stability and precision of calculations with the large sample size.

Figure 1. Geographic distribution of Internet use and health and cancer information seeking.



Results

Sociodemographic Characteristics

Table 1 summarizes weighted percentages and sample frequencies for sociodemographic characteristics and health care access for the total sample and by self-reported use of the Internet. Estimates summarized herein are weighed to be representative of the adult population in Puerto Rico; thus, direct calculation of percentages from the sample frequencies given does not equal the values given. Of 639 respondents, 450 (weighted percentage 53.7%) were female and 443 (weighted percentage 40.2%) were aged 50 years or over. Of the 600 respondents who reported their education, 430 (weighted percentage 76.0%) had at least a high school education. Only 390 of the total sample of 639 respondents reported their income level; of those 390 respondents, 117 (weighted percentage 35.9%) reported annual incomes of $25,000 or greater. Of 598 respondents reporting on their employment, 165 (weighted percentage 40.8%) were employed. Of 603 respondents, 310 (weighted percentage 48.5%) were married.

Access to Health Care and Health Status

Of 636 respondents, 604 (weighted percentage 92.1%) reported that they had health care coverage (either private or government-based coverage), and 507 of 637 respondents (weighted percentage 74.1%) reported having a regular source of health care. Of 604 respondents, 166 (weighted percentage 34.8%) rated their health as excellent or very good.

Use of the Internet

Only 142 of 639 respondents (weighted percentage 32.7%) indicated that they ever went online to access the Internet. We observed several significant differences in the demographic profile of Internet users versus nonusers (Table 1). Generally, Internet users tended to be younger, to be more highly educated, to be employed, and to have higher annual incomes than non-Internet users. While fewer Internet users than nonusers reported having a regular health care provider, differences in health insurance status or perceived health were not statistically significant.

Information Seeking

Table 1 also summarizes weighted frequencies for health and cancer information seeking for the total sample and by Internet use. Of 639 respondents, 219 (weighted percentage 32.9%) reported that they had ever looked for information about health or medical topics and 101 (weighted percentage 28.1%) reported that they had ever looked for information about cancer. Internet users engaged in health information seeking and cancer information seeking with over twice the frequency of non-Internet users. Among 142 Internet users, 101 (weighted percentage 59.6%) reported that they had ever looked for information on health or medical topics compared with 118 of 497 nonusers (weighted percentage 20.0%). The pattern was similar for cancer information seeking, wherein 76 of 142 Internet users (weighted percentage 47.2%) reported that they had ever looked for cancer information compared with 105 of 497 non-Internet users (weighted percentage 18.8%).

Table 2 summarizes results of two multivariable logistic regression models to test for significant associations of health and cancer information seeking with Internet access controlling for other variables that were significantly associated with Internet use in bivariate analyses (excluding income). Model 1 examined the association of having ever sought health or medical information with use of the Internet, controlling for age, education, employment, and regular source of health care. Model 2 examined the association of having ever sought cancer information with use of the Internet, controlling for age, education, employment, and regular source of health care. In both models, the same factors emerged as significant: information seeking, age, and education were independently and significantly associated with Internet use. The odds of using the Internet were over 5 times as likely (odds ratio 5.11, $P < .001$) among respondents who sought health information and 2.5 times as likely (odds ratio 2.5, $P < .05$) among respondents who sought cancer information. Age and education level were also independently and significantly associated with Internet use in both models, with Internet users being younger and more educated. The odds of using the Internet significantly decreased with each increase in specified age group and significantly increased with each increase in specified education level. These trends were confirmed in an analysis of the independent variable effects for both age and education, such that each of the age and education categories differed significantly from their respective referent categories in a linear fashion.

Table 1. Weighted estimates and sample frequencies for sociodemographic characteristics, health care access, and health information-seeking patterns of respondents who reported using the Internet and those who did not.

Characteristic	Total (n = 639)	"Do you ever go online to access the Internet or World Wide Web…?"		P value[a]
		Yes (32.7%) (n = 142)	No (67.3%) (n = 497)	
	%	%	%	
Sex	n = 639	n = 142	n = 497	
Female	53.7	50.3	55.3	.44
Male	46.3	49.7	44.7	
Age range (years)	n = 639	n = 142	n = 497	
18–34	31.9	56.7	19.9	<.0001
35–49	27.9	31.6	26.2	
50–64	21.4	8.9	27.6	
65+	18.7	2.9	26.4	
Educational level	n = 600	n = 136	n = 464	
Less than high school	24.0	1.1	35.4	<.0001
High school graduate	27.8	20.0	31.6	
Some college	27.9	42.2	20.9	
College graduate	20.3	36.7	12.2	
Employment status	n = 598	n = 135	n = 463	
Employed	40.8	52.1	35.2	.01
Not employed	59.2	47.9	64.8	
Income ($US)	n = 390	n = 94	n = 296	
<15,000	36.9	18.5	46.3	<.0001
15,000–24,999	27.2	23.2	29.3	
25,000–34,000	14.4	16.3	13.4	
35,000+	21.5	42.0	11.1	
Marital status	n = 603	n = 136	n = 467	
Married/living as married	48.5	39.5	52.9	.07
Not married	51.5	60.5	47.1	
Regular health care provider	n = 637	n = 142	n = 495	
Yes	74.1	66.0	78.0	.047
No	25.9	34.0	22.0	
Health insurance coverage	n = 636	n = 142	n = 494	
Yes	92.1	87.7	94.2	.104
No	8.9	12.3	5.8	
Health status	n = 604	n = 136	n = 468	
Excellent or very good	34.8	46.6	29.0	.054
Good	31.0	27.0	33.0	
Fair or poor	34.2	26.5	38.0	
Looked for health information	n = 639	n = 142	n = 497	
Yes	32.9	59.6	20.0	<.001
No	67.1	40.4	80.1	
Looked for cancer information	n = 639	n = 142	n = 497	

Characteristic	Total (n = 639)	"Do you ever go online to access the Internet or World Wide Web…?"		*P* value[a]
		Yes (32.7%) (n = 142)	No (67.3%) (n = 497)	
	%	%	%	
Yes	28.1	47.2	18.8	.0001
No	71.9	52.8	81.2	

[a] *P* values associated with chi-square tests of independence for sociodemographic, health care access, and information-seeking variables by Internet use.

Figure 2. Geographic distribution of poverty and education in Puerto Rico. Data set: 2005-2009 Puerto Rico Community Survey 5-year estimates (public-use microdata).

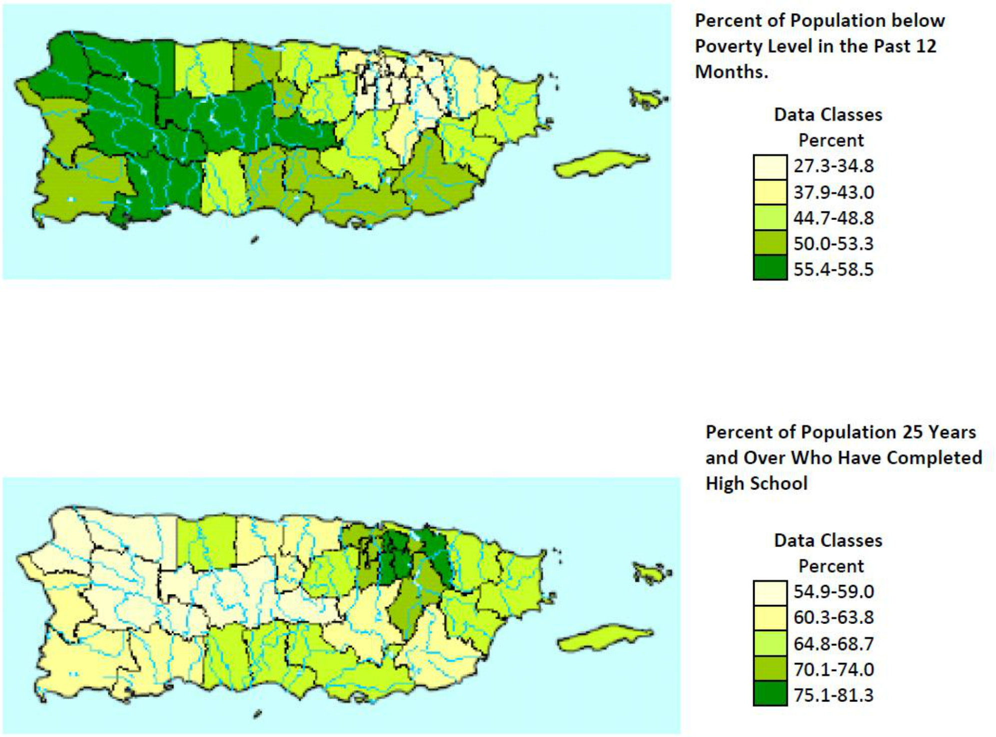


Table 2. Multivariable logistic models of independent associations with Internet use.

Model		Odds ratio	95% CI[a]
1	**(n = 595), overall model *F* 9 = 12.3, *P* < .0001**		
	Looked for health information		
	No	1.00	1.00
	Yes	5.11	2.18–11.98
	Age range (years)		
	18–34	1.00	1.00
	35–49	0.31	0.13–0.70
	50–64	0.07	0.02–0.18
	65+	0.05	0.01–0.16
	Educational level		
	High school or less	1.00	1.00
	Some college	4.55	1.18–17.52
	College graduate or more	7.79	2.66–22.83
	Employment status		
	Employed	1.00	1.00
	Not employed	0.93	0.44–1.96
	Regular health care provider		
	No	1.00	1.00
	Yes	0.76	0.30–1.91
2	**(n = 595), overall model *F* 9 = 15.4, *P* < .001**		
	Looked for cancer information		
	No	1.00	1.00
	Yes	2.51	1.02–6.16
	Age range (years)		
	18–34	1.00	1.00
	35–49	0.30	0.13–0.70
	50–64	0.80	0.03–0.22
	65+	0.06	0.02–0.19
	Educational level		
	High school or less	1.00	1.00
	Some college	4.55	1.22–16.92
	College graduate or more	9.72	3.35–28.15
	Employment status		
	Employed	1.00	1.00
	Not employed	0.94	0.46–1.91
	Regular health care provider		
	No	1.00	1.00
	Yes	0.74	0.30–1.80

[a] 95% Confidence intervals (CIs) not containing 1 were considered statistically significant.

Geographic Location

Figure 1 shows the geographic distributions of Internet use, and health and cancer information seeking. Geographic variability in Internet use and information seeking was observed, with relatively more frequent Internet use in the Fajardo and Metro regions than in other regions. The frequency of Internet use was lowest in the Mayaguez region. Health information seeking occurred most frequently in the Bayamon, Caguas, and Fajardo regions. The frequency of health information seeking was lowest in the Ponce and Metro regions. Cancer information seeking was most frequent in the Bayamon, Fajardo, and Metro regions, and least frequent in Mayaguez.

To place the observed geographic trends in Internet penetration into a context informed by related sociodemographic variables, we explored known geographic trends in key indicators of socioeconomic status: poverty and education. Figure 2 summarizes data from the American Community Survey (2005–2009) for poverty and education rates in Puerto Rico [29]. The highest levels of poverty and lowest levels of education were observed in Aguadilla, Arecibo, Ponce, and Mayaguez.

Discussion

Only one-third of the population in Puerto Rico reported going online to access the Internet or World Wide Web in 2009. This finding is consistent with 2011 Internet penetration rates for Puerto Rico (37.3%) obtained through market research [31] and prior HINTS analysis documenting low use of the Internet among US Spanish-speaking Hispanics [1]. This percentage is considerably lower than the comparable estimate of 68.4% from HINTS 2008 data for the general population of the US mainland and 49% for Hispanics living in the US mainland [17,24] and well below the Healthy People 2020 goal of 75% [5]. Consistent with previous research, Internet users in Puerto Rico tended to be younger and more highly educated, to be employed, and to have higher annual incomes than non-Internet users [14,31].

Prior research analyzing HINTS data from Puerto Rico revealed that the population seems to be split in its degree of trust in the Internet as a source of health information [17]. About a quarter of the population in Puerto Rico indicated “a lot” of trust in the Internet as a source of health information, while a similar proportion indicated that they did not trust the Internet “at all” as a source of health information and the remaining half indicated “somewhat” or “a little” trust in the Internet. This distribution of responses to trust in the Internet is distinct from the distribution of responses to trust in other sources of mediated health information such as television and radio. While the percentage of the population indicating a lot of trust in the Internet is similar to that for television and radio (and several other sources), population trust in the Internet seems much more divided, with a much greater proportion of the population indicating not trusting the Internet at all than the percentage who indicated not trusting other sources of information. This finding speaks to the need to employ a range of outreach and educational mechanisms that emphasize use of traditional media and interpersonal sources of health information.

Health and cancer information seeking varied significantly between those islanders who accessed the Internet and those who did not. Persons who had sought health information and those who had sought cancer information had 5 times the odds and 2.5 times the odds, respectively, of using the Internet. These findings point to a potential digital divide and are consistent with previous research indicating that Hispanic populations are less likely to be health information seekers [1,4,17]. Although prior research has demonstrated that the most frequently identified source of health information in Puerto Rico is the Internet [17], overall Internet penetration on the island is fairly low, and a significant segment of the population does not trust the Internet as a source of health information [17], thereby limiting access to the growing body of online health and cancer information.

Patterns of information-seeking behavior by sociodemographic characteristics observed in this study are also consistent with those of prior research showing associations of cancer and health information seeking with educational attainment and age [4,17,32]. Overall, approximately one-third of the population in Puerto Rico had ever looked for information about health or cancer topics from any source; these findings are similar to those observed in the HINTS 2008 continental US Hispanic population (health, 50.9%; cancer, 26.4%) [17,33]. Moreover, cancer and health information-seeking behavior varied significantly between those who accessed the Internet and those who did not.

Geographic variability in Internet use and health and cancer information seeking was considerable on the island. Internet use was more frequent in the San Jaun metropolitan area (Fajardo and Metro regions) than on the rest of the island, where frequency of Internet use was substantially lower. Health information seeking occurred most frequently in the regions surrounding the San Juan metro area (Bayamon, Caguas, and Fajardo); curiously, health information seeking was somewhat lower in the Metro region proper. Cancer information seeking was most frequent in San Juan metro area (Bayamon, Fajardo, and Metro regions) relative to the more rural areas of the island. When comparing our results with demographic data from the Current Population Survey, we observed a pattern of a higher prevalence of Internet use in regions with a greater proportion of persons with higher educational attainment and a lower proportion of persons living below the poverty level. This pattern is consistent with the greater likelihood of Internet use observed in our models among persons with higher education.

The geographic patterns of Internet use and information seeking are consistent with prior research in Puerto Rico demonstrating socioeconomic disparities in incidence and mortality of selected cancers by an area-based socioeconomic position index [24]. Specifically, socioeconomic deprivation is greater in the centrally located municipalities of the island than in the municipalities in the San Juan metropolitan area [24]. It has been proposed that populations residing in the central mountainous regions of Puerto Rico face greater obstacles to health care, including economic, environmental, cultural, and social barriers [24]. Such barriers are likely shared in common with barriers to the Internet and health information.

Recognition of the geographic pattern of Internet access provides an excellent resource for cancer information communication planning on the island and can guide efforts to expand existing infrastructure to rural areas or promote adoption of mobile technologies to support health information seeking. Furthermore, these data can inform efforts to tailor information sources, including print, traditional, and new media, to varying geographic regions. The data also support the continued use of traditional and trusted sources of health information such as interpersonal communication between patient and doctor, and utilization of trusted and trained community health workers. These interpersonal communication strategies must be leveraged to ensure that those who are offline are connected to health information. They can also be used to point consumers to reliable and credible sources of online health information.

The following limitations are worth noting. The survey design for HINTS Puerto Rico is cross-sectional; therefore, definitive conclusions about causal associations are not appropriate. In this respect, it is important to note that the direction and nature of the association between Internet access and health information seeking cannot be determined from these data. Limitations are also inherent in the use of random digit dial telephone methods and self-report measures. Additionally, this study was a one-time snapshot of Internet usage during a time of rapid change in cell phone usage and advances in mobile technology. Over a third of the sample did not report their income; therefore, we excluded the income variable from the multivariable model, preventing us from examining the independent association of income with Internet use and information seeking. With current trends in penetration of cell phones with Internet access, the level of Internet usage has increased in Puerto Rico since the time of the study. Survey data collected in 2011 indicate that 44% of respondents who have a cell phone use it to surf the Web [34]. This number is double that of 2009, pointing to a rapid increase in use of mobile devices for Internet access [34]. Despite these limitations, the data were derived from a stratified representative sample from the eight geographic regions of the Puerto Rico Department of Health, providing an adequate sample size to enable “state”-level analyses. Furthermore, the response rates for this survey were quite high, lending further confidence to the generalizability of the findings.

Understanding health information-seeking behavior in relationship to use of the Internet is timely and important, given the rapid increase in the amount of information available online and the increasing influence of online health information seeking on health behaviors, health processes, and health outcomes [1,4,7,32,35]. Much of the content available online is published in English. Increasing Spanish-language content available online would remove a key barrier to access and likely increase use among Spanish-speaking populations. Allocation of resources to enable access to online information has considerable potential to reach a broad audience of consumers. Efforts to increase access to the Internet through expanding the urban infrastructure to rural regions and through use of mobile devices and applications are encouraged. In Puerto Rico, observation of distinct regions with low Internet penetration and information seeking (eg, Mayaguez) could help to inform both infrastructure (eg, broadband penetration in rural areas) and educational efforts to ensure that health information is made more available via traditional modalities. In addition to long-term planning around communication and health care infrastructure, health education and outreach efforts should consider using available and trusted [17] methods of dissemination such as radio and television, as well as community-based health care providers and organizations. Our analyses provide important insights into Internet use and health information-seeking behaviors and experiences of the population in Puerto Rico and contribute to the evidence base for health and cancer communication planning for the island.

Acknowledgments

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

Conflicts of Interest

None declared.

References

1. Viswanath K. Science and society: the communications revolution and cancer control. Nat Rev Cancer 2005 Oct;5(10):828-835. [doi: 10.1038/nrc1718] [Medline: 16195753]
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. Rutten LJ, Squiers L, Hesse B. Cancer-related information seeking: hints from the 2003 Health Information National Trends Survey (HINTS). J Health Commun 2006;11 Suppl 1:147-156. [doi: 10.1080/10810730600637574] [Medline: 16641080]

5. Healthy People 2020. US Department of Health & Human Services. 2012 Jul 3 URL: http://healthypeople.gov/2020/ [accessed 2012-07-13] [WebCite Cache ID 6983WUAyT]
6. Shim M, Kelly B, Hornik R. Cancer information scanning and seeking behavior is associated with knowledge, lifestyle choices, and screening. J Health Commun 2006;11 Suppl 1:157-172. [doi: 10.1080/10810730600637475] [Medline: 16641081]
7. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]
8. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
9. Czaja R, Manfredi C, Price J. The determinants and consequences of information seeking among cancer patients. J Health Commun 2003;8(6):529-562. [doi: 10.1080/716100418] [Medline: 14690888]
10. Viswanath K. Public communications and its role in reducing and eliminating health disparities. In: In: Thomson GE, Mitchell F, Williams MB, editors: Examining the Health Disparities Research Plan of the National Institutes of Health: Unfinished Business. Washington, DC: National Academy Press; 2006:215-253.
11. Viswanath K, Breen N, Meissner H, Moser RP, Hesse B, Steele WR, et al. Cancer knowledge and disparities in the information age. J Health Commun 2006;11 Suppl 1:1-17. [doi: 10.1080/10810730600977517] [Medline: 16641071]
12. Clayman ML, Manganello JA, Viswanath K, Hesse BW, Arora NK. Providing health messages to Hispanics/Latinos: understanding the importance of language, trust in health information sources, and media use. J Health Commun 2010;15 Suppl 3:252-263. [doi: 10.1080/10810730.2010.522697] [Medline: 21154097]
13. Fox S, Livingston G. Latinos online. Washington, DC: Pew Internet & American Life Project; 2011 Mar 14. URL: http://www.pewinternet.org/~/media//Files/Reports/2007/Latinos_Online_March_14_2007.pdf.pdf [accessed 2011-09-16] [WebCite Cache ID 61k9Xmd9u]
14. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
15. Livingston G, Parker K, Fox S. Latinos Online-2008: narrowing the gap. Washington, DC: Pew Hispanic Center; 2006. URL: http://pewhispanic.org/files/reports/119.pdf [accessed 2011-09-16] [WebCite Cache ID 61k9soMfJ]
16. Vanderpool RC, Kornfeld J, Rutten LF, Squiers L. Cancer information-seeking experiences: the implications of Hispanic ethnicity and Spanish language. J Cancer Educ 2009;24(2):141-147. [doi: 10.1080/08858190902854772] [Medline: 19431032]
17. Tortolero-Luna G, Finney Rutten LJ, Hesse BW, Davis T, Kornfeld J, Sanchez M, et al. Health and cancer information seeking practices and preferences in Puerto Rico: creating an evidence base for cancer communication efforts. J Health Commun 2010;15 Suppl 3:30-45. [doi: 10.1080/10810730.2010.522698] [Medline: 21154082]
18. Ortiz AP, Soto-Salgado M, Calo W, Nogueras G, Tortolero-Luna G, Hebl S, et al. Disparities in breast cancer in Puerto Rico and among Hispanics, non-Hispanic whites, and non-Hispanics blacks in the United States, 1992-2004. Breast J 2010 Dec;16(6):666-668. [doi: 10.1111/j.1524-4741.2010.00990.x] [Medline: 21070450]
19. Ortiz AP, Soto-Salgado M, Calo WA, Tortolero-Luna G, Pérez CM, Romero CJ, et al. Incidence and mortality rates of selected infection-related cancers in Puerto Rico and in the United States. Infect Agent Cancer 2010;5:10 [FREE Full text] [doi: 10.1186/1750-9378-5-10] [Medline: 20470399]
20. Ortiz AP, Pérez J, Otero-Domínguez Y, García-Rodríguez O, Garced-Tirado S, Escalera-Maldonado F, et al. Endometrial cancer in Puerto Rico: incidence, mortality and survival (1992-2003). BMC Cancer 2010;10:31 [FREE Full text] [doi: 10.1186/1471-2407-10-31] [Medline: 20128912]
21. Romero Marrero C, Ortiz AP, Pérez CM, Pérez J, Torres EA. Survival of hepatocellular carcinoma in Puerto Rico. P R Health Sci J 2009 Jun;28(2):105-113. [Medline: 19530551]
22. Soto-Salgado M, Suárez E, Calo W, Cruz-Correa M, Figueroa-Vallés NR, Ortiz AP. Incidence and mortality rates for colorectal cancer in Puerto Rico and among Hispanics, non-Hispanic whites, and non-Hispanic blacks in the United States, 1998-2002. Cancer 2009 Jul 1;115(13):3016-3023 [FREE Full text] [doi: 10.1002/cncr.24340] [Medline: 19402167]
23. Suárez E, Calo WA, Hernández EY, Diaz EC, Figueroa NR, Ortiz AP. Age-standardized incidence and mortality rates of oral and pharyngeal cancer in Puerto Rico and among Non-Hispanics Whites, Non-Hispanic Blacks, and Hispanics in the USA. BMC Cancer 2009;9:129 [FREE Full text] [doi: 10.1186/1471-2407-9-129] [Medline: 19400958]
24. Torres-Cintrón M, Ortiz AP, Ortiz-Ortiz KJ, Figueroa-Vallés NR, Pérez-Irizarry J, Díaz-Medina G, et al. Using a socioeconomic position index to assess disparities in cancer incidence and mortality, Puerto Rico, 1995-2004. Prev Chronic Dis 2012 Jan;9:E15 [FREE Full text] [Medline: 22172182]
25. Hugo MH, Velasco G. A demographic portrait of Puerto Ricans, 2009. Washington, DC: Pew Hispanic Center; 2011 Jun 13. URL: http://www.pewhispanic.org/files/2011/06/143.pdf [accessed 2012-02-08] [WebCite Cache ID 65IbYVpk]
26. Rivera-Hernandez N, Andino-Ortiz VL. National Council of La Raza. 2011. 2010 Kids Count: Puerto Rico Data Book URL: http://www.nclr.org/index.php/publications/2010_kids_count_-_puerto_rico_data_book/ [accessed 2012-02-08] [WebCite Cache ID 65IpYOk8v]

27. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. Am J Epidemiol 2002 Sep 1;156(5):471-482 [FREE Full text] [Medline: 12196317]
28. Research Triangle Institute. RTI International. SUDAAN version 9. 0 URL: http://www.rti.org/sudaan/page.cfm/SUDAAN_9 [accessed 2012-07-13] [WebCite Cache ID 6983Zn5Hg]
29. US Census Bureau. 2010 Sep 27. Current Population Survey URL: http://www.census.gov/cps/ [accessed 2011-10-19] [WebCite Cache ID 62YgQHzJz]
30. Kaw AK, Kalu EE, Besterfield G. Numerical Methods with Applications: 2nd edition. Raleigh, NC: Lulu Enterprises, Inc; 2008.
31. Internet World Stats. Miniwatts Marketing Group. 2011. Usage and Population Statistics URL: http://www.internetworldstats.com/ [accessed 2011-09-29] [WebCite Cache ID 623kf7DcU]
32. Fox S. Online Health Search 2006. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. URL: http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Online_Health_2006.pdf [accessed 2011-09-16] [WebCite Cache ID 61k9ytRo4]
33. Health Information Trends Survey. National Cancer Institute. All HINTS Questions URL: http://hints.cancer.gov/questions.aspx [accessed 2012-07-16] [WebCite Cache ID 69CYwlsRi]
34. SMEI home page. Sales & Marketing & Executives International, Inc. 2010 URL: http://www.smei.org/ [accessed 2012-07-13] [WebCite Cache ID 6983eNBj1]
35. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System
HINTS: Health Information National Trends Survey

Edited by G Eysenbach; submitted 22.11.11; peer-reviewed by N Atkinson; comments to author 15.01.12; revised version received 13.02.12; accepted 10.03.12; published 19.07.12.

Please cite as:
Finney Rutten LJ, Hesse BW, Moser RP, Ortiz Martinez AP, Kornfeld J, Vanderpool RC, Byrne M, Tortolero Luna G
Socioeconomic and Geographic Disparities in Health Information Seeking and Internet Use in Puerto Rico
J Med Internet Res 2012;14(4):e104
URL: http://www.jmir.org/2012/4/e104/
doi:10.2196/jmir.2007
PMID:22849971

Original Paper

Computer Skills and Internet Use in Adults Aged 50-74 Years: Influence of Hearing Difficulties

Helen Henshaw[1], BSC(Hons), PhD; Daniel P A Clark[1], BSC(Hons), MRes, PhD; Sujin Kang[1], MSc; Melanie A Ferguson[2], BSC, MSc

[1]NIHR National Biomedical Research Unit in Hearing, School of Clinical Sciences, University of Nottingham, Nottingham, United Kingdom
[2]NIHR National Biomedical Research Unit in Hearing, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom

Corresponding Author:
Helen Henshaw, BSC(Hons), PhD
NIHR National Biomedical Research Unit in Hearing
School of Clinical Sciences
University of Nottingham
Ropewalk House
113 The Ropewalk
Nottingham, NG1 5DU
United Kingdom
Phone: 44 (0)115 8232600
Fax: 44 (0)115 8232615
Email: helen.henshaw@nottingham.ac.uk

Abstract

Background: The use of personal computers (PCs) and the Internet to provide health care information and interventions has increased substantially over the past decade. Yet the effectiveness of such an approach is highly dependent upon whether the target population has both access and the skill set required to use this technology. This is particularly relevant in the delivery of hearing health care because most people with hearing loss are over 50 years (average age for initial hearing aid fitting is 74 years). Although PC skill and Internet use by demographic factors have been examined previously, data do not currently exist that examine the effects of hearing difficulties on PC skill or Internet use in older adults.

Objective: To explore the effect that hearing difficulty has on PC skill and Internet use in an opportunistic sample of adults aged 50-74 years.

Methods: Postal questionnaires about hearing difficulty, PC skill, and Internet use (n=3629) were distributed to adults aged 50-74 years through three family physician practices in Nottingham, United Kingdom. A subsample of 84 respondents completed a second detailed questionnaire on confidence in using a keyboard, mouse, and track pad. Summed scores were termed the “PC confidence index.” The PC confidence index was used to verify the PC skill categories in the postal questionnaire (ie, never used a computer, beginner, and competent).

Results: The postal questionnaire response rate was 36.78% (1298/3529) and 95.15% (1235/1298) of these contained complete information. There was a significant between-category difference for PC skill by PC confidence index (*P*<.001), thus verifying the three-category PC skill scale. PC and Internet use was greater in the younger respondents (50-62 years) than in the older respondents (63-74 years). The younger group’s PC and Internet use was 81.0% and 60.9%, respectively; the older group’s PC and Internet use was 54.0% and 29.8%, respectively. Those with slight hearing difficulties in the older group had significantly greater odds of PC use compared to those with no hearing difficulties (odds ratio [OR]=1.57, 95% confidence interval [CI] 1.06-2.30, *P*=.02). Those with moderate+ hearing difficulties had lower odds of PC use compared with those with no hearing difficulties, both overall (OR=0.58, 95% CI 0.39-0.87, *P*=.008) and in the younger group (OR=0.49, 95% CI 0.26-0.86, *P*=.008). Similar results were demonstrated for Internet use by age group (older: OR=1.57, 95% CI 0.99-2.47, *P*=.05; younger: OR=0.32, 95% CI 0.16-0.62, *P*=.001).

Conclusions: Hearing health care is of particular relevance to older adults because of the prevalence of age-related hearing loss. Our data show that older adults experiencing slight hearing difficulty have increased odds of greater PC skill and Internet use than those reporting no difficulty. These findings suggest that PC and Internet delivery of hearing screening, information, and intervention is feasible for people between 50-74 years who have hearing loss, but who would not typically present to an audiologist.

(J Med Internet Res 2012;14(4):e113) doi:10.2196/jmir.2036

KEYWORDS

hearing loss; hearing difficulties; screening; intervention; hearing health care; eHealth; personal computer; Internet use.

Introduction

The use of personal computers (PCs) and the Internet to provide health care and health-related information to patients and the public has increased substantially over the last decade [1,2], to the point where the Internet is now a major source of health information [2]. Hearing health is no exception. Current examples of PC and Internet delivery of hearing health care include hearing screening [3-5], auditory training [6], counseling [7], education [8], and information delivery [9,10]. A review of eHealth in audiology suggests that published evidence assessing online hearing health care finds these applications to be both reliable and effective [11].

Sensorineural hearing loss (SNHL) is highly associated with age, sex, and socioeconomic status (SES) [12]. Hearing loss affects approximately 1 in 4 people (27%) between 55-74 years, it increases in severity with age, is worse in males than females, and is more prevalent in those with lower SES [13]. This is a growing problem in our aging society, with the number of people with SNHL set to increase even further in the future [12,14]. The main clinical management intervention for people with hearing loss is hearing aids, but most adults aged 55-74 years (80%) do not use them [12]. The reasons for this are broad and varied, including inadequate access to audiology services because of low referral rates by family practice physicians or general practitioners (GPs), a perception that poor hearing is not that bad, acceptance of hearing loss as a normal consequence of aging, a lack of awareness of hearing loss and what to do about it, and simply not wanting to address the issue [12]. The average age of a first-time hearing aid user is 74 years, yet these adults may have suffered with significant hearing loss for an average of 10 years before receiving hearing aids [12]. A major advantage of delivering hearing health care through the Internet is the potential to increase accessibility to large numbers of people with hearing loss, many of whom do not, or cannot, access current hearing health care sources.

Hearing aids are not the only form of intervention available for hearing loss. Computerized auditory training, such as Listening and Communication Enhancement (LACE) [15], which can change a person's ability to process sounds and can improve auditory performance on a trained task [16], is an example of an alternative intervention strategy that may help alleviate hearing difficulties [17,18] and is ideally suited for Internet delivery [6]. Other types of online hearing-related interventions include Internet discussion groups, counseling, and information provision. These interventions have demonstrated effectiveness in improving self-reported hearing abilities and satisfaction with amplification in experienced hearing aid users (mean age 63.5 years) [8], and in reinforcing positive adjustment behaviors in new hearing aid users (mean age 68 years) [7]. Yet online hearing information and hearing health care can only be effective if they are accessible and usable by the target population. The earlier an adult with SNHL begins a program of rehabilitation, the greater chance their hearing disability will be reduced and their quality of life improved [19]. Currently, there are no guidelines for the screening of age-related hearing loss for adults despite the success of newborn and school-based hearing screening programs that have been highly effective in identifying children who are deaf or hearing impaired [20,21]. Online hearing screening for adults (eg, the UK Action on Hearing Loss hearing screen [3] and the Dutch functional hearing-screening test [5]) at the onset of hearing difficulties (typically around 50 years of age) may offer a cost-effective strategy to promote early diagnosis of progressive SNHL and, in turn, may lead to earlier intervention and better quality of life [22].

Severity of age-related SNHL increases with age, thus the hearing health care requirements of adults are likely to change as they get older. Davis et al [12] showed that approximately 5% of 55-64 year olds experience a significant hearing impairment of ≥35 decibels (dB) in both ears. This increases for those between 65-69 years and between 70-74 years, with the prevalence of a significant hearing loss of ≥35dB hearing level (HL) rising to 15% and 20%, respectively. Adults over the age of 65 years with hearing loss are statistically more likely to benefit from amplification than those younger than 65 years [12]. As such, hearing aid intervention and support related to hearing aids are likely to be appropriate forms of hearing health care for adults in their mid-60s and 70s. However, 1 in 3 adults between 55-64 years experiences hearing losses between 20-34dB [23]. Although classed as “mild” impairment, this can still lead to reduced social interaction, participation, and quality of life [12]. Consequently, it is likely that adults in their 50s and early 60s would benefit from hearing screening and advice regarding hearing loss designed to address the onset of age-related hearing difficulties.

Recent evidence suggests that the Internet may provide a means to facilitate communication in people with hearing loss because it removes the auditory barrier [24] and appeals to those with text-based communication preferences [25]. Nevertheless, research focusing on the association between Internet use and hearing impairment has been confined to an adolescent population to date [24]. It is unknown whether these findings are applicable to the majority of people with age-related hearing loss, especially those who are over the age of 50 years and have mild to moderate losses. Although there is some evidence that PC and Internet use is affected by age, SES, and sex [26,27], there is no published evidence examining whether PC and Internet use are affected by hearing difficulty in older adults. As hearing loss is also associated with age, sex, and SES, it is important to examine the combined effects of all four factors on levels of PC skill and Internet use. This will enable any effects of hearing difficulties to be identified while controlling for any confounding demographic factors.

The primary aim of the present study was to explore the relationships among hearing difficulties and both PC skill and

Internet use in an older adult population after accounting for the confounding demographic factors of age, SES, and sex. Should hearing difficulties be related to levels of PC skill and Internet use, it is hypothesized that individuals with hearing difficulty will have greater PC skill and Internet use than those reporting no difficulty. Furthermore, those effects may be greater as the degree of hearing difficulty increases.

Level of PC skill was defined in this study by using a three-category scale (ie, never used a computer, beginner, or competent). The category "never used a computer" is clear. However, the distinction between the "beginner" and the "competent" PC user categories is less well defined. As degree of computer use has been demonstrated to affect computer attitude measures including confidence with computers [28,29], a secondary aim was to verify this three-category PC skill scale by examining user confidence with PCs.

Methods

Postal Questionnaire

The primary measure was a 16-item postal questionnaire (Appendix 1) designed as a recruitment tool for a study assessing the benefits of auditory training for adults aged between 50-74 with hearing loss. The questionnaire included five items on hearing difficulties that were used in a hearing screening questionnaire by Davis and colleagues [12]. Previous research using these hearing screening questions found that reports of slight difficulty were associated with hearing losses <35dBHL at 3kHz, whereas reports of moderate difficulty or greater were predictive of hearing loss ≥35dBHL at 3kHz, with sensitivity of 78% [12]. The postal questionnaire also requested details regarding respondents' PC skill level, Internet use, demographic information, and permission to contact for further research. No details regarding the auditory training study were included in the postal questionnaire, thus eliminating any potential for the auditory training study to influence questionnaire response rates. Instead, those respondents who agreed to be contacted for further research were later contacted to inform them about the study.

Three family practice physician offices in Nottingham, United Kingdom, participated in study recruitment. Postal questionnaires were sent to a total of 3529 patients on the family practice patient registers aged between 50-74, together with an invite letter from the family practice physician (GP in the United Kingdom) and the lead researcher (MAF). Patients were invited to complete and return the questionnaire in an enclosed reply-paid envelope. Non-respondents were not followed up; non-response was assumed to indicate a desire not to participate in the survey.

Participants

The response rate to the postal questionnaire was 36.77% (1298/3529), which is comparable with response rates for a recent national postal survey of family practice patients [30]. A total of 63 questionnaires were excluded from the analysis: 18 because the respondents did not complete the PC or Internet questions and 45 because the respondents fell outside of the target age bracket (50-74 years). Data from 1235 respondents were used in further analyses.

Respondents' ages ranged from 50-74 years with a mean of 62.2 years (SD6.6 years). There were more female respondents (54.49%, 673/1235) than male (45.26%, 559/1235), but 3/1235 (0.25%) of respondents failed to report their sex. SES was determined by using the Index of Multiple Deprivation (IMD) score based on respondents' postal codes. The IMD is a measure of deprivation by area, with higher IMD scores equating to lower SES, or greater deprivation. In our sample, IMD scores ranged between 3.65 and 78.37 with a mean of 27.87 (SD17.10). Nationally, IMD scores range between 0.99 and 84.22 [31].

PC Confidence Questionnaire

PC skill was rated in the postal questionnaire on a three-category scale (ie, never used a computer, beginner, or competent). We aimed to verify this PC skill scale by examining whether individuals' confidence in using a PC differed significantly among these three categories of PC skill.

A second 11-item PC confidence questionnaire (Appendix 2) was developed to obtain more detailed information regarding respondents' PC use and was administered to a subsample of the postal questionnaire respondents (n=85) who were invited to take part in the auditory training study. Participants completed the questionnaires in the waiting room. The questionnaire comprised of closed-set questions to assess overall confidence using a PC and confidence in using a keyboard, a mouse, and a laptop track pad. Ratings ranged from 0-3: 0 is not confident at all; 1 is I usually need help; 2 is it takes me a while but I can manage; and 3 is confident. Ratings for the four confidence items were summed to form a PC confidence index that ranged from 0-12 points. These data were used to address our secondary aim, which was to verify the three-category PC skill scale used in the primary postal questionnaire. One respondent did not complete the PC confidence questions; therefore, data from 84 respondents were included in subsequent analyses.

Participants

Respondents ranged from 50-74 years with a mean age of 63.8 years (SD6.4). A total of 52/84 (62%) respondents were male and 32/84 (38%) were female. IMD scores ranged from 3.84-67.73 with a mean of 25.14 (SD17.45).

Statistical Analyses

For all analyses, alpha<.05 was considered statistically significant. Initial correlations using Spearman rank correlation (ρ) were conducted to identify associations among PC skill, Internet use, better-ear hearing difficulty (BEHD), and demographic factors.

Main analyses assessed levels of PC skill and Internet use by reported hearing difficulty. Associated factors of age, SES, and sex were included within these analyses. Initially, univariate explorations were used to assess the individual relationship between each factor and PC skill and Internet use. All factors were then pooled within multivariate regression analyses to control for any confounding effects between factors. Through a backward elimination process, factors that were considered to be statistically significant (Wald z-statistic; P<.05) were retained within the multivariate logistic regression analyses, and the Akaike information criterion (AIC) was then used to assess relative goodness-of-fit to determine an optimized model. Finally, respondents were divided by median age to investigate the relationship among hearing difficulties and levels of PC skill and Internet use in the "younger" (50-62 years) and the "older" (63-74 years) portions of the sample.

Secondary analyses assessed the three PC skill levels by using respondents' PC confidence indexes from the subset of respondents who completed the PC confidence questionnaire (n=84). The PC confidence index for each respondent was compared to their selected PC skill category. Analysis of variance (ANOVA) was used to assess any significant differences between categories. These analyses sought to identify whether self-selected PC skill categories were an accurate reflection of the respondents' confidence in using a PC.

Results

Verification of Self-reported PC Skill Categories

Although verification of the self-reported PC skill categories was a secondary aim, the results for the PC confidence questionnaire respondents (n=84) are presented first because they form the basis of the main analyses to follow.

Mean PC confidence scores by reported PC skill categories are shown in Table 1. A 1-way ANOVA revealed a highly significant between-category difference ($F_{2,81}=69.78$; $P<.001$). Respondents who classed themselves as competent PC users scored highest on the PC confidence index, followed by beginners, and then those who had never used a computer (Figure 1). This suggests that respondents were selecting their PC skill levels appropriately based upon their confidence in using a PC and it provides evidence that the PC skill levels selected by respondents in the postal questionnaire were valid reflections of their confidence in using a PC.

Factors Associated with PC Skill and Internet Use

The primary objective of this study was to explore the effects of BEHD on PC skill and Internet use for the 1235 postal questionnaire respondents.

Degree of hearing difficulty was categorised on reported hearing difficulties in the better ear: no hearing difficulty, slight difficulty, and moderate or greater difficulties (moderate+ difficulty). Table 2 shows the frequency of postal questionnaire respondents by BEHD level, age, SES, and sex by PC skill and Internet use. Figure 2 presents these data as percentage of respondents.

Prevalence of hearing difficulties was 26.80% overall (331/1235), which is comparable to data from a large UK hearing population study [12], and was slightly greater in females (27.6%, 186/673) than in males (26.1%, 146/559). For those respondents who reported no hearing difficulty, 45.7% (413/904) were male and 53.9% (487/904) were female. Despite a greater number of female than male respondents overall, for those reporting slight difficulty, 56.4% (128/904) were male and 43.6% (99/904) were female, and for those reporting moderate+ hearing difficulty, 63.4% (64/904) were male and 36.6% (37/904) were female. This represents a statistically significant difference in the prevalence of hearing difficulties, with difficulties being reported more often by male than by female respondents ($\chi^2_2=32.6$, $P<.001$).

Over two-thirds of our sample (839/1235, 67.94%) reported being PC users, either beginner or competent, and 45.83% (566/1235) used the Internet. These figures are consistent with existing literature on PC and Internet use in older adults [26,27]. There was a decline in both PC and Internet use with increasing age. In the younger group (50-62 years), PC use (81.0%, 516/637) and Internet use (60.9%, 388/637) was greater than PC use (54.0%, 323/598) and Internet use (29.8%, 178/598) in the older group (63-74 years). For the youngest 5-year age range (50-54 years), 84.6% (165/195) used PCs and 65.6% (128/195) were Internet users, whereas for those respondents in the oldest 5-year age range (70-74 years)—representing the typical ages of first-time hearing aid users—PC use was 36.3% (77/212) and Internet use was 17.5% (37/212).

Respondents who reported slight hearing difficulty were equally likely to rate their PC skill as "never used a computer" (30.6%, 70/229) as those reporting no hearing difficulty (30.8%, 278/904), and equally likely to rate their PC skill as "competent" (40.2%, 92/229) as those reporting no hearing difficulty (40.0%, 362/904). However, those respondents who reported moderate+ hearing difficulty were more likely to rate their PC skill as "never used a computer" (47.1%, 48/102) and less likely to rate their PC skill as "competent" (24.5%, 25/102).

Similarly, for Internet use, respondents who reported slight hearing difficulty were equally likely to use the Internet (45.9%, 105/229) as those reporting no hearing difficulty (47.5%, 429/904). However, those respondents who reported moderate+ hearing difficulty were less likely to report being an Internet user (31.4%, 32/102).

Table 1. Mean PC confidence scores (range 0-3; 0=not at all confident to 3=confident) and PC confidence index (range 0-12) by PC skill level for computer confidence questionnaire respondents (n=84).

PC confidence measure	Self-reported PC skills mean (SD)			Mean confidence mean (SD)
	Never	Beginner	Competent	
Keyboard	0.24 (0.56)	2.03 (0.85)	2.75 (0.60)	1.98 (1.17)
Mouse	0.64 (1.01)	2.21 (0.92)	2.91 (0.28)	2.25 (1.09)
Track pad	0.31 (0.79)	1.05 (1.25)	2.19 (0.97)	1.40 (1.28)
Overall PC confidence	0.24 (0.44)	1.23 (0.77)	2.49 (0.66)	1.56 (1.10)
PC confidence index[a]	1.22 (1.93)	6.10 (3.01)	9.94 (2.47)	6.70 (4.21)

[a] sum of four confidence scores

Figure 1. Mean and median PC confidence index by PC skill level for computer confidence questionnaire respondents (n=84).

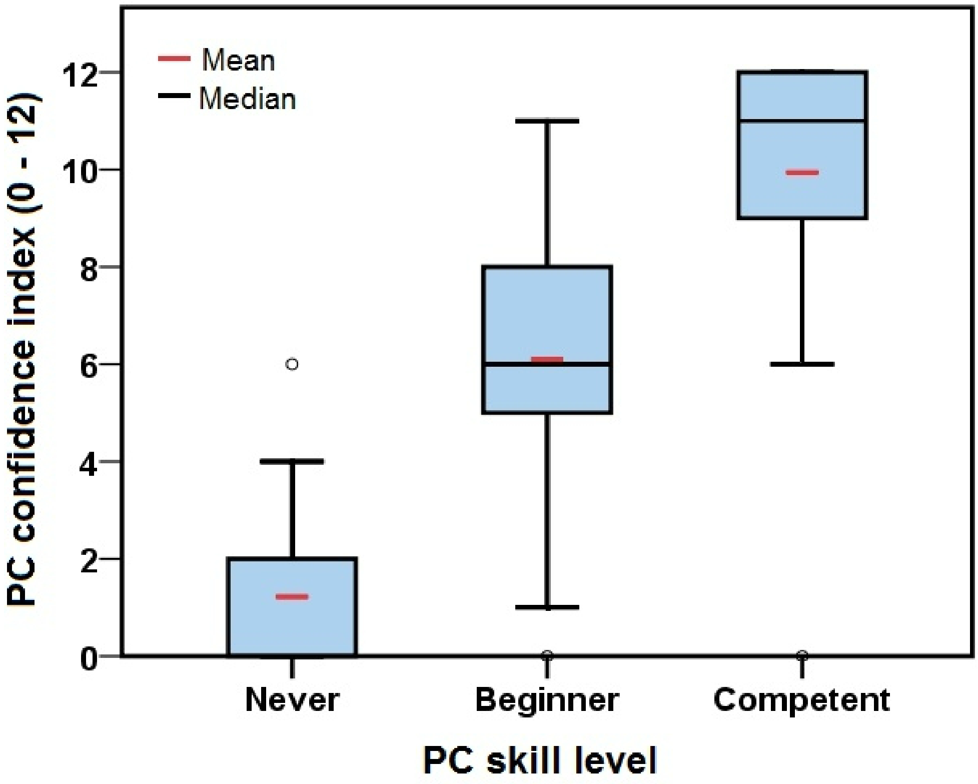
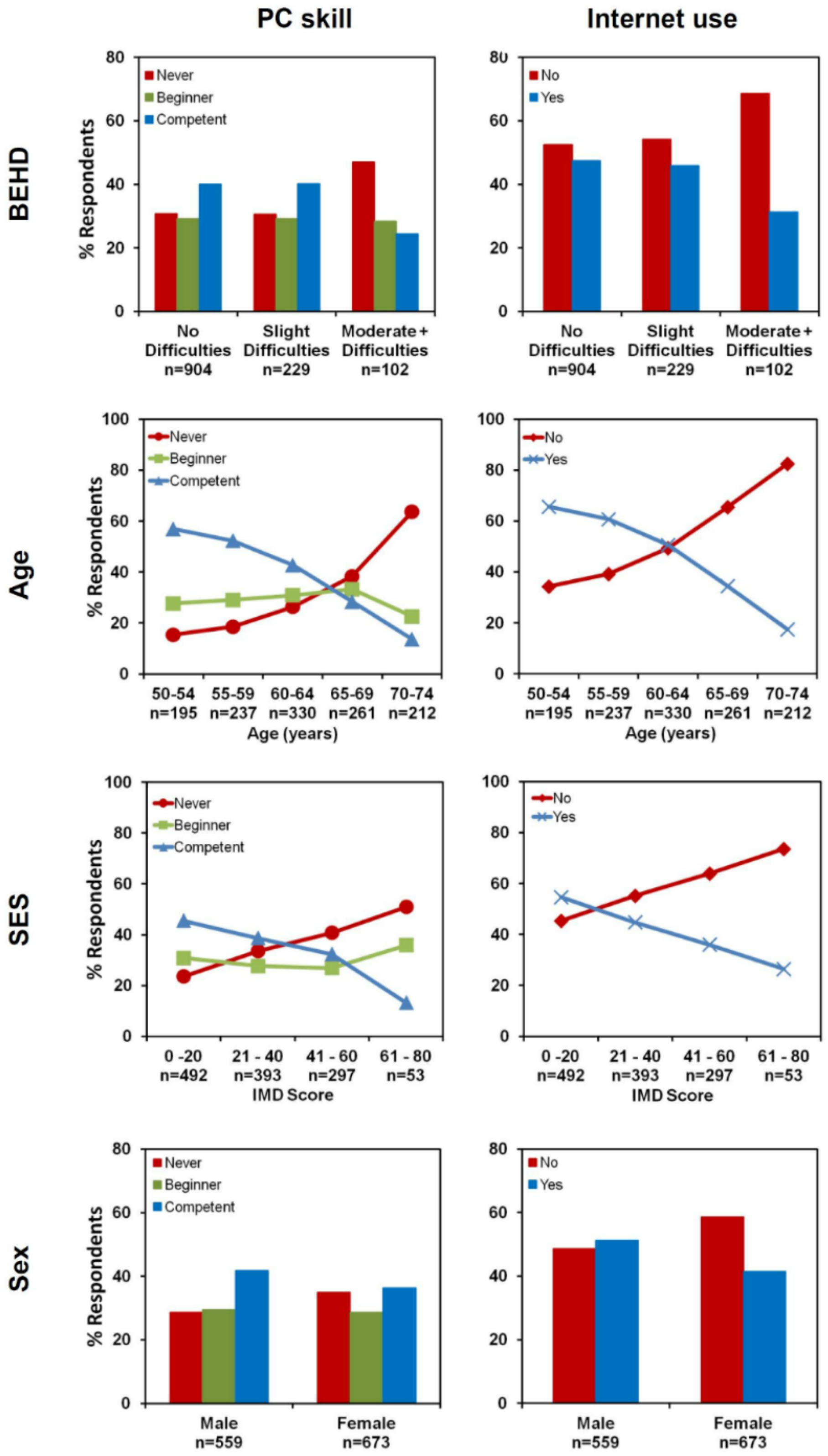


Table 2. Better-ear hearing difficulty (BEHD) and demographics of postal questionnaire respondents (n=1235) by PC skill and Internet use.

Characteristics		Total n (%)	PC skill n (%)			Internet use n (%)	
			Never 396 (32.1)	Beginner 360 (29.1)	Competent 479 (38.8)	No 669 (54.2)	Yes 566 (45.8)
BEHD	None	904 (73.20)	278 (70.2)	264 (73.3)	362 (75.6)	475 (71.0)	429 (75.8)
	Slight	229 (18.54)	70 (17.7)	67 (18.6)	92 (19.2)	124 (18.5)	105 (18.6)
	Moderate+	102 (8.26)	48 (12.1)	29 (8.1)	25 (5.2)	70 (10.5)	32 (5.6)
Age (years)	50-54	195 (15.79)	30 (7.6)	54 (15.0)	111 (23.2)	67 (10.0)	128 (22.6)
	55-59	237 (19.19)	44 (11.1)	69 (19.2)	124 (25.9)	93 (13.9)	144 (25.5)
	60-64	330 (26.72)	87 (21.9)	102 (28.3)	141 (29.4)	163 (24.4)	167 (29.5)
	65-69	261 (21.11)	100 (25.3)	87 (24.2)	74 (15.4)	171 (25.5)	90 (15.9)
	70-74	212 (17.17)	135 (34.1)	48 (13.3)	29 (6.1)	175 (26.2)	37 (6.5)
SES	0-20	492 (39.84)	116 (29.3)	152 (42.2)	224 (46.8)	223 (33.3)	269 (47.5)
	21-40	393 (31.82)	132 (33.3)	109 (30.3)	152 (31.7)	217 (32.4)	176 (31.1)
	41-60	297 (24.05)	121 (30.6)	80 (22.2)	96 (20.0)	190 (28.4)	107 (18.9)
	61-80	53 (4.29)	27 (6.8)	19 (5.3)	7 (1.5)	39 (5.8)	14 (2.5)
Sex[a]	Male	559 (45.26)	160 (40.4)	165 (45.8)	234 (48.9)	272 (40.7)	287 (50.7)
	Female	673 (54.49)	235 (59.3)	193 (53.6)	245 (51.1)	394 (58.9)	279 (49.3)

[a] 0.25% (3/1235) of respondents failed to report their sex

Figure 2. Percentage of postal questionnaire respondents (n = 1235), by PC skill level and Internet use, by better-ear hearing difficulty (BEHD) and demographic factors.



PC Skill Level

Although PC skill was shown to be significantly associated with Internet use in our sample (ρ=0.74, n=1235, *P*<.001), the present research aimed to assess any effects of hearing difficulties on both the skill set and the access of older adults to hearing health care delivered through PCs or the Internet. As such, factors relating to PC skill and Internet use were independently assessed within this study.

Univariate ordinal logistic regression analyses were conducted to examine any significant effects of BEHD, age, SES, and sex on reported PC skill. Results revealed a significant effect of moderate+ hearing difficulty (odds ratio [OR]=0.49, 95% confidence interval [CI] 0.34-0.72, *P*=.001), but no effect of slight hearing difficulty (OR=1.00, 95% CI 0.76-1.30, *P*=.97) on PC skill. There were also significant effects of age (OR=0.90, 95% CI 0.88-0.91, *P*<.001), SES (OR=0.98, 95% CI 0.97-0.99, *P*<.001), and sex (OR=0.77, 95% CI 0.63-0.95, *P*=.015) on PC skill.

All factors were combined within a multivariate ordinal logistic regression model to assess the effects of hearing difficulty on PC skill while controlling for any confounding effects attributable to age, SES, and sex. To aid interpretation, PC skill was categorized as either PC "use" (those with beginner or competent PC skills) or "non-use" (respondents who have never used a computer). Within the categories of BEHD and sex, "none" (no difficulty) and "male" were used as the baselines for comparison, respectively. Results are presented in Table 3 for the whole sample, and for the younger and older groups.

Table 3. Logistic regression coefficients (beta) and odds ratio estimates from the multivariate ordinal logistic regression models for factors affecting PC skill level.

Characteristics		All (50-74 years) n=1235			Younger (50-62 years) n=637			Older (63-74 years) n=598		
		beta	Odds ratio (95% CI)	*P*	beta	Odds ratio (95% CI)	*P*	beta	Odds ratio (95% CI)	*P*
BEHD	None	–	1.00 (–)	–	–	1.00 (–)	–	–	1.00 (–)	–
	Slight	.13	1.14 (0.86-1.51)	.37	-.24	0.79 (0.52-1.20)	.26	.45	1.57 (1.06-2.30)	.02
	Moderate+	-.54	0.58 (0.39-0.87)	.008	-.74	0.47 (0.26-0.86)	.008	-.36	0.70 (0.40-1.20)	.20
Age		-.11	0.90 (0.88-0.91)	<.001	-.04	0.96 (0.92-0.99)	.05	-.16	0.85 (0.81-0.89)	<.001
SES		-.02	0.98 (0.97-0.99)	<.001	-.02	0.98 (0.97-0.99)	<.001	-.02	0.98 (0.97-0.99)	<.001
Sex	Male	–	1.00 (–)	–	–	1.00 (–)	–	–	1.00 (–)	–
	Female	-.28	0.75 (0.61-0.94)	.01	-.19	0.83 (0.61-1.12)	.23	-.39	0.68 (0.49-0.93)	.02

All Respondents

Results from the multivariate analyses revealed a significant effect of moderate+ hearing difficulty on PC skill level when no difficulty was used as the baseline measure (OR=0.58, 95% CI 0.39-0.87, *P*=.008). This suggests that the odds of being a PC user are 0.58 times less for those reporting moderate+ hearing difficulty than those reporting no hearing difficulty. There was no significant effect of slight hearing difficulty (OR=1.14, 95% CI 0.86-1.51, *P*=.37), suggesting that PC skill did not significantly differ between respondents with slight hearing difficulty and those with no hearing difficulty.

Despite the slight hearing difficulty category showing no significant association with PC skill, a comparison using the likelihood ratio test (LRT) on AIC estimates of a model with and without BEHD revealed a significant difference between models (LRT=8.71, df=2, *P*=.01), suggesting that BEHD was a significant factor associated with PC skill in the multivariate model.

Younger Group (50-62 Years)

There was a significant effect of moderate+ hearing difficulty (OR=0.47, 95% CI 0.26-0.86, *P*=.01), but no significant effect of slight hearing difficulty (OR=0.79, 95% CI 0.52-1.20, *P*=.26) on PC skill level in the younger group.

Although age and SES were significant factors related to PC skill, there was no difference between levels of PC skill for males and females (OR=0.83, 95% CI 0.61-1.12, *P*=.23). Further analysis by using the LRT on AIC estimates showed a model including the factor sex did not perform significantly better than a model with this factor removed (LRT=1.47, df=1, *P*=.23). Therefore, sex was eliminated from the final model.

The final model showed a significant effect of moderate+ hearing difficulty (OR=0.49, 95% CI 0.27-0.89; *P*=.02), age (OR=0.96, 95% CI 0.92-0.99; *P*=.04), and SES (OR=0.98, 95% CI 0.97-0.99; *P*<.001) on PC skill level. These results suggest that for those younger respondents with moderate+ hearing difficulty, the odds of being a PC user over a non-user were significantly less than for those with no difficulty hearing.

Older Group (63-74 Years)

For the older group, there were no significant effects of moderate+ hearing difficulty (OR=0.70, CI 0.40-1.20, *P*=.20) on PC skill. However, for those with slight hearing difficulty, the odds of being a PC user over a non-user were significantly greater (OR=1.57, 95% CI 1.06-2.30, *P*=.02) than for those with no hearing difficulty.

Internet Use

Univariate logistic regression analyses revealed a significant effect of moderate+ hearing difficulty (OR=0.51, 95% CI 0.32-0.78, *P*=.01), but no effect of slight hearing difficulty (OR=0.94, 95% CI 0.70-1.25, *P*=.66), on Internet use. Thus, there were no differences between Internet use for respondents with slight hearing difficulty and those with no hearing difficulty. However, for those respondents reporting moderate+ hearing difficulty, the odds of being an Internet user were 0.51 times less than those reporting no hearing difficulty. Demographic factors of age (OR=0.90, 95% CI 0.88-0.91, *P*<.001), SES (OR=0.98, 95% CI 0.97-0.99, *P*<.001), and sex (OR=0.67, 95% CI 0.54-0.84, *P*=.001) were also shown to be significantly associated with Internet use in the univariate analyses.

All factors were combined within a multivariate ordinal logistic regression model to assess the effects of hearing difficulty on Internet use while controlling for any confounding effects of age, SES, and sex. Results are presented in Table 4 for the whole sample, and the younger and older groups.

Table 4. Logistic regression coefficients (beta) and odds ratio estimates from the multivariate logistic regression model for factors affecting Internet use.

Characteristics		All (50-74 years) n=1235			Younger (50-62 years) n=637			Older (63-74 years) n=598		
		beta	Odds ratio (95% CI)	*P*	beta	Odds ratio (95% CI)	*P*	beta	Odds ratio (95% CI)	*P*
BEHD	None	–	1.00 (–)	–	–	1.00 (–)	–	–	1.00 (–)	–
	Slight	.02	1.02 (0.74-1.14)	.90	-.35	0.70 (0.45-1.10)	.12	.45	1.57 (0.99-2.47)	.05
	Moderate+	-.55	0.58 (0.36-0.92)	.02	-.74	0.32 (0.16-0.62)	.001	.05	1.05 (0.54-1.96)	.88
Age		-.11	0.89 (0.88-0.91)	<.001	-.04	0.96 (0.92-0.99)	.04	-.13	0.87 (0.83-0.92)	<.001
SES		-.02	0.98 (0.97-0.98)	<.001	-.02	0.98 (0.97-0.99)	<.001	-.03	0.97 (0.96-0.98)	<.001
Sex	Male	–	1.00 (–)	–	–	1.00 (–)	–	–	1.00 (–)	–
	Female	-.47	0.62 (0.49-0.80)	<.001	-.45	0.64 (0.46-0.89)	.009	-.51	0.60 (0.41-0.87)	.007

All Respondents

There was a significant effect of moderate+ hearing difficulty on Internet use (OR=0.58, 95% CI 0.36-0.92, *P*=.02). However, as with the findings for PC skill, there was no significant effect of slight hearing difficulty (OR=1.02, 95% CI 0.74-1.14, *P*=.90), suggesting Internet use did not significantly differ between those respondents with slight hearing difficulty and those with no hearing difficulty.

Further analysis by using error sum of squares (SSE) on AIC estimates revealed that a model including BEHD did not perform significantly better than a model with this factor removed (SSE=5.49, df=2, *P*=.06). As such, BEHD was eliminated from the final model.

Results from the final model show a significant effect of age (OR=0.89, 95% CI 0.88-0.91, *P*<.001), SES (OR=0.98, 95% CI 0.97-0.98, *P*<.001), and sex (OR=0.64, 95% CI 0.50-0.82, *P*<.001) on Internet use. These results suggest that despite respondents with slight hearing difficulty having significantly greater odds of being Internet users than those with no hearing difficulty overall, this was not a significant predictor of Internet use in the multivariate model.

Younger Group (50-62 Years)

There was a significant effect of moderate+ hearing difficulty (OR=0.32, 95% CI 0.16-0.62, *P*=.001), but no significant effect of slight hearing difficulty (OR=0.70, 95% CI 0.45-1.10, *P*=.12) on Internet use. This indicates that for the younger group, the odds of being an Internet user over a non-user were significantly less for respondents with moderate+ hearing difficulties when compared with those with no hearing difficulty. These results are similar to those shown for PC skill.

Older Group (63-74 Years)

For the older group, there was a significant effect of slight hearing difficulty (OR=1.57, 95% CI 0.99-2.47, *P*=.05), but no significant effect of moderate+ hearing difficulty (OR=1.05, 95% CI 0.54-1.96, *P*=.89) on Internet use. This indicates that for those reporting slight hearing difficulties in the older group, the odds of being an Internet user over a non-user were greater than for those with no hearing difficulty. These results are similar to those for PC skill.

Discussion

Hearing-related interventions are of particular relevance to older adults because of the increasing prevalence of hearing loss with age, particularly beyond 50 years. Our postal questionnaire survey identified that for the younger group (aged 50-62 years), levels of PC and Internet use were high (81.0% and 60.9%, respectively), whereas for the older group (63-74 years), PC and Internet use was considerably less (54.0% and 29.8%).

Hearing difficulties were shown to be significantly associated with PC skill and Internet use after controlling for effects of age, sex, and SES, all of which are significant factors in hearing loss. The findings can be categorized into two main effects. First, those with slight hearing difficulties had significantly increased PC and Internet use compared with those with no hearing difficulties in the older group only. Second, those with moderate+ hearing difficulty (ie, moderate/great/cannot hear at all) had lower PC and Internet use than those with no hearing difficulties, both for the whole sample and within the younger group. Although literature exists that examines the effects of age on PC and Internet use [26,27], any interplay between demographic factors and hearing difficulties on PC and Internet use have not been examined previously in an adult population to our knowledge. The current study offers a novel investigation as to how demographic factors and hearing difficulties are associated with PC skill and Internet use in a large sample of adults aged 50-74 years(n=1235). This provides evidence to underpin and establish the potential for PC and Internet delivery of hearing health care information and interventions for people with hearing difficulties.

The postal questionnaire sample was similar to those in other studies for response rate [30], prevalence of hearing difficulties [12], and PC and Internet use [26,27]. For example, a European study of adults aged 55-74 years reported Internet use at 37% [26], which was similar to the equivalent age group in our sample (35.5%). A US study with a similar sample size for those over 50 years (no upper limit), showed PC use (including iPad and smartphone use) was 73% and Internet use was 48.6% [27], whereas Internet use in our present sample (50-74 years) was 45.8%.

PC skill and Internet use differed according to age and hearing difficulty in our sample. Respondents with slight hearing difficulties in the older group had increased odds of PC use (OR=1.57, 95% CI 1.06-2.30, *P*=.02) and Internet use (OR=1.57, 95% CI 0.99-2.47, *P*=.05) compared with those reporting no hearing difficulties. This is consistent with previous research reporting increased motivation to use the Internet and increased Internet use in adolescents with hearing loss [24]. One reason for the increased Internet use suggested by Barak and Sadovsky [24] is that PCs and the Internet provide people with hearing loss with a means of communication that is primarily visual. This puts them on an equal communication basis with people without hearing loss so their hearing disability ceases to be an issue when communicating in the visual modality. The use of PCs and the Internet has been shown to be important for the communication needs of older adults with hearing loss. Pilling and Barrett [25] showed that text-based communication preferences of older adults with severe to profound hearing loss differed from those of adolescents with comparable hearing losses. Almost half (44%) of the adults aged 50-69, preferred to communicate through email, whereas adolescents aged 15-18 preferred short message service (SMS) text messages (64%). Only 12% of adolescents selected email as their preferred means of text-based communication.

There are 10 million people in the United Kingdom with significant hearing loss, yet only 2 million have hearing aids and just 1.4 million use their hearing aids regularly [14]. The result is a huge unmet need, which could—at least in part—be addressed by online hearing health care either outside or within the current UK model of National Health Service (NHS) hearing care provision. Untreated hearing loss can be a significant problem for both the person with hearing loss and their family and friends, leading to reduced social interaction, participation, and quality of life [12]. Given that the typical age for first fitting of hearing aids is approximately 74 years and with many of these adults experiencing hearing difficulties for an average of 10 years prior to hearing aid fitting [12], untreated hearing loss is of particular importance for those aged 63-74 years. Our findings indicate that PC and Internet use was greater in this age group for those with early signs of hearing difficulties. This suggests that PCs and the Internet could be used to target specific hearing health care needs for this group.

Two potential health care approaches for the older group with slight hearing difficulties are delivery of effective information and hearing screening. Information and advice to educate these older adults about the effects of hearing loss and the benefits of hearing aids could be delivered in the form of short, easy-to-use video tutorials [9]. It is anticipated that this increased awareness will encourage at least some in this age group to seek appropriate interventions at an earlier age than is typical, with all the attendant benefits such as improved communication, participation, and quality of life [13]. This may be as simple as understanding the process of how to get a hearing test and what to expect afterwards [32]. Such information could be accessed easily either through a PC or streamed online through the Internet. In addition, levels of user interactivity with health care information can be substantially increased by delivery through PCs and the Internet, with users being able to revisit and review previously encountered material, which may lead to greater learning [33].

Remote hearing screening through the Internet is another highly relevant intervention for this cohort, and it has been shown to provide early benefits [12]. Screening enables the detection of those who are likely to benefit from hearing aid amplification and can help encourage individuals to attend audiology services for appropriate support. There are already good examples of screening initiatives delivered through the Internet and telephone. To date, the hearing check provided by Action on Hearing Loss has provided nearly one million hearing checks [3].

Finally, for those with mild hearing losses where hearing aids are unlikely to offer substantial benefit compared to those with greater levels of hearing loss, PCs and the Internet could be used to deliver alternative interventions such as auditory training to help alleviate the difficulties associated with hearing loss [17,18]. The main advantage of Internet delivery of auditory training is that accessibility to this intervention would be relatively high. A further advantage of online delivery of auditory training, which is not routinely available in the NHS, is that it would not necessarily require the user to access this through audiology services.

In the younger age group (50-62 years), in which PC and Internet use was highest, there was no difference in PC or Internet use between respondents with slight hearing difficulty and those with no hearing difficulty. We suggest that this is a result of there being a high level of PC and Internet use in this group already, which masks specific differences due to hearing difficulties. The relatively high skill set and Internet use in this age group suggests that online delivery of hearing health care is feasible for this younger age group with hearing loss. Appropriate online hearing health care for these adults within this age range may include online information and advice on the detrimental effects of untreated hearing loss, with a view to promote earlier awareness of hearing-related interventions including auditory training.

For respondents in the older age group (63-74 years), in which PC and Internet use was lowest, those with moderate+ hearing difficulties were equally likely to use PCs and the Internet as those with no hearing difficulties. This suggests that online hearing-related interventions may not be particularly effective if targeting this group as a whole. Subsequently, hearing health care tailored for those with moderate+ hearing losses, which is most likely to be hearing aid provision and information or advice relating to amplification, would be best supplemented with additional methods of information support (eg, printed materials or video tutorials delivered through DVD for those who do not or cannot access PC and Internet technology). Of course, for the 30% in the older age group who do use the Internet, information and advice would still be a valid option. The relatively low PC and Internet use in this group does not mean that online delivery of hearing health care will not be suitable for most of this cohort in the future. As PC and Internet use is becoming more prevalent in older adults over time, online delivery of hearing health care may provide a cost-effective, efficient method of providing hearing health care for older adults with a moderate+ hearing difficulties in future years [9].

Some limitations of the present study should be highlighted. First, our sample is limited to a small demographic (adults aged 50-74 years living in Nottingham, United Kingdom) because we were specifically interested in the pre-hearing aid user population. Nevertheless, our respondents have been shown to be representative of published data in terms of the prevalence of hearing loss and PC and Internet use. The World Health Organization lists adult-onset hearing loss as the most common cause of disability worldwide, with presbycusis (age-related hearing loss) the leading cause of adult-onset hearing loss [34]. Because eHealth offers the potential for hearing health care to be delivered globally, further studies may wish to assess the relationships among hearing difficulties and PC and Internet use in geographically remote and hard-to-reach populations, or those under the age of 50 years with significant hearing loss. Second, our postal questionnaire did not request details about our respondent's employment history. Those employed in manual professions are less likely to have used PCs or to use the Internet regularly at work [35]. This may have affected overall levels of PC skill and Internet use. Future studies may wish to control for employment as a potential confounding factor. Third, PC skill and Internet use were assessed subjectively in this study. Internet use was dichotomized as either “yes” or “no” without any information on frequency or proficiency of Internet use or their information delivery preferences because this was not the purpose of the postal questionnaire at that time. When considering online delivery of hearing health care for older adults, basic levels of PC skill and Internet proficiency are likely to be sufficient to allow access to online information and intervention [17]. However, further investigation of older adults' information technology skill levels, access, and information delivery preferences will help to inform specific delivery and content of online hearing health care for those older than 50 years.

In the present digital era, delivery of health care information and intervention through PCs and the Internet is common and the traditional method of clinical or medical health care delivery is supplemented increasingly by online information and support [3-10]. Advantages of PC and Internet delivery of supplementary information, hearing screening, and other interventions include the ability to reach those who do not or cannot present to an audiologist. This is of particular relevance given that approximately 47% of adults aged 55-74 years who visit their family practice physician about their hearing difficulties fail to be referred to an audiologist or hearing specialist [12]. Further advantages include both time and cost efficiency, with patients being able to access information at a time or place that suits them. Findings from this study suggest that delivery of hearing health care through the Internet can potentially target a substantial proportion of adults aged 50-74 years with age-related hearing loss, many of whom may not typically present to an audiologist. Therefore, PC and Internet delivery of hearing health care could help address the huge unmet need in those over the age of 50 years who have hearing loss, but do not currently have access to intervention or receive intervention.

Acknowledgments

We would like to thank the practice managers of the general practices who took part in this study for their assistance with the postal survey: Linda Lawton, practice manager, Leen View Surgery, Bulwell, Nottingham, United Kingdom; Jonathon Cummins, practice manager, Torkard Hill Medical Centre, Hucknall, Nottingham, United Kingdom; and Sayed Rashanmoniri, practice manager, Wollaton Vale Health Centre, Wollaton, Nottingham, United Kingdom.

This research was funded by the National Institute for Health Research (NIHR).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Postal questionnaire (hearing questions).

[PDF File (Adobe PDF File), 62KB - jmir_v14i4e113_app1.pdf]

Multimedia Appendix 2

Computer confidence questionnaire.

[PDF File (Adobe PDF File), 62KB - jmir_v14i4e113_app2.pdf]

References

1. Tu HT, Cohen GR. Striking jump in consumers seeking health care information. Track Rep 2008 Aug(20):1-8. [Medline: 18770913]
2. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
3. Action on Hearing Loss. 2011. Hearing screening test URL: http://www.actiononhearingloss.org.uk/your-hearing/look-after-your-hearing/check-your-hearing/take-the-check.aspx [WebCite Cache ID 62ihWsSyJ]
4. Krumm M, Ribera J, Klich R. Providing basic hearing tests using remote computing technology. J Telemed Telecare 2007;13(8):406-410. [doi: 10.1258/135763307783064395] [Medline: 18078552]
5. Smits C, Merkus P, Houtgast T. How we do it: The Dutch functional hearing-screening tests by telephone and internet. Clin Otolaryngol 2006 Oct;31(5):436-440. [doi: 10.1111/j.1749-4486.2006.01195.x] [Medline: 17014457]
6. Sweetow RW, Sabes JH. The Case for LACE: listening auditory communication enhancement training. Hear J 2004;57(3):32-35.
7. Laplante-Lévesque A, Pichora-Fuller MK, Gagné JP. Providing an internet-based audiological counselling programme to new hearing aid users: a qualitative study. Int J Audiol 2006 Dec;45(12):697-706. [doi: 10.1080/14992020600944408] [Medline: 17132558]
8. Thorén E, Svensson M, Törnqvist A, Andersson G, Carlbring P, Lunner T. Rehabilitative online education versus internet discussion group for hearing aid users: a randomized controlled trial. J Am Acad Audiol 2011 May;22(5):274-285. [doi: 10.3766/jaaa.22.5.4] [Medline: 21756843]
9. Ferguson MA, Leighton PA, Brandreth ME, Wharrand HJ. Development of evidence-based interactive videos for first-time hearing aid users. Int J Audiol 2012 (forthcoming).
10. HearCom. The HearCom project URL: http://hearcom.eu/main.html [WebCite Cache ID 63vLkgPoX]
11. Swanepoel de W, Hall JW. A systematic review of telehealth applications in audiology. Telemed J E Health 2010 Mar;16(2):181-200. [doi: 10.1089/tmj.2009.0111] [Medline: 20187743]
12. Davis A, Smith P, Ferguson M, Stephens D, Gianopoulos I. Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. Health Technol Assess 2007 Oct;11(42):1-294 [FREE Full text] [Medline: 17927921]
13. Davis A. Hearing in Adults. London: Whurr; 1995.
14. Action on Hearing Loss. Hearing Matters. 2011. URL: http://www.actiononhearingloss.org.uk/~/media/Documents/Research%2520and%2520policy/Hearing%2520matters/Hearing%2520matters_pdf.ashx [WebCite Cache ID 63vLci2Y5]
15. Neurotone. 2011. Listening and communication enhancement (LACE) URL: http://www.neurotone.com/ [WebCite Cache ID 63vKhFGPv]
16. Moore DR, Halliday LF, Amitay S. Use of auditory learning to manage listening problems in children. Philos Trans R Soc Lond B Biol Sci 2009 Feb 12;364(1515):409-420 [FREE Full text] [doi: 10.1098/rstb.2008.0187] [Medline: 18986969]
17. Ferguson MA, Henshaw H, Riley A, Clark DPA. Auditory training: does it work in people with hearing loss? Int J Audiol 2011;50(10):719.
18. Sweetow RW, Sabes JH. The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. J Am Acad Audiol 2006 Sep;17(8):538-558. [Medline: 16999250]
19. Cacciatore F, Napoli C, Abete P, Marciano E, Triassi M, Rengo F. Quality of life determinants and hearing function in an elderly population: Osservatorio Geriatrico Campano Study Group. Gerontology 1999;45(6):323-328. [Medline: 10559650]
20. Bamford J, Uus K, Davis A. Screening for hearing loss in childhood: issues, evidence and current approaches in the UK. J Med Screen 2005;12(3):119-124. [doi: 10.1258/0969141054855256] [Medline: 16156941]

21. Watkin PM, Baldwin M. Identifying deafness in early childhood: requirements after the newborn hearing screen. Arch Dis Child 2011 Jan;96(1):62-66. [doi: 10.1136/adc.2010.185819] [Medline: 21047829]
22. Ciorba A, Bianchini C, Pelucchi S, Pastore A. The impact of hearing loss on the quality of life of elderly adults. Clin Interv Aging 2012;7(6):159-163 [FREE Full text] [doi: 10.2147/CIA.S26059] [Medline: 22791988]
23. Davis A, Davis KA. Epidemiology of aging and hearing loss related to other chronic illnesses. Chicago: Phonak; 2009 Presented at: Hearing Care for Adults - The Challenge of Aging; 2010; Chicago.
24. Barak A, Sadovsky Y. Internet use and personal empowerment of hearing-impaired adolescents. Computers in Human Behaviour 2008;24:1802-1815.
25. Pilling D, Barrett P. Text communication preferences of deaf people in the United Kingdom. J Deaf Stud Deaf Educ 2008;13(1):92-103 [FREE Full text] [doi: 10.1093/deafed/enm034] [Medline: 17602163]
26. Seybert H, Loof A. Eurostat. 2010. Internet usage in 2010 – Households and Individuals URL: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-QA-10-050/EN/KS-QA-10-050-EN.PDF [accessed 2011-11-07] [WebCite Cache ID 631Lfmqjx]
27. Koppen J. American Association of Retired Pensioners (AARP). 2010. Social media and technology use among adults 50+ URL: http://assets.aarp.org/rgcenter/general/socmedia.pdf [accessed 2011-11-07] [WebCite Cache ID 631LYbvGJ]
28. Woodrow JEJ. The development of computer-related attitudes of secondary students. J Educat Comp Res 1994;11:307-338.
29. Levine T, Donitsa-Schmidt S. Computer Use, Confidence, Attitudes, and Knowledge: A Causal Analysis. Computers in Human Behavior 1998;41(1):125-146.
30. Roland M, Elliott M, Lyratzopoulos G, Barbiere J, Parker RA, Smith P, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ 2009;339:b3851 [FREE Full text] [Medline: 19808811]
31. Noble M, McLennan D, Wilkinson K, Whitworth A, Barnes H. Communities and Local Government. London: Department for Communities and Local Government; 2008. The English Indices of Deprivation 2007 URL: http://www.communities.gov.uk/documents/communities/pdf/733520.pdf [accessed 2011-12-14] [WebCite Cache ID 63vLQzrgC]
32. Smeeth L, Fletcher AE, Ng ES, Stirling S, Nunes M, Breeze E, et al. Reduced hearing, ownership, and use of hearing aids in elderly people in the UK--the MRC Trial of the Assessment and Management of Older People in the Community: a cross-sectional survey. Lancet 2002 Apr 27;359(9316):1466-1470. [Medline: 11988245]
33. Zhang D, Zhou L, Briggs RO, Nunamaker JF. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. Information & Management 2006;43:15-27.
34. World Health Organization. The global burden of disease: 2004 update. Geneva: WHO Press; 2008. URL: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf [accessed 2012-04-25] [WebCite Cache ID 67BmFDcrK]
35. Palmer KT, Griffin MJ, Syddall HE, Davis A, Pannett B, Coggon D. Occupational exposure to noise and the attributable burden of hearing difficulties in Great Britain. Occup Environ Med 2002 Sep;59(9):634-639 [FREE Full text] [Medline: 12205239]

Abbreviations

AIC: Akaike information criterion
ANOVA: analysis of variance
BEHD: better-ear hearing difficulty
CI: confidence interval
GP: general practitioner
HL: hearing level
IMD: Index of Multiple Deprivation
LRT: likelihood ratio test
NHS: National Health Service
NIHR: National Institute for Health Research
OR: odds ratio
PC: personal computer
SES: socioeconomic status
SMS: short message service (text message for mobile phones)
SNHL: sensorineural hearing loss
SSE: error sum of squares

Edited by G Eysenbach; submitted 22.12.11; peer-reviewed by C Quinn, W Anderson; comments to author 30.01.12; revised version received 30.04.12; accepted 24.05.12; published 24.08.12.

Please cite as:
Henshaw H, Clark DPA, Kang S, Ferguson MA
Computer Skills and Internet Use in Adults Aged 50-74 Years: Influence of Hearing Difficulties
J Med Internet Res 2012;14(4):e113
URL: http://www.jmir.org/2012/4/e113/
doi:10.2196/jmir.2036
PMID:22954484

Original Paper

Speech Perception Benefits of Internet Versus Conventional Telephony for Hearing-Impaired Individuals

Georgios Mantokoudis[1*], MD; Patrick Dubach[1], MD; Flurin Pfiffner[1], PhD; Martin Kompis[1], MD, PhD; Marco Caversaccio[1], MD; Pascal Senn[1*], MD

Cochlear Implant Division, Department of Otorhinolaryngology, Head & Neck Surgery, Inselspital, University of Bern, Bern, Switzerland

[*]these authors contributed equally

Corresponding Author:
Pascal Senn, MD
Cochlear Implant Division
Department of Otorhinolaryngology, Head & Neck Surgery, Inselspital
University of Bern
Inselspital
Freiburgstrasse
Bern, 3010
Switzerland
Phone: 41 31 632 3347
Fax: 41 31 632 3416
Email: pascal.senn@insel.ch

Abstract

Background: Telephone communication is a challenge for many hearing-impaired individuals. One important technical reason for this difficulty is the restricted frequency range (0.3–3.4 kHz) of conventional landline telephones. Internet telephony (voice over Internet protocol [VoIP]) is transmitted with a larger frequency range (0.1–8 kHz) and therefore includes more frequencies relevant to speech perception. According to a recently published, laboratory-based study, the theoretical advantage of ideal VoIP conditions over conventional telephone quality has translated into improved speech perception by hearing-impaired individuals. However, the speech perception benefits of nonideal VoIP network conditions, which may occur in daily life, have not been explored. VoIP use cannot be recommended to hearing-impaired individuals before its potential under more realistic conditions has been examined.

Objective: To compare realistic VoIP network conditions, under which digital data packets may be lost, with ideal conventional telephone quality with respect to their impact on speech perception by hearing-impaired individuals.

Methods: We assessed speech perception using standardized test material presented under simulated VoIP conditions with increasing digital data packet loss (from 0% to 20%) and compared with simulated ideal conventional telephone quality. We monaurally tested 10 adult users of cochlear implants, 10 adult users of hearing aids, and 10 normal-hearing adults in the free sound field, both in quiet and with background noise.

Results: Across all participant groups, mean speech perception scores using VoIP with 0%, 5%, and 10% packet loss were 15.2% (range 0%–53%), 10.6% (4%–46%), and 8.8% (7%–33%) higher, respectively, than with ideal conventional telephone quality. Speech perception did not differ between VoIP with 20% packet loss and conventional telephone quality. The maximum benefits were observed under ideal VoIP conditions without packet loss and were 36% (*P* = .001) for cochlear implant users, 18% (*P* = .002) for hearing aid users, and 53% (*P* = .001) for normal-hearing adults. With a packet loss of 10%, the maximum benefits were 30% (*P* = .002) for cochlear implant users, 6% (*P* = .38) for hearing aid users, and 33% (*P* = .002) for normal-hearing adults.

Conclusions: VoIP offers a speech perception benefit over conventional telephone quality, even when mild or moderate packet loss scenarios are created in the laboratory. VoIP, therefore, has the potential to significantly improve telecommunication abilities for the large community of hearing-impaired individuals.

(J Med Internet Res 2012;14(4):e102) doi:10.2196/jmir.1818

KEYWORDS

VoIP; Internet telephony; hearing impaired; communication

Introduction

Engaging in telephone conversations is a challenge for many hearing-impaired individuals, including hearing aid and cochlear implant users [1-4]. Reduced social connectivity likely mediates the well-known associations between hearing loss and depression, cognitive decline, reduced quality of life, and increased morbidity and mortality [5-8]. Any improvement in telecommunication would therefore be of great importance and may affect millions of people, as hearing loss is a common disease in modern societies: it affects 1 to 4 of every 1000 people born [9], approximately 16% of adults aged 20 to 69 years [10], and more than 80% of older adults aged 80 to 92 years [11].

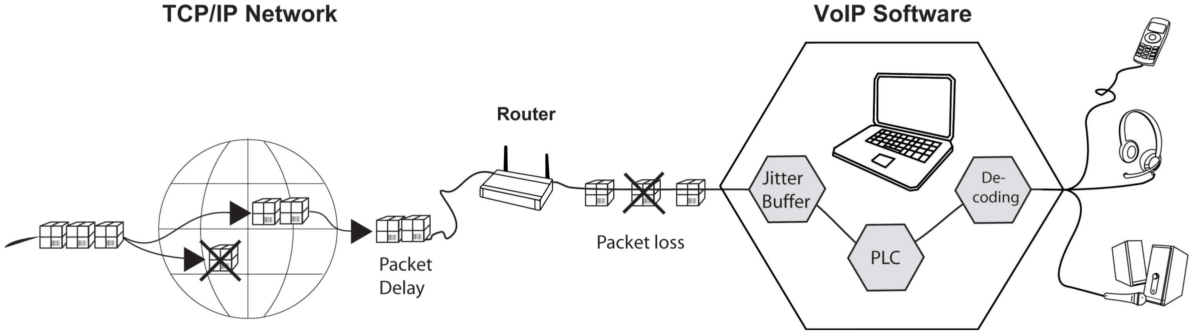
The two main technical limitations of conventional telephones for hearing-impaired individuals are, first, the frequency restriction and second, the digital data compression used in conventional telephony to maximize the network infrastructure utilization.

In addition, issues related to coupling the telephone to a hearing aid or a cochlear implant may further reduce speech perception by the hearing-impaired end user, even when assistive telephone listening devices are used [12-14]. Finally, hearing-impaired individuals face additional intelligibility problems when the caller or receiver is surrounded by environmental noise.

From a theoretical perspective, Internet telephony (voice over Internet protocol [VoIP]) should offer improved speech perception to the end user, as the transmitted frequency range is double that of conventional telephones (0.1–8 kHz vs 0.3–3.4 kHz). The association between improved speech perception and the presentation of speech at higher bandwidths has been repeatedly shown in the literature [15-17]. Additionally, modernized versions of signal processing strategies for coding and decoding speech, so-called codecs, have recently been implemented into VoIP software with the aim of improving intelligibility, particularly under adverse network conditions.

These technical advantages of Internet telephony translate into improved speech perception by hearing-impaired and normal-hearing adults, at least when simulated under ideal laboratory conditions [18]. However, ideal network conditions are not constantly present in the World Wide Web, especially when Internet lines are overloaded [19]. Under adverse conditions, digital data packets may be delayed or lost. Lost data packets are usually described as a percentage of the original number of data packets sent through the network [20]. Despite software solutions designed to recover lost data packets (known as packet loss concealment; Figure 1), sound quality and speech perception may be negatively affected. For hearing-impaired individuals, even small differences in audio quality may have an important impact on speech perception. However, to date, no reports have addressed the relationship between packet loss in Internet telephony and speech perception by hearing-impaired individuals. To fill this gap, we assessed speech perception using standardized speech test material presented at a variety of simulated adverse VoIP qualities and compared it with speech perception under ideal conventional telephony conditions.

Figure 1. Each data packet originating in a voice over Internet protocol (VoIP)-sending device (not shown) takes a different route through the Internet (TCP/IP network) before arriving at the receiver. Data packets may be delayed or lost on the way. The VoIP software includes two solutions to improve audio quality in these cases: the jitter buffer collects the maximum number of data packets by waiting as long as needed and keeping the time delay to a minimum; and packet loss concealment (PLC) aims to reconstruct lost data packets. Finally, digital data packets are decoded and delivered to a VoIP-compatible user interface, such as a VoIP handheld telephone, a headset, or external loudspeakers.



Methods

We conducted all tests between January and June 2009 at the University Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital, Bern, Switzerland. The study protocol was approved by the local institutional review board. All patients gave written informed consent.

Test Participants

Participants in the study were 20 hearing-impaired adults, consisting of 10 users of cochlear implants and 10 users of hearing aids, and 10 normal-hearing adults. All test participants were at least 18 years old and were selected from the institution’s clinical database. Mean age was 46 years in the cochlear implant group, 68 years in the hearing-impaired group, and 35 years in the normal-hearing group. A total of 90% of participants who were tested in our previous experimental study [18] consented to participate in the present study and came to the laboratory for a second series of independent experiments. The same inclusion criteria used in the former study were applied; these were based on pure-tone audiograms, speech audiometry performance (the German Freiburger monosyllable test), and technical device specifications. Eligible cochlear implant users had processors that allowed stimulation beyond 5 kHz and aided minimum monosyllabic word discrimination scores of 50% or more at a 60-dB sound pressure level (Table 1). Eligible hearing aid users had moderate bilateral hearing loss with unaided sloping pure-tone audiometry thresholds, aided pure-tone audiometry with hearing gains beyond 3.4 kHz, and aided minimum monosyllabic word discrimination scores of 15% and 70% at sound pressure levels of 60 and 75 dB, respectively. Table 1 shows the aided speech discrimination scores for each individual and Figure 2 shows the aided hearing thresholds in free field, which reflects the degree of compensation for their hearing impairments and their ability to hear the telephone speech signal. The pure-tone audiometry thresholds and monosyllabic word discrimination scores of the normal-hearing control participants were within normal limits (100%).

Table 1. Clinical data for cochlear implant users and hearing aid users.

Participant	Sex	Age (years)	Hearing loss etiology	Device brand and model	Aided German monosyllable word discrimination score (%)			
					Ear	60 dB	75 dB	80 dB
Cochlear implant users								
1	M[a]	77	Progressive	MED-EL Pulsar/Opus 2	Left	77.5	NA[b]	97.5
2	M	17	Postmeningitic	MED-EL Pulsar/Opus 2	Left	77.5	NA	87.5
3	F[c]	39	Congenital	MED-EL Pulsar/Opus 2	Right	62.5	NA	72.5
4	F	69	Progressive	MED-EL C40+ Tempo+	Left	72.5	NA	85
5	F	48	SHL[d]	MED-EL C40+ Tempo+	Right	77.5	NA	77.5
6	F	61	Progressive	MED-EL Pulsar/Opus 2	Left	50	NA	65
7	F	22	SHL	MED-EL C40+ Tempo+	Right	55	NA	75
8	M	50	Congenital	MED-EL C40C Tempo+	Right	85	NA	80
9	M	58	Meningitis	MED-EL C40C Tempo+	Left	70	NA	65
10	F	23	Progressive	MED-EL C40C Tempo+	Right	55	NA	52.5
Hearing aid users								
1	F	66	Presbycusis	BTE/Oticon Tego Pro VC	Left	100	100	NA
2	M	77	Presbycusis	ITE/Bernafon Symbio XT	Left	100	100	NA
3	M	86	Presbycusis	BTE/Phonak Piconet 2	Left	90	95	NA
4	F	79	Presbycusis	ITE/Bernafon Neo 315	Left	75	95	NA
5	M	91	Presbycusis	BTE/Widex Inteo	Left	25	85	NA
6	M	62	Presbycusis	BTE/Phonak Una M AZ	Right	30	80	NA
7	M	76	Progressive	BTE/Phonak Extra	Left	70	85	NA
8	F	36	Congenital	BTE/GN ReSound Air	Right	100	100	NA
9	M	63	Progressive	BTE/Phonak micro eXtra	Left	90	100	NA
10	M	41	Progressive	BTE/Phonak Audéo	Right	70	95	NA

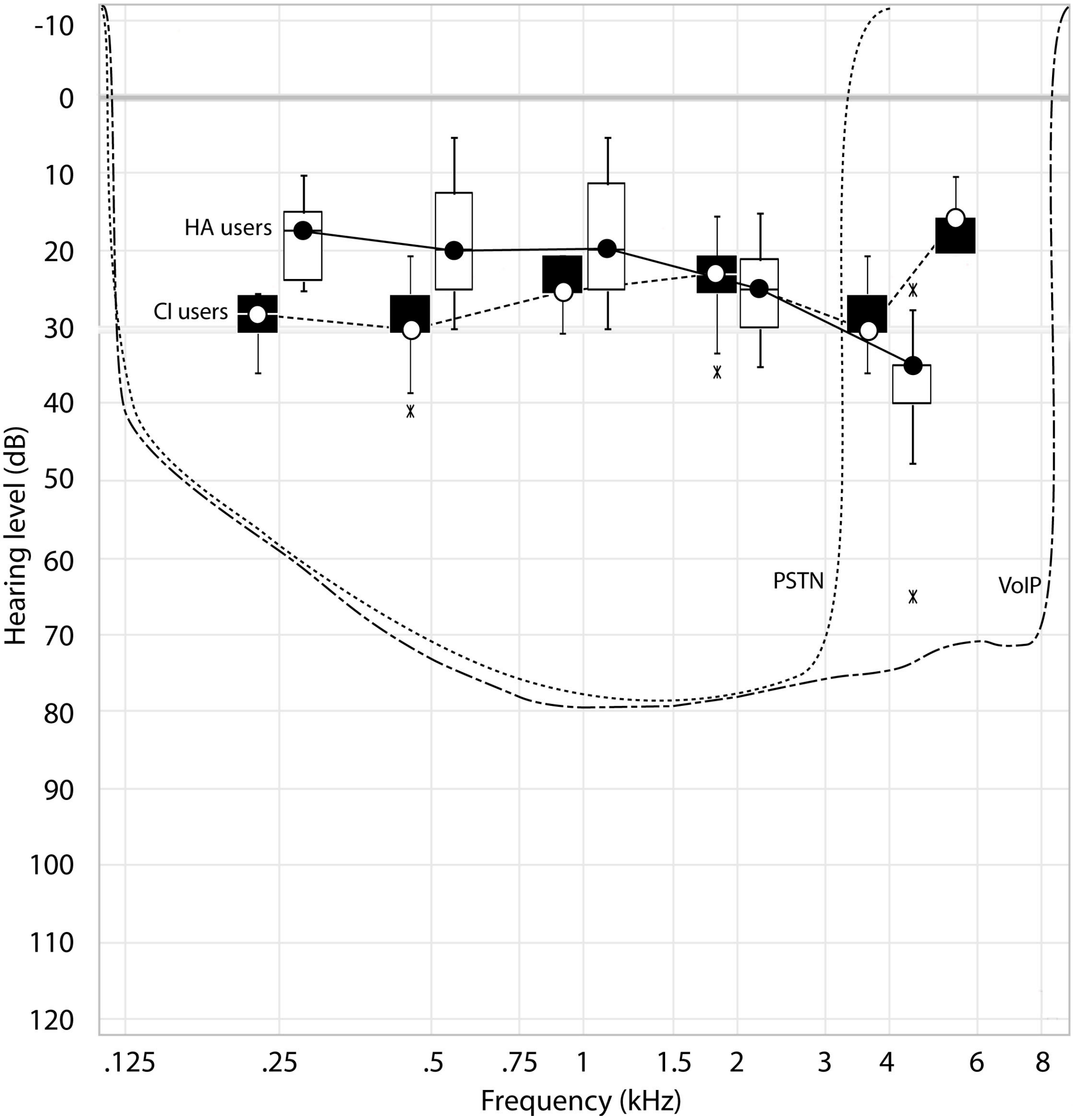
[a] Male.

[b] Not applicable.

[c] Female.

[d] Sudden hearing loss.

Figure 2. Lower quartile, median, upper quartile, and 1.5*interquartile range (X = outliers) of aided hearing thresholds in the free sound field for cochlear implant (CI) and hearing aid (HA) users. The analog telephone (public switched telephone network [PSTN]) speech signals are shown in dB hearing level as dotted lines. VoIP = voice over Internet protocol.



Speech Perception Test Protocol

All tests were performed in the free sound field in a sound-treated room. The speech test material was played on an audio compact disc (CD) player connected to an audiometer (GSI 61; Grason-Stadler, Milford, NH, USA) and was reproduced using a pair of loudspeakers (Type 1030 A; Genelec Oy, Iisalmi, Finland) situated 1 m from the front of the patient's head.

All tests were conducted monaurally. The most suitable ear was selected based on the inclusion criteria; if both ears were equally suitable, the ear commonly used for telephony was used. The opposite ear canal was occluded with an earplug (E.A.R. Classic, Aearo Technologies, Stockport, UK). The specified average attenuation of these earplugs is 24.6–41.6 dB in the 250- to 4000-Hz range. Bilateral cochlear implant users had to switch off one of their devices to produce homogeneous and comparable data. Monaural testing was necessary because not all individuals had the same degree of hearing loss on both sides, and speech perception performance varied between the first and second listening device. In addition, monaural testing more realistically reflects the use of a conventional telephone handset.

We used the standardized German Hochmair-Schulz-Moser (HSM) Sentence Test [21] in quiet and competing noise to test speech perception. This test consists of 30 lists with 20 sentences and 106 words per list to avoid learning effects. The sentences were presented in 4 different VoIP qualities (with 0%, 5%, 10%, and 20% packet loss) and ideal conventional telephone quality (Table 2) in quiet and with competing noise with 4 different signal to noise ratios (SNRs): 15, 10, 5, and 0 dB. We set the defined broadband noise signal to a constant level of 70 dB sound pressure level and changed the speech signal according to the SNR required by the standardized test protocol (range 70–85 dB sound pressure level). The audibility of the speech signal in relation to the aided hearing threshold is shown in Figure 2 for VoIP and analog telephone signals. Participants had to orally repeat the presented sentences, and the percentage of correctly understood words was assessed and used for comparison across conditions. A random permutation of all individuals was performed in conjunction with test order randomization to avoid order effects and selection bias. The participants were unaware of which condition was being tested at any time of the assessment.

Table 2. Frequency and filter characteristics for each audio quality

Codec	Frequency range (kHz)	Sampling rate	Fp (Hz)[a]	Fs (Hz)[b]	Ap (dB)[c]	As (dB)[d]
PSTN[e] codec G.711	0.1–3.4	8 kHz G.711 A-Law	3900	4400	1	60
iPCMwb codec	0.1–8	16 kHz PCM[f]	8000	8500	1	60
CD low-pass filtered	0.1–8	16 kHz PCM	8000	8500	1	60

[a] Frequency at the edge of the pass band.

[b] Frequency at the beginning of the stop band.

[c] Amount of ripple allowed in the pass-band (also called Apass).

[d] Stop-band attenuation.

[e] Public switched telephone network.

[f] Pulse code modulation.

Digital Generation of VoIP Audio Signals

The original audio CD files of the HSM Sentence Test were converted into a wave-format audio file using the Switch Audio File Converter software, version 1.05 (NCH Software Pty Ltd, Canberra, ACT, Australia). The speech and noise channels were mixed to mono wave files with a sampling rate of 44.1 kHz, thereby allowing identical signal processing for speech and noise. Before encoding, the audio files were low-pass filtered using MATLAB software, version 7.9.0.529 (The MathWorks, Inc, Natick, MA, USA).

The audio files were then converted again into raw files using Switch Audio File Converter software. To generate a VoIP simulation with different extents of packet loss, the raw data were encoded using simulation software (a voice engine demonstration application) in conjunction with a modern iPCMwb codec (0.1–8 kHz; Global IP Solutions, San Francisco, CA, USA). Table 2 shows the filter parameters. The public switched telephone network (PSTN) transmission has an upper limit of 3.4 kHz. In contrast to this, the frequency transmission of VoIP extends to 8 kHz, thereby preserving the high-frequency content of speech.

Table 3. Parameters of the voice over Internet protocol (VoIP) simulations.

Condition	Description (packet loss)	Loss rate (p)	Loss length (packets)	BurstR	Frame length (bytes)	q
0	Perfect	0.0	1.0			
1	Mild loss (5%)	0.05	1.15	1.1	640	0.87
2	Medium loss (10%)	0.10	1.30	1.2	640	0.77
3	Severe loss (20%)	0.20	2.0	1.43	640	0.5

We simulated 4 different scenarios: 1 scenario without packet loss and 3 scenarios with packet losses of 5%, 10%, and 20%. In Table 3, the parameter p represents the probability of packet loss. The average loss length defines the number of lost packets. The parameter BurstR is a measure of packet loss "burstiness," as defined by Raake [22]. BurstR = 1.0 indicates that there was no correlation across packet losses, which were all independent and identically distributed with probability p (obtained by omitting the loss length parameter from the command line). BurstR >1.0 means that the packet losses tended to come in bursts; the larger the BurstR, the longer the bursts of losses. The variable BurstR is dimensionless. The parameter q is simply a translation of loss length into the transition probability from a *lost* to a *found* state of the underlying Markov model: q = 1 / loss length [23,24]. Finally, the encoded data sizes for a frame, the frame length or payload, are indicated in bytes.

Digital Generation of PSTN Audio Signals

To simulate conventional telephone audio quality, we implemented a PSTN G.711 A-Law codec, which is a standard in PSTNs, in the Switch Audio File Converter software. We coded the files at a sampling rate of 8 kHz. Before encoding, we used MATLAB software to limit the upper frequency range of the original wave files to 4 kHz.

All 5 audio CDs (4 VoIP simulations and 1 PSTN quality simulation) were reproduced using an active loudspeaker system (Genelec Type 1030 A). They were calibrated in the free sound field using a Type 2636 measuring amplifier and a Type 4133 FF measuring microphone connected to a Type 2619 preamplifier (all from Brüel & Kjær Sound & Vibration Measurement A/S, Nærum, Denmark). We used a Pistonphone Brüel & Kjær 4288 precision calibrator to calibrate the measurement arrangement. The final measurements showed no difference in the sound pressure levels of speech signals across different audio signals.

Statistical Analysis

W used a 2-tailed Wilcoxon matched-pairs signed rank test to compare the scores obtained under different VoIP versus telephone quality simulations. For the condition with no packet loss (condition 0, Table 3), a 1-tailed test was applied because of the expected superiority of VoIP under this condition [18,25]. $P < .05$ was considered significant. When applying a Bonferroni correction for multiple testing, P_{Bonf} £ .0125 was considered significant.

Results

Across all test groups, speech perception scores assessed with different VoIP qualities versus conventional telephone quality were higher in 39 out of 60 test conditions (Figure 3 and Table 4). The average advantage of VoIP in the 39 conditions was 14.6% (range 1%–53%), and the differences were statistically significant in 23 conditions ($P < .05$). Mean speech perception scores were 15.2% (range 0%–53%) higher in the sentence test using VoIP with no packet loss, 10.6% (4%–46%) with mild packet loss, and 8.8% (7%–33%) with medium packet loss across the 3 groups (Table 4). Scores obtained under VoIP conditions with severe packet loss were similar to those obtained with conventional telephone quality.

Cochlear implant users showed improved speech perception using VoIP in 19 out of 20 test conditions (Figure 3, part A). On average, speech perception scores were 15.3% (range 1%–36%) higher with VoIP. The differences reached statistical significance in 13 conditions (Table 4). In only 1 condition with severe packet loss did we find a disadvantage for VoIP versus conventional telephone quality; this 1% difference was not statistically significant.

Hearing aid users had improved speech perception scores with different VoIP qualities in half of the test conditions (Figure 3, part C). The mean advantage was 7.1% (range 2%–18%); this advantage was statistically significant in 3 test conditions with no packet loss (Table 4). In 9 other test conditions with increased packet loss, speech perception scores were on average 6.2% (1%–17%) lower with VoIP. The negative differences were statistically significant under 4 conditions with severe packet loss (Table 4).

Normal-hearing adults showed an average benefit of 20.8% (range 1%–53%) with different VoIP qualities under half of the conditions (Figure 3, part E). Under 7 conditions, the differences were statistically significant. Under 6 conditions with medium and severe packet loss, an average disadvantage of 2.2% (1%–6%) was found for VoIP; one of these conditions showed a statistically significant disadvantage (Table 4).

When we experimentally increased VoIP packet loss from 0% to 10%, speech perception scores dropped only mildly (Figure 3, parts B, D, and F). When the packet loss increased to 20%, speech perception scores dropped more sharply. The scores achieved with VoIP qualities of 5% to 10% packet loss were typically still better than those obtained with simulations of ideal conventional telephone quality (represented by black triangles in Figure 3, parts B, D, and F).

Table 4. Mean differences in speech perception scores (D%) assessed with different voice over Internet protocol qualities (degree of packet loss) versus conventional telephone quality using the Hochmair-Schulz-Moser (HSM) Sentence Test.

Participant group	SNR (dB)[a]	Packet loss None (0%) D%	P value	Mild (5%) D%	P value	Medium (10%) D%	P value	Severe (20%) D%	P value
CI[b] users	0	+2	NA[c]	+1	NA	+2	NA	+2	NA
	5	+21	.002[d]	+11	.01[d]	+9	.008[d]	+1	.81
	10	+36	.001[d]	+34	.002[d]	+30	.002[d]	+4	.04[e]
	15	+36	.001[d]	+23	.01[e]	+25	.006[d]	+6	.25
	Quiet	+16	.001[d]	+16	.002[d]	+16	.002[d]	–1	.92
HA[f] users	0	+4	.16	+2	.69	+0	>.99	–1	.02[e]
	5	+18	.002[d]	+4	.56	+6	.38	–6	.16
	10	+16	.003[d]	+9	.08	+4	.32	–12	.02[e]
	15	+6	.004[d]	–4	.65	–7	.25	–17	.004[d]
	Quiet	+2	.16	–1	.31	–1	.56	–7	.004[d]
NHA[g]	0	+53	.001[d]	+46	.002[d]	+33	.002[d]	+22	.002[d]
	5	+17	.003[d]	+16	.002[d]	+16	.004[d]	–3	.49
	10	+1	.21	+3	.06	+1	.47	–6	.009[d]
	15	0	NA	0	NA	–1	NA	–1	>.99
	Quiet	0	NA	0	NA	–1	NA	–1	NA

[a] Signal to noise ratio.

[b] Cochlear implant.

[c] Not applicable.

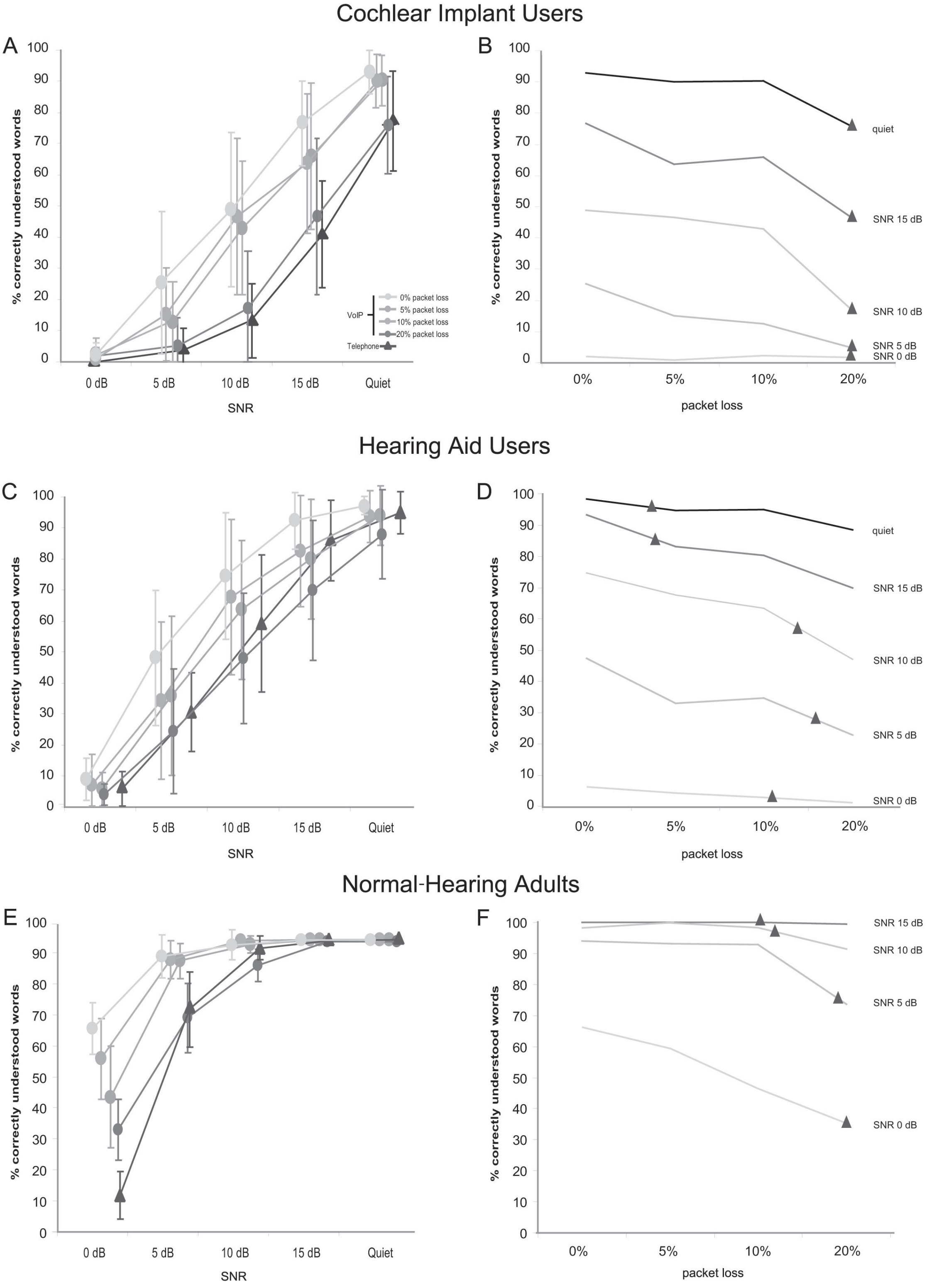
[d] Statistically significant with Bonferroni correction.

[e] Statistically significant without Bonferroni correction.

[f] Hearing aid.

[g] Normal-hearing adults.

Figure 3. Speech perception scores assessed with the HSM Sentence Test are plotted against different signal to noise ratios (SNRs) for cochlear implant users (A), hearing aid users (C), and normal-hearing adults (E) for 4 different VoIP qualities (0%, 5%, 10%, and 20% packet loss) and 1 ideal conventional telephone quality. The impact of different network conditions with increasing data packet loss (x-axis) on word discrimination scores is shown for different SNRs in B, D, and F. The black triangle indicates the speech perception level corresponding to a conventional telephone with a constant and stable transmission. VoIP = voice over Internet protocol.



Discussion

Key Findings

The present study confirmed that simulations of Internet versus conventional telephony quality are associated with improved speech perception by hearing-impaired and normal-hearing adults under perfect network conditions without packet loss or delay in the laboratory. The advantage for cochlear implant users and normal-hearing adults persists, even when the VoIP quality is reduced by 5% and 10% packet loss. Similarly, hearing aid users also scored higher with VoIP than with conventional telephony, but the differences did not reach statistical significance.

In general, speech perception scores assessed with VoIP quality remained good until a packet loss of 10%. Interestingly, VoIP simulation under severely adverse network conditions (with a packet loss of 20%) was not inferior to a perfect conventional telephone simulation in the majority of test conditions and for most of the tested participants.

Impact of Frequency Range, Coding Strategy, and SNR on Speech Perception

Our research group earlier showed the superiority of VoIP versus conventional telephony simulations under perfect network conditions [18]. The main reason for this advantage is the enlarged frequency range offered by VoIP (0.1–8 kHz vs 0.3–3.4 kHz). The association of improved speech perception with presentation of speech at higher bandwidths has been repeatedly shown [15-17].

Cochlear implant users had the lowest speech perception scores of all 3 tested groups in our study (Figure 3), which is explainable because they have the highest degree of hearing loss. Interestingly, cochlear implant users reached the highest gain when tested with VoIP versus conventional telephony. The main reason for this is probably technical. Cochlear implant devices transmit a broad frequency range (on average 0.2–7.9 kHz, as verified prior to testing; Figure 2). Therefore, cochlear implant users can fully exploit the enlarged frequency range offered by VoIP. In contrast, most hearing aids used by our test participants do not transmit up to 8 kHz (on average 0.2–6.7 kHz according to the manufacturer's specifications). Hearing aid users can therefore not exploit the full frequency range offered by VoIP. The enlarged frequency range is audible to hearing aid users with functional thresholds below 40 dB at 4 and 6 kHz, respectively (Figure 2). They still benefit from the enlarged frequency range, but to a lesser extent.

The second reason for the advantage of VoIP is the conservation of high audio quality through digital signal processing using the chosen iPCMwb codec. In the first study by our group, speech perception scores that were assessed with VoIP quality were equal to scores obtained with frequency-restricted (0.1–8 kHz), uncompressed audio CD quality [18]. This means that through compression of digital data no relevant information for speech perception is lost. Additionally, modern VoIP codecs offer a constant full frequency range transmission even under adverse network conditions. The bit rate of the chosen iPCM-wb codec was variable (minimum at 36 kbit/s for silence and low levels) at an average rate of 80 kb/s and constant sampling rate of 16 kHz, which ensured a high-quality audio performance over heavily loaded packet networks.

Speech perception is more challenging with increasing competing noise, or decreasing SNRs. In particular, elderly hearing-impaired individuals with predominant high-frequency hearing loss suffer in noisy test conditions; in addition to complication from competing noise, the high-frequency content of speech is missing. VoIP may be helpful, because it transmits the high-frequency content of speech and because it offers the possibility of presenting the speech signal simultaneously to both ears through external loudspeakers, thereby allowing binaural hearing, which is a well-known advantage for speech perception in noise. Additionally, wired or wireless links enabling binaural hearing from 1 telephone signal are already available for hearing aids and cochlear implants.

Packet Loss

The measurement of packet loss under real VoIP transmission is a challenge for many VoIP companies because there is no constant data transmission over the Internet [26]. Packet loss of speech data may have a significant impact on speech audio quality [27]. The network transmission of voice data packets depends on the Internet infrastructure and transmission capacity, both of which may vary across companies and countries [28]. The transmission capacity may be reduced when Internet lines are overcharged. This situation can occur during rush hours, for example.

A decade ago, when VoIP telephony was not so highly developed, the average packet loss for a large number of measurement traces has been reported to be below 8% ($p < 0.08$ and $q > 0.8$) [24]. Since then, the Internet infrastructure, VoIP-compatible devices, and VoIP software solutions have been drastically improved. In 2004, one author already postulated that the packet loss should be held lower than 1% to ensure excellent VoIP transmission [29]. Nowadays, this request seems to be met, at least for most of the highly industrialized countries.

It can therefore be assumed that telecommunication using VoIP should substantially improve speech perception compared with conventional telephony under real network conditions, since the benefit of VoIP was measurable for most of the test participants up to an experimental packet loss of 10% in our study. A packet loss of more than 10% has a significant impact on sound quality with tone bursts, interruptions, extended time delay, and jitter of the audio signal. This is shown in Figure 3, parts B, E, and H, with degradation of speech intelligibility beyond 10% packet loss.

The calculations and models of packet loss depend on the measuring method used [30-35]. However, passive or active real-time packet loss monitoring and measurements are still challenging for many researchers and network engineers.

To our knowledge, no other group has assessed the speech perception of hearing-impaired individuals using Internet telephony under adverse network conditions. The results of the present study therefore fill an important gap. Measuring packet loss under controlled laboratory conditions offers the opportunity to systematically address a highly variable phenomenon in the real network.

Jitter Buffer, Time Delay, and Packet Loss Concealment

Many technical parameters that may further influence speech perception by the end user have not been addressed in the present study. Every conversation can be disturbed when data packets arrive late to the receiver [28]. A jitter buffer is a part of the software solution for this problem. It collects all relevant voice data packets by waiting as long as needed and minimizing the time delay (Figure 1). Packet losses can also be induced locally by setting a low playout delay in the jitter buffer. This may lead to bursts of packet loss and increased jitter, with further degradation of audio speech quality. Additionally, time delays and echo may affect speech perception; however, we did not test these factors. Nonetheless, experiences with VoIP and broadband access networks have shown only small packet delay variations with minimal delay and jitter [29]. Similarly, packet loss concealment algorithms (Figure 1) in the VoIP software reconstruct speech information when data packets are lost [28,29]. Because speech perception may vary with the use of different packet loss concealment settings and codecs, we cannot generalize our results for all VoIP software available on the market. VoIP audio quality may further be influenced by wireless network conditions [34].

Practical Usefulness

Our test results may be important for hearing-impaired individuals, including hearing aid and cochlear implant users, because there is now strong experimental evidence for real improvement in speech perception when using VoIP instead of conventional telephones. The study is also important for physicians, audiologists, cochlear implant engineers, speech therapists, and other professionals who care for hearing-impaired individuals. Professionals should encourage hearing-impaired individuals to try VoIP, which is typically downloadable at no cost from most providers. Patients who already own a computer may be able to gain the benefits of VoIP at no cost. The use of external loudspeakers connected to the computer may further improve speech perception by permitting bilateral hearing and additional amplification through the volume control, which should be mentioned to the patients. Hearing aid and cochlear implant accessories, such as an FM transmitter and 3.5-mm audio jack, may also be helpful for coupling the computer directly to the hearing device. Patients should be advised that both sender and receiver should have a good microphone and loudspeaker system to take advantage of VoIP's broadband advantage over conventional telephony.

Conclusions

Speech perception by hearing-impaired individuals and normal-hearing adults is improved when using perfect VoIP versus perfect conventional telephony transmission under controlled laboratory conditions. The superiority of VoIP persists even under experimental adverse network conditions, but not to the same extent and not for all tested individuals. Cochlear implant users seem to benefit more than hearing aid users because their devices are better suited to exploit VoIP's broadband frequency range.

Acknowledgments

The authors would like to thank Christoph Schmid, engineer, for technical support and Aisha Spring for editing the figures. The authors would like to thank Global IP Solutions (GIPS, San Francisco, CA, USA) for providing their simulation software and the iPCMwb codec. We are very grateful to Michael Poppler and Tina le Grand, GIPS engineers, for their suggestions regarding the manuscript.

Financial support for this study was obtained from the foundation Stiftung für die Erforschung von Hör- und Sprachstörungen, Bern.

Conflicts of Interest

None declared.

References

1. Holmes AE, Frank T. Telephone listening ability for hearing-impaired individuals. Ear Hear 1984;5(2):96-100. [Medline: 6724177]
2. Kepler LJ, Terry M, Sweetman RH. Telephone usage in the hearing-impaired population. Ear Hear 1992 Oct;13(5):311-319. [Medline: 1487091]
3. Adams JS, Hasenstab MS, Pippin GW, Sismanis A. Telephone use and understanding in patients with cochlear implants. Ear Nose Throat J 2004 Feb;83(2):96, 99-100, 102. [Medline: 15008443]
4. Cray JW, Allen RL, Stuart A, Hudson S, Layman E, Givens GD. An investigation of telephone use among cochlear implant recipients. Am J Audiol 2004 Dec;13(2):200-212. [Medline: 15903146]
5. Barnett S, Franks P. Deafness and mortality: analyses of linked data from the National Health Interview Survey and National Death Index. Public Health Rep 1999;114(4):330-336. [Medline: 10501133]
6. Cacciatore F, Napoli C, Abete P, Marciano E, Triassi M, Rengo F. Quality of life determinants and hearing function in an elderly population: Osservatorio Geriatrico Campano Study Group. Gerontology 1999;45(6):323-328. [Medline: 10559650]

7. Dalton DS, Cruickshanks KJ, Klein BE, Klein R, Wiley TL, Nondahl DM. The impact of hearing loss on quality of life in older adults. Gerontologist 2003 Oct;43(5):661-668. [Medline: 14570962]
8. Gates GA, Cobb JL, Linn RT, Rees T, Wolf PA, D'Agostino RB. Central auditory dysfunction, cognitive dysfunction, and dementia in older people. Arch Otolaryngol Head Neck Surg 1996 Feb;122(2):161-167. [Medline: 8630210]
9. Connolly JL, Carron JD, Roark SD. Universal newborn hearing screening: are we achieving the Joint Committee on Infant Hearing (JCIH) objectives? Laryngoscope 2005 Feb;115(2):232-236. [doi: 10.1097/01.mlg.0000154724.00787.49] [Medline: 15689741]
10. Agrawal Y, Platz EA, Niparko JK. Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the National Health and Nutrition Examination Survey, 1999-2004. Arch Intern Med 2008 Jul 28;168(14):1522-1530. [doi: 10.1001/archinte.168.14.1522] [Medline: 18663164]
11. Cruickshanks KJ, Tweed TS, Wiley TL, Klein BE, Klein R, Chappell R, et al. The 5-year incidence and progression of hearing loss: the epidemiology of hearing loss study. Arch Otolaryngol Head Neck Surg 2003 Oct;129(10):1041-1046. [doi: 10.1001/archotol.129.10.1041] [Medline: 14568784]
12. Latzel M, Gebhart TM, Kiessling J. Benefit of a digital feedback suppression system for acoustical telephone communication. Scand Audiol Suppl 2001(52):69-72. [Medline: 11318487]
13. Lowe RG, Goldstein DP. Acoustic versus inductive coupling of hearing aids to telephones. Ear Hear 1982;3(4):227-234. [Medline: 6214443]
14. Plyler PN, Burchfield SB, Thelin JW. Telephone communication with in-the-ear hearing aids using acoustic and electromagnetic coupling. J Am Acad Audiol 1998 Dec;9(6):434-443. [Medline: 9865776]
15. Baer T, Moore BC, Kluk K. Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. J Acoust Soc Am 2002 Sep;112(3 Pt 1):1133-1144. [Medline: 12243160]
16. Liu C, Fu QJ, Narayanan SS. Effect of bandwidth extension to telephone speech recognition in cochlear implant users. J Acoust Soc Am 2009 Feb;125(2):EL77-EL83. [doi: 10.1121/1.3062145] [Medline: 19206836]
17. Milchard AJ, Cullington HE. An investigation into the effect of limiting the frequency bandwidth of speech on speech recognition in adult cochlear implant users. Int J Audiol 2004 Jun;43(6):356-362. [Medline: 15457818]
18. Mantokoudis G, Kompis M, Dubach P, Caversaccio M, Senn P. How internet telephony could improve communication for hearing-impaired individuals. Otol Neurotol 2010 Sep;31(7):1014-1021. [doi: 10.1097/MAO.0b013e3181ec1d46] [Medline: 20634771]
19. Sun LF, Wade G, Lines BM, Ifeachor EC. Impact of Packet Loss Location on Perceived Speech Quality. 2001 Presented at: Proceedings of 2nd IP-Telephony Workshop; 2001; New York.
20. Friedl A, Ubik S, Kapravelos A, Polychronakis M, Markatos EP. Realistic Passive Packet Loss Measurement for High-Speed Networks. 2009 Presented at: Proceedings of the 1st Traffic Monitoring and Analysis Wokshop (TMA); 2009; Aachen, Germany.
21. Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M. The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. Am J Otol 1997 Nov;18(6 Suppl):S83. [Medline: 9391610]
22. Raake A. Speech Quality of VoIP: Assessment and Prediction. 1st ed. New York: John Wiley & Sons, Ltd; 2006.
23. Yajnik M. Measurement and modeling of packet loss in the internet: University of Massachussetts Amherst, 2000. Measurement and modeling of packet loss in the internet: University of Massachussetts Amherst, 2000 2000.
24. Yajnik M, Moon S, Kurose J, Towsley D. MeasurementModelling of the Temporal Dependence in Packet Loss. 1999 Presented at: INFOCOM99, Eighteenth Annual Joint Conference of the IEEE ComputerCommunications Societies; 1999; New York, USA.
25. Trond U, Stafsnes F. VoIP speech quality - Better than PSTN? Telektronikk 2006;1:119-129.
26. Borella MS, DSwider D, Uludaq S, Brewster GB. Internet Packet Loss: Measurement and Implications for End-to-End QoS. : Measurement and Implications for End-to-End QoS. International Conference on Parallel Processing Workshops (ICPPW'98); 1998 Presented at: International Conference on Parallel Processing Workshops (ICPPW'98); 1998; Minneapolis, USA.
27. Ding L, Goubran RA. Assessment of effects of packet loss on speech quality in VoIP. Haptic, Audio and Visual Environments and Their Applications (HAVE), Proceedings. The 2nd IEEE Internatioal Workshop; Ottawa, Canada 2003:49-54.
28. Skoglund J, Kozica E, Linden J, Hagen R, Kleijn WB. Voice over IP: Speech Transmission over Packet Networks. In: Handbook of Speech Processing. Heidelberg: Springer; 2008:307-330.
29. James JH, Chen B, Garrison L. Implementing VoIP: A Voice Transmission Performance Progress Report. IEEE Communications Magazine 2004;42(7):36-41. [doi: 10.1109/MCOM.2004.1316528]
30. Sommers J, Barford P, Duffield N, Ron A. Improving accuracy in end-to-end packet loss measurement. Computer Communication Review (CCR) 2005;35(4):157-168. [doi: 10.1145/1080091.1080111]
31. Haßlinger G, Hohlfeld O. The gilbert-elliott model for packet loss in real time services on the internet. 2008 Presented at: 14th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB); 2008; Dortmund, Germany.

32. Jiang W, Schulzrinne. Modeling of Packet Loss and Delay and their Effect on Real-Time Multimedia Service Quality. 2000 Presented at: ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV); ; Chapel Hill, North; 2000; Carolina, USA. p. #27.
33. Singh HP, Singh S, Singh J. Computer Modeling & Performance Analysis of VoIP under Different Strategic Conditions. 2010 Presented at: Second International Conference on Computer Engineering and Applications (ICCEA); 2010; Bali Island.
34. Tsompanidis I, Fortetsanakis G, Papadopouli M. Analysis of the perceived VoIP quality under various wireless network conditions. Technical Report. Heraklion, Greece: Institute of Computer Science (ICS): TR 401 2010.
35. Xi B, Chen H, Cleveland WS, Telkamp T. Statistical analysis and modeling of Internet VoIP traffic for network engineering. Electronic Journal of Statistics 2010;4:58-116. [doi: 10.1214/09-EJS473]

Abbreviations

CD: compact disc
HSM: Hochmair-Schulz-Moser
PSTN: public switched telephone network
SNR: signal to noise ratio
VoIP: voice over Internet protocol

Edited by G Eysenbach; submitted 27.03.11; peer-reviewed by H McDermott; comments to author 27.07.11; revised version received 20.10.11; accepted 09.03.12; published 16.07.12.

Please cite as:
Mantokoudis G, Dubach P, Pfiffner F, Kompis M, Caversaccio M, Senn P
Speech Perception Benefits of Internet Versus Conventional Telephony for Hearing-Impaired Individuals
J Med Internet Res 2012;14(4):e102
URL: http://www.jmir.org/2012/4/e102/
doi:10.2196/jmir.1818
PMID:22805169

Original Paper

Animated Graphics for Comparing Two Risks: A Cautionary Tale

Brian J Zikmund-Fisher[1,2,3,4], PhD; Holly O Witteman[3], PhD; Andrea Fuhrel-Forbis[3], MA; Nicole L Exe[3], MPH; Valerie C Kahn[3], MPH; Mark Dickson[3], MA

[1]Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[2]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States

[3]Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI, United States

[4]Risk Science Center, School of Public Health, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Brian J Zikmund-Fisher, PhD
Department of Health Behavior and Health Education
School of Public Health
University of Michigan
1415 Washington Heights
Ann Arbor, MI, 48109-2029
United States
Phone: 1 734 936 9179
Fax: 1 734 763 7379
Email: bzikmund@umich.edu

Abstract

Background: The increasing use of computer-administered risk communications affords the potential to replace static risk graphics with animations that use motion cues to reinforce key risk messages. Research on the use of animated graphics, however, has yielded mixed findings, and little research exists to identify the specific animations that might improve risk knowledge and patients' decision making.

Objective: To test whether viewing animated forms of standard pictograph (icon array) risk graphics displaying risks of side effects would improve people's ability to select the treatment with the lowest risk profile, as compared with viewing static images of the same risks.

Methods: A total of 4198 members of a demographically diverse Internet panel read a scenario about two hypothetical treatments for thyroid cancer. Each treatment was described as equally effective but varied in side effects (with one option slightly better than the other). Participants were randomly assigned to receive all risk information in 1 of 10 pictograph formats in a quasi-factorial design. We compared a control condition of static grouped icons with a static scattered icon display and with 8 Flash-based animated versions that incorporated different combinations of (1) building the risk 1 icon at a time, (2) having scattered risk icons settle into a group, or (3) having scattered risk icons shuffle themselves (either automatically or by user control). We assessed participants' ability to choose the better treatment (choice accuracy), their gist knowledge of side effects (knowledge accuracy), and their graph evaluation ratings, controlling for subjective numeracy and need for cognition.

Results: When compared against static grouped-icon arrays, no animations significantly improved any outcomes, and most showed significant performance degradations. However, participants who received animations of grouped icons in which at-risk icons appeared 1 at a time performed as well on all outcomes as the static grouped-icon control group. Displays with scattered icons (static or animated) performed particularly poorly unless they included the settle animation that allowed users to view event icons grouped.

Conclusions: Many combinations of animation, especially those with scattered icons that shuffle randomly, appear to inhibit knowledge accuracy in this context. Static pictographs that group risk icons, however, perform very well on measures of knowledge and choice accuracy. These findings parallel recent evidence in other data communication contexts that less can be more—that is, that simpler, more focused information presentation can result in improved understanding. Decision aid designers and health educators should proceed with caution when considering the use of animated risk graphics to compare two risks, given that evidence-based, static risk graphics appear optimal.

(J Med Internet Res 2012;14(4):e106) doi:10.2196/jmir.2030

KEYWORDS

Risk; patient education as topic; patient-provider communication; decision aids; visual aids

Introduction

The most basic way to communicate risk to patients is to provide them with a risk number. All numerical formats are not equivalent, however, and considerable research has compared, for example, the pros and cons of using frequency formats instead of percentages [1-4] and the potential pitfalls of specific formats such as 1-in-X [5,6] and number needed to treat [7]. In addition, research on affective influences on risk perceptions make it clear that the same risk number can lead to very different feelings about that risk based on different circumstances, presentations, or contexts [8,9].

In an effort to make risk statistics more intuitive, both researchers and health educators have increasingly turned to visual displays of risk. While this literature is evolving, comparative studies have shown that icon-based displays (often called pictographs or icon arrays) appear to have significant advantages over other displays such as bar graphs and pie charts [10-14]. In particular, displays that visually represent the part-whole relationships (ie, the number of risk events in comparison with the entire at-risk population) appear to be better understood [14,15] and may be of particular help to people with lower numeracy skills [16].

Animated Risk Displays

While most numerical and visual formats for communicating risk are easily implementable in traditional paper-and-pencil materials, the last decade has seen an enormous increase in the use of electronic applications designed to communicate risk to patients. Today, we regularly see computers in clinical consultation rooms, patients can go to innumerable websites that will estimate their health risks, and a growing suite of mobile device applications are available that purport to support healthy living in various ways.

Such technologies open the door to using multimedia techniques such as animation for communicating risk. Animation is commonly used in online applications to call users' attention to a particular area of the screen [17] or portion of the content [18]. This type of signaling may help people make sense of verbal information [19], may help them better learn how a complex system such as a turbofan jet engine works [20], and may help somewhat in the acquisition of complex cognitive skills such as doing experimental research or designing a house [21].

Motion Cues

Animation of risk graphics could be particularly useful because it allows motion cues to draw attention to specific elements of the visual display. For example, instead of simply showing the proportion of area or icons affected by a risk, animation could be used to sequentially draw the viewer's eye to each new risk event, thereby adding a time cue (how quickly the set of events occurs) to reinforce the smallness or largeness of the risk. Such cues may be particularly useful when comparing multiple risks.

Transformations

When people are at risk for a health outcome, they will either experience the outcome or not. The randomness of this occurrence or nonoccurrence is one of the conceptual challenges of risk communication. Past research has noted that icon array displays that scatter event icons randomly convey this sense of randomness but make it difficult to grasp exactly how large the risk is [3,22], whereas displays that group event icons are easy to count (allowing for faster interpretation) [11]. Animation could be used to transform one arrangement into another (for example, a scattered display into a grouped display, or one type of scatter into another), which might enable viewers to have the best of both worlds.

Potential Concerns

While animated displays may have potential advantages, there are also several reasons for caution and further research regarding their use in risk communication applications. Many types of animation exist, each of which may reinforce different types of gist messages. In the absence of research clarifying what types of animation are taken to imply different types of conceptual understanding, use of animated graphics has the potential to cause unintended negative effects. Reviews of other types of animated graphics have found decidedly mixed results [23], providing reason to question whether animated graphics will support improved understanding over static graphs. It may be the case that their utility depends on user characteristics. For example, animated graphs have previously been demonstrated to help people who had performed well on a short mathematics test administered prior to an experiment that tested ability to transform graphs of mathematical functions, while hindering those who had performed more poorly on the mathematics test [24]. The benefits or drawbacks of adding animation cues may also depend on the complexity of the visual stimulus, since human factors research has shown that excessively complicated displays can reduce people's ability to attend to particular cues [18].

Existing Research

Very little work has been done to assess whether animation of risk graphics can improve people's understanding of their health risk, or which types of animation might improve or inhibit accuracy of risk knowledge and risk perceptions. One notable exception is the study of Han et al, which showed that a dynamic scattered display increased subjective uncertainty about a risk, making participants less certain about their interpretations of risk information [25]. Other researchers have examined using interactive risk displays to engage people in learning about risk [26-28], but these studies have focused on the effect of different types of participant tasks (eg, manually graphing a provided risk number or playing within a game-like environment) rather than the visual cues potentially provided by animation.

To begin to explore whether viewing specific types of motion cues and visual transformations in animated risk graphics could improve people's knowledge accuracy, decision making or choice accuracy, and graph evaluation ratings as compared with viewing static images, we conducted a randomized survey experiment comparing various animated and static displays in the context of a hypothetical medical treatment scenario. Our primary research question was whether animated displays with different types of motion would increase or decrease participants' ability to identify which of two treatment options had lower side effect risks (choice accuracy). A secondary research question involved determining, via graph evaluation ratings, whether people like or dislike these types of animation for receiving medical risk information.

Methods

Recruitment

We selected a stratified random sample of US adults age 21 years and older from a panel of Internet users administered by Survey Sampling International (Shelton, CT, USA), which recruits panel members through a variety of opt-in methods. To ensure demographic diversity (though not necessarily representativeness) and offset large expected variations in response rates, we drew distinct subsamples by both age and race (thereby roughly approximating the distributions of these characteristics in the US population), and dynamically adjusted the number of email invitations in each demographic subsample until all quotas were achieved. Selected panel members received an email invitation with a personalized link to complete the online survey with one reminder email for nonresponders. Survey Sampling International tracked participation via unique identification numbers to prevent duplicate uses of the same link to participate. We recruited for a 3-week period in fall 2010. On completion, participants were entered into both an instant-win contest and a monthly drawing administered by Survey Sampling International for modest prizes.

Design of the Study

Respondents read a revised version of a short vignette previously used in a study of interactive graphics [28] in which they imagined being given a diagnosis of thyroid cancer and discussing two types of hypothetical external beam radiation treatments with their doctor. The two treatments, called focal beam therapy and crossed beam therapy, were each briefly described and then presented as being equally effective in treating the patient's type of thyroid cancer. Each therapy also had the same risk (11%) of causing one side effect: fatigue. However, the treatments were described as differentially likely to cause a second side effect, mouth and throat problems: one therapy caused mouth and throat problems in 16% of patients, whereas the other therapy caused mouth and throat problems in only 14% of patients. We randomly assigned which therapy had the higher risk to prevent our scenario descriptions from biasing our results. The scenario referred to a less common disease (thyroid cancer, instead of breast or prostate cancer) and included hypothetical types of radiation beam therapy in order to minimize respondent preconceptions about treatment options or their associated risks.

Our primary research question was to determine whether different graphical formats would increase respondents' ability to recognize which treatment option was less risky (choice accuracy). To do so, we implemented a quasi-factorial design to experimentally vary the type of risk graphic used to present each of the two side effects in a side-by-side presentation. Participants were randomly assigned to 1 of 10 experimental conditions summarized in Table 1.

Table 1. Experimental conditions.

Version		Animated?	Initial arrangement	Animation type		
				Built?	Settle into group?	Shuffle?
V1	Static grouped	No	Grouped	NA[a]	NA	NA
V2	Static scattered	No	Scattered	NA	NA	NA
V3	Scatter, settles	Yes	Scattered	No	Yes	No
V4	Grouped, built	Yes	Grouped	Yes	NA	No
V5	Scatter, built	Yes	Scattered	Yes	No	No
V6	Scatter, built, settles	Yes	Scattered	Yes	Yes	No
V7	Scatter, auto shuffles	Yes	Scattered	No	No	Automatic
V8	Scatter, auto shuffles, settles	Yes	Scattered	No	Yes	Automatic
V9	Scatter, user shuffles	Yes	Scattered	No	No	User controlled
V10	Scatter, user shuffles, settles	Yes	Scattered	No	Yes	User controlled

[a] Not applicable.

Approximately 20% of participants in total were assigned to 1 of 2 static display conditions. Participants in the baseline condition (V1: static grouped) saw side-by-side static icon arrays (a 10 × 10 matrix of blocks) in which all of the colored blocks used to represent event occurrence (ie, experience of fatigue, or mouth or throat problems) were grouped at the bottom of the display (Figure 1). A second group of participants (V2: static scattered) viewed a scattered static display in which the event icons were randomly distributed within the matrix to help convey the underlying random distribution of events. Previous research has suggested that this type of scattered display does, in fact, help convey randomness, but at the expense of a sense of the magnitude of the risk [3,11,22,25]. We included this design factor to explore whether we might achieve the best of both worlds by displaying randomness without sacrificing a sense of quantity.

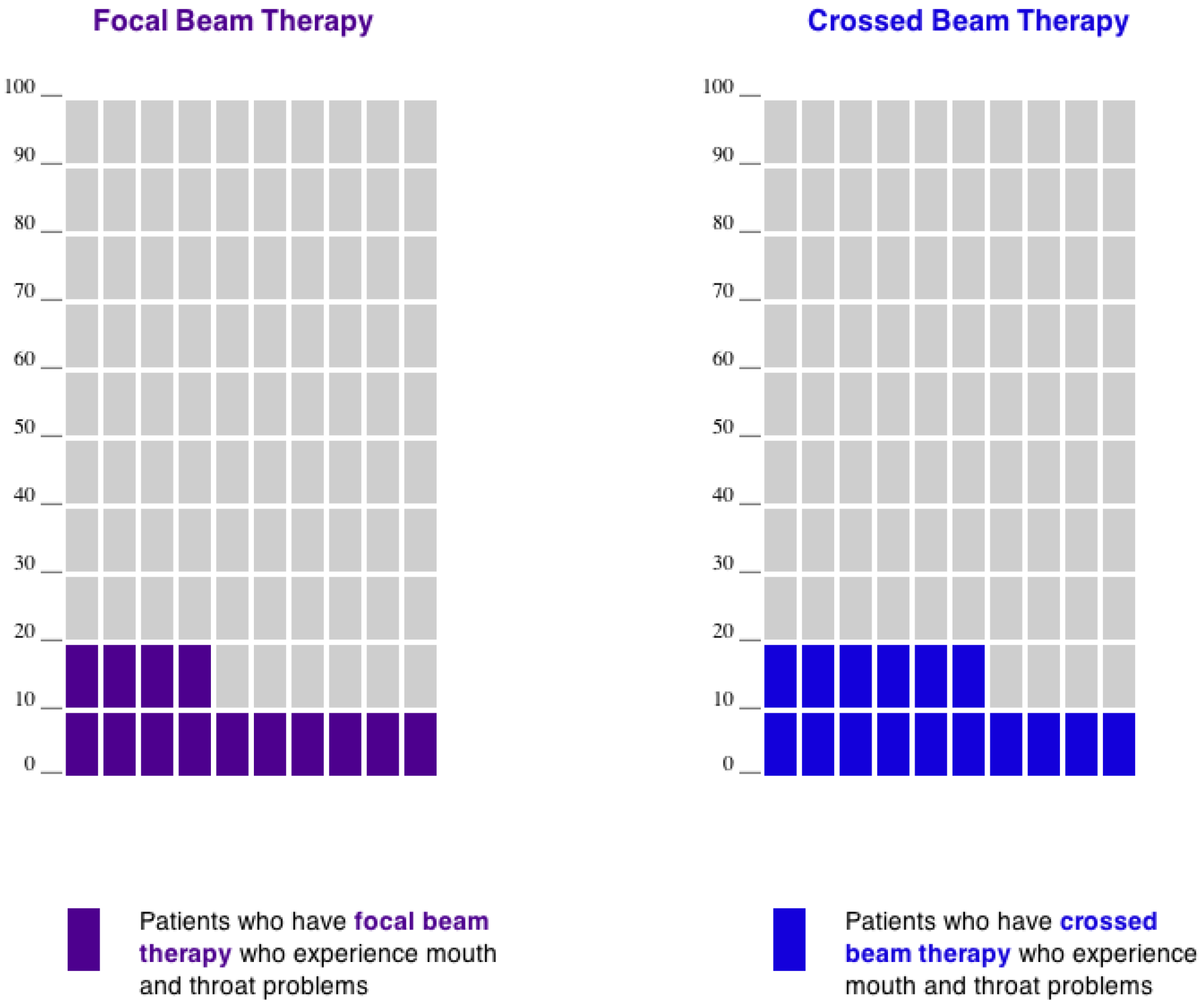
The remaining 80% of participants viewed 1 of 8 animated displays that included 1 or more animations based on the different types of potentially useful motion cues discussed above. Some groups (V4, V5, and V6) viewed built displays that were either grouped or scattered according to the above description. In these versions, participants initially viewed an empty array (ie, all icons were gray) but then saw colored icons representing risk events appear sequentially (1 every 450 milliseconds) in each of the 2 graphs until the final level of risk was reached. Note that due to different levels of risk between the two treatment options, this animation meant that, for the display of mouth and throat problem risks, colored blocks finished appearing in 1 of the 2 arrays before the other one, creating a motion cue to reinforce which treatment had a larger risk.

Participants in the scattered graphs conditions (built or not) were further subdivided based on whether they saw 2 other animations. First, some participants in scattered conditions saw the scattered risk elements remain still (V2 and V5); others (V7 and V8) saw these colored blocks shuffle (redisplay themselves repeatedly in new randomly generated positions) in a manner similar to the dynamic random visual used by Han et al [25] to promote subjective uncertainty; and others (V9 and V10) were required to press a button that caused the blocks to shuffle a few times before they could proceed. We included this last condition to test whether having user control of the random scattering process would affect participants' perceptions of the risk. Second, while some participants saw only the pictographs with the risk scattered (V2, V5, V7, and V9), others initially saw a scattered display that then showed the colored units settling down toward the bottom of the array and then arranging themselves into the same grouping seen by the grouped conditions. In other words, we used animation in these settle conditions (V3, V6, V8, and V10) to enable participants to observe both an initial scattered visual (which may promote understanding of randomness) and an ending grouped visual (to facilitate assessment of risk magnitude).

All survey versions were pretested by study team members for functionality and to estimate time to complete prior to survey launch. We also randomly varied which treatment was shown on the left as well as what color was used to refer to each treatment to prevent order effects. Example movies of V4 (grouped, built), V6 (scatter, built, settles), and V9 (scatter, user shuffles), which collectively demonstrate all of the animation types, are available as multimedia appendices (see Multimedia Appendix 1, Multimedia Appendix 2, and Multimedia Appendix 3).

On entering our survey, participants were given an introduction page that explained the purpose of the study, the anonymous nature of the research, and the expected time to take the survey. The survey consisted of 57 total questions over 20 webpages (between 0 and 9 questions per page). Participants also completed between 3 and 8 webpages of survey materials for unrelated studies (cross-randomized across all 10 arms of this study) after completing the primary and secondary measures for this study but before completing individual difference measures (eg, numeracy) and demographics. This design received institutional review board exempt status approval as anonymous survey research.

Figure 1. Icon arrays (version V1: static grouped) displaying the risk of mouth and throat problems for both hypothetical treatments.



Measures and Covariates

Our primary outcome measure was the preferred treatment choice (focal beam or crossed beam). We also asked respondents two gist knowledge questions in which they were to indicate which therapy had a higher risk of fatigue (both equal), or mouth or throat problems (which varied based on randomization).

We asked 3 graph evaluation items to evaluate user preferences about the different risk graphics. These questions asked respondents to use a 10-point Likert-type scale to rate how well the graphs described the risk of different side effects, how helpful the graphs were, and whether the respondent would like to see risk information in this type of graph. Our a priori intention was to combine these ratings into a 3-item scale.

Because ample evidence exists that even highly educated adults can have poor numeracy skills (ie, facility and comfort with quantitative health information such as risk statistics) [29-31], all study participants also completed the Subjective Numeracy Scale (SNS) [32], a validated measure of quantitative ability and preference for receiving information in numerical form. The SNS has previously been shown to correlate with the ability to recall and comprehend both textual and graphical risk communications [33,34]. A participant's SNS score is calculated as his or her mean rating across the 8 SNS questions and ranges from 1 (least numerate) to 6 (most numerate).

We also assessed participants' need for cognition using a shortened version (7 of 18 questions) of the Need for Cognition Scale [35] due to concerns about survey duration. Responses were averaged to create a single scale.

Statistical Analyses

Since our main goal for this study was to explore the effect of different types of animation on treatment choices, primary analyses focused on the percentage of accurate choices (ie, percentage choosing the treatment with the lowest risk profile), knowledge accuracy, and graph evaluation ratings. To calculate the significance of the observed variations, we used logistic regression models (linear regressions for graph ratings) that included graph version (with V1: static grouped as reference), as well as SNS score and Need for Cognition Scale score as covariates. We also analyzed subsets for the effect of numeracy by splitting at the median, grouping participants into lower-numeracy and higher-numeracy subgroups, and rerunning the logistic regression analyses for each numeracy subgroup (while still controlling for need for cognition). All analyses were performed using Stata version 11 [36], and all tests of significance were 2-sided and used alpha = .05.

Results

Sample Description

In total, 6240 people age 21 years and older reached the survey website and viewed the first content page. Of these, 38 reported having a diagnosis of thyroid cancer and were excluded as having preexisting knowledge of related treatment options, leaving 6202 possible participants.

Overall, 4198 (67.7%) of participants completed the entire survey (range across the 10 survey versions: 64%–71%), including questions on demographics, which came toward the end of the survey instrument. We restricted our analyses to this subsample. Characteristics of those participants who answered each demographic question are reported in Table 2. We observed a wide range of educational achievement, with 1540 participants (36.7%) having a bachelor’s or higher college degree but also 809 (19.3%) having completed high school or less education. The SNS numeracy measure showed high reliability (Cronbach's alpha = .86), as did the shortened need for cognition measure (Cronbach alpha = .83). Mean SNS score was 4.30, with substantial variation (range 1.5–6.0; SD 1.03). Because questions about participants’ demographics came at the end of the survey, we do not know whether the demographics of those who dropped out differed from those who completed the survey.

Table 2. Participant demographic characteristics (n = 4198)[a].

Characteristic	Category	Frequency (%)	Mean (SD)
Age (years)	21–29	698 (16.7%)	49.1 (16.1)
	30–39	663 (15.8%)	
	40–49	588 (14.0%)	
	50–59	848 (20.2%)	
	60–69	1035 (24.7%)	
	70+	359 (8.6%)	
Gender	Male	1936 (46.2%	
	Female	2255 (53.8%)	
Ethnicity	Hispanic (any race)	485 (11.7%)	
Race[b]	White	3267 (78.0%)	
	African American	592 (14.1%)	
	All other	384 (9.2%)	
Education	< High school	86 (2.1%)	
	High school only	723 (17.3%)	
	Some college/trade	1832 (43.8%)	
	Bachelor’s degree	1017 (24.3%)	
	Master’s/doctorate	523 (12.5%)	
Subjective Numeracy Scale score	1.00–1.99	102 (2.4%)	4.30 (1.03)
	2.00–2.99	337 (8.1%)	
	3.00–3.99	970 (23.3%)	
	4.00–4.99	1491 (35.8%)	
	5.00–5.99	1179 (28.3%)	
	6.00	91 (2.2%)	

[a] Reports results only for those respondents who completed each question or measure.

[b] Respondents could mark more than one race.

Treatment Choice Accuracy

Table 3 reports the percentage of participants who correctly chose the dominant treatment option (ie, the therapy with the lower risk of mouth and throat problems), stratified by numeracy level. Lower-numeracy participants selected the best treatment option about 75% of the time, and varying the graphic used to display the side effect risks did not result in any significant differences in treatment choices based on the logistic regression analysis. There was much more variation, however, among higher-numeracy participants. While participants in the baseline V1 group (static grouped) made the correct treatment choice about 85% of the time, participants in most of the other experimental conditions were less likely to pick the best treatment choice. In the logistic regression analysis, we observed significantly different selection rates for participants in all of the scattered but not settled groups (V2, V5, V7, and V9) as well as V10 (scatter, user shuffles, settles). Most of the remaining experimental groups also were less likely than the V1 control group to choose the best treatment, although the differences were not statistically significant. The only group that was more likely (albeit not statistically significant) than the static grouped condition to choose optimally was V4 (built, grouped).

Table 3. Percentage choosing best treatment option, by graph version and respondent numeracy.

Version		Lower numeracy			Higher numeracy		
		%	OR[a] (95% CI[b])	*P* value	%	OR (95% CI)	*P* value
V1	Static grouped	74.3	Reference	–	84.8	Reference	–
V2	Static scattered	72.2	0.90 (0.59–1.36)	.62	76.6	0.58 (0.34–1.00)	.05
V3	Scatter, settles	75.5	1.04 (0.68–1.59)	.85	80.0	0.72 (0.42–1.24)	.23
V4	Grouped, built	75.2	1.02 (0.67–1.55)	.93	86.3	1.14 (0.64–2.04)	.66
V5	Scatter, built	68.3	0.73 (0.48–1.11)	.14	74.4	0.53 (0.31–0.90)	.02
V6	Scatter, built, settles	77.3	1.14 (0.74–1.77)	.55	81.0	0.77 (0.44–1.34)	.35
V7	Scatter, auto shuffles	72.3	0.87 (0.56–1.34)	.52	74.2	0.52 (0.31–0.89)	.02
V8	Scatter, auto shuffles, settles	76.4	1.08 (0.70–1.67)	.72	82.9	0.87 (0.50–1.50)	.61
V9	Scatter, user shuffles	72.9	0.92 (0.60–1.40)	.69	67.6	0.37 (0.22–0.63)	<.001
V10	Scatter, user shuffles, settles	73.3	0.94 (0.62–1.43)	.77	75.8	0.56 (0.33–0.96)	.03

[a] Odds ratio reported from logistic regression model controlling for need for cognition.

[b] Confidence interval.

Gist Knowledge Accuracy

Table 4 reports the percentage of participants (by graph version and numeracy level) who accurately identified that both treatments had equal risks of fatigue. Here, the pattern of results is quite similar for both lower- and higher-numeracy respondents. In both numeracy groups, the baseline group (V1) that saw static grouped graphs had either the highest or next-highest level of knowledge. Knowledge was similar to the baseline in V4 (grouped, built) for lower-numeracy participants, and in V8 (scatter, shuffles, settles) for all participants. Participants in the remaining groups all showed lower knowledge rates, with statistically significant differences observed in the logistic regression analyses for V2, V5 (higher numeracy only), V7 (lower numeracy only), V9 and V10 (higher numeracy only).

The pattern of results was quite similar for accurately identifying the treatment with the higher rates of mouth and throat problems (Table 5), with lower knowledge rates observed for versions V2, V5, V7, and V9 in particular. However, people who saw the V4 (grouped, built) risk graphic had somewhat higher knowledge than those who saw the baseline V1 (static grouped) version, although this difference was not significant for either lower- or higher-numeracy participants.

Table 4. Percentage correctly identifying that both treatments had an equal risk of fatigue, by graph version and respondent numeracy.

Version		Lower numeracy			Higher numeracy		
		%	OR[a] (95% CI[b])	*P* value	%	OR (95% CI)	*P* value
V1	Static grouped	78.8	Reference	–	86.3	Reference	–
V2	Static scattered	68.8	0.59 (0.38–0.91)	.02	73.3	0.44 (0.26–0.75)	.003
V3	Scatter, settles	71.6	0.67 (0.43–1.03)	.07	81.2	0.69 (0.39–1.22)	.21
V4	Grouped, built	79.3	1.01 (0.64–1.60)	.97	82.1	0.74 (0.42–1.32)	.31
V5	Scatter, built	74.0	0.75 (0.48–1.18)	.21	75.0	0.49 (0.28–0.84)	.01
V6	Scatter, built, settles	75.3	0.80 (0.51–1.26)	.34	83.1	0.79 (0.44–1.43)	.44
V7	Scatter, auto shuffles	70.9	0.64 (0.40–1.00)	.05	84.1	0.85 (0.47–1.54)	.60
V8	Scatter, auto shuffles, settles	78.2	0.94 (0.59–1.49)	.79	86.2	0.99 (0.55–1.79)	.98
V9	Scatter, user shuffles	66.1	0.52 (0.34–0.80)	.003	70.2	0.38 (0.22–0.64)	<.001
V10	Scatter, user shuffles, settles	75.3	0.81 (0.52–1.27)	.36	77.7	0.56 (0.32–0.98)	.04

[a] Odds ratio reported from logistic regression model controlling for need for cognition.

[b] Confidence interval.

Table 5. Percentage correctly identifying the treatment with the higher risk of mouth and throat problems, by graph version and respondent numeracy.

Version		Lower numeracy			Higher numeracy		
		%	OR[a] (95% CI[b])	*P* value	%	OR (95% CI)	*P* value
V1	Static grouped	46.9	Reference	–	65.1	Reference	–
V2	Static scattered	44.9	0.92 (0.64–1.34)	.68	55.7	0.68 (0.44–1.04)	.07
V3	Scatter, settles	45.5	0.93 (0.64–1.33)	.68	62.3	0.89 (0.58–1.36)	.59
V4	Grouped, built	54.9	1.35 (0.94–1.95)	.11	69.8	1.25 (0.81–1.95)	.31
V5	Scatter, built	42.4	0.81 (0.55–1.19)	.29	55.0	0.67 (0.43–1.02)	.06
V6	Scatter, built, settles	50.9	1.15 (0.79–1.67)	.46	62.9	0.92 (0.59–1.42)	.70
V7	Scatter, auto shuffles	39.4	0.71 (0.48–1.05)	.09	52.6	0.60 (0.39–0.93	.02
V8	Scatter, auto shuffles, settles	53.0	1.24 (0.85–1.80	.26	64.9	0.99 (0.65–1.52)	.96
V9	Scatter, user shuffles	34.6	0.59 (0.41–0.87)	.007	48.6	0.51 (0.33–0.78)	.002
V10	Scatter, user shuffles, settles	42.9	0.84 (0.58–1.22)	.36	54.8	0.76 (0.50–1.16)	.21

[a] Odds ratio reported from logistic regression model controlling for need for cognition.

[b] Confidence interval.

Graph Evaluation Ratings

As planned, we combined our 3 graph evaluation rating questions into a 3-item scale based on the average rating, which had high reliability (Cronbach's alpha = .93). Table 6 reports the mean graph evaluation ratings for each graph type. Here, a clear pattern emerges: consistent with the previous results for knowledge accuracy and treatment choice accuracy, participants in 2 conditions, V1 (static grouped) and V4 (grouped, built), reported the highest evaluation ratings for both lower- and higher-numeracy participants. Participants in the V3 (scatter, settles) and V6 (scatter, built, settles) groups had slightly lower ratings, though the differences were not statistically significant. The remaining 6 graph types received significantly lower graph evaluation ratings in the linear regression models, with all differences highly significant (all *P* < .001) versus the baseline static grouped condition.

Table 6. Graph evaluation ratings[a], by graph version and respondent numeracy.

Version		Lower numeracy			Higher numeracy		
		Mean	Coefficient[b] (95% CI)[c]	*P* value	Mean	Coefficient (95% CI)	*P* value
V1	Static grouped	5.99	Reference	–	7.24	Reference	–
V2	Static scattered	4.68	–1.30 (–1.75 to –0.86)	<.001	5.07	–2.17 (–2.68 to –1.65)	<.001
V3	Scatter, settles	5.72	–0.29 (–0.73 to 0.15)	.19	6.72	–0.51 (–1.03 to 0.00)	.05
V4	Grouped, built	6.31	0.30 (–0.14 to 0.75)	.19	6.98	–0.25 (–0.77 to 0.26)	.33
V5	Scatter, built	4.89	–1.12 (–1.59 to –0.66)	<.001	5.60	–1.64 (–2.15 to –1.12)	<.001
V6	Scatter, built, settles	5.90	–0.12 (–0.57 to 0.34)	.62	6.81	–0.43 (–0.95 to 0.09)	.11
V7	Scatter, auto shuffles	4.02	–2.01 (2.48 to –1.55)	<.001	4.60	–2.64 (–3.16 to –2.12)	<.001
V8	Scatter, auto shuffles, settles	5.27	–0.75 (–1.20 to –0.29)	<.001	6.08	–1.16 (–1.66 to –0.65)	<.001
V9	Scatter, user shuffles	4.09	–1.91 (–2.36 to –1.46)	<.001	4.77	–2.47 (–2.98 to –1.95)	<.001
V10	Scatter, user shuffles, settles	4.88	–1.12 (–1.57 to –0.67)	<.001	5.34	–1.90 (–2.41 to –1.39)	<.001

[a] Ratings are the average of 3 questions, each reported on a 0–9 scale.

[b] Coefficient from linear regression model controlling for need for cognition.

[c] Confidence interval.

Discussion

Principal Results

In this study, we evaluated 8 different animated icon array risk graphics that incorporated different combinations of 3 basic animations: building risk 1 unit at a time, settling scattered risk into a grouping to ease assessment of magnitude, and shuffling scattered risk to reinforce randomness. When compared against the type of static, grouped icon pictographs that have been previously shown to support high levels of risk knowledge [10], the animated graphics consistently fell short. No animated display resulted in significantly improved knowledge or evaluation ratings versus the static grouped control display, and significant deficits were observed for most of the animated versions. Only the building animation that presented the colored icons representing risk events 1 at a time (eg, V4: grouped, built) showed even the slightest promise of improving understanding, and this was not consistent across outcome measures.

Consistent with some prior research [22], scattered risk displays generally resulted in poorer knowledge and graph evaluation ratings. Shuffling the event icons often made things worse and dramatically lowered evaluation ratings. Adding an animation to allow a scattered risk to settle into a grouping did help, though such animations did not convey any advantages over displays that started in a grouped orientation to begin with. However, a parallel study from our research group (personal communication with HO Witteman, et al, December 8, 2011) suggested that animated displays of scattered icons that include both the building and settling animations may increase sensitivity to differences in risk magnitude. In addition, Han et al found that a dynamic scattered icon array resulted in increased subjective uncertainty about cancer risks [25]. If so, a scatter-plus-settle animation may have practical value even if it does not confer intrinsic improvements in risk knowledge or preference.

Limitations

Our study has several key limitations. First, we recruited participants from an online survey panel and gave them a hypothetical medical treatment scenario. As a result, participants may well have been less motivated to learn about the risk levels and more easily distracted by the animations. This account is consistent with our findings that increased complexity of animations (eg, shuffling) particularly decreased participant knowledge accuracy. It is certainly possible that patients facing actual medical treatment decisions would be better focused on the risk knowledge and, as a result, have smaller deficits or perhaps improvements in understanding over static graphs. We note, however, we found our strongest variations not in knowledge, but in our participants' graph evaluation ratings. Additionally, it is possible that patients facing actual medical treatment decisions would be more susceptible to distraction due to the complexity of animations because of increased cognitive burden or stress brought about by their illness. Although it is plausible that more complicated, "cooler" animations might yield higher evaluation ratings, especially among participants who were only taking a survey and not making real medical decisions, in fact we observed the opposite pattern: the most complex graphics were least preferred by our study participants.

Second, the task used in our experiment required comparing two risks. Animated graphs were presented side-by-side, making it possible that it was the dual presentation of animation, rather than the animation per se, that hampered the communicative effectiveness of the graphics. We selected a comparison task because many risk evaluation and decision processes require balancing competing risks and benefits, making this a plausible application for such graphics. It may be that, although the animation was harmful in this context, it might still hold value in the context of a single risk, or when presenting 1 risk at a time.

Third, our analyses focused here on differences between higher-numeracy and lower-numeracy participants. However, recent research has shown that interpretation of risk graphics is also mediated by graphical literacy skills, which are only moderately correlated with numeracy [37,38]. It is possible that some of the effects we attribute to numeracy are in fact graphical literacy effects. While we did not collect graphical literacy measures here (because these data were collected prior to publication of the scale), we intend to measure both numeracy and graphical literacy in follow-up research.

Comparison With Prior Work

Our study is placed in the context of previous work, most notably Tversky et al's 2002 extensive review of animated graphics [23]. Using their delineation, the task of specifying a dominant option in our study is made more challenging by the fact that health risks are not inherently visuospatial concepts. In their review, the authors noted mixed effects of animated graphics, and suggested that fair comparisons between animated graphics and static graphics require information content that can be adequately conveyed in the static form. This was the case in our context, and this may explain the lack of benefit demonstrated, similar to previous studies where static graphics were able to effectively convey information (eg, [39]). Tversky et al [23] also reviewed other experimental studies, based on which they argued, reasonably, that the benefits of some animated graphics may be attributable to the additional information content that could be conveyed via the animated movement. Therefore, potential benefits of animation may exist, but are concentrated in contexts in which static graphics cannot communicate all the information of the animated versions. The near equivalence of some of our animated versions in performance measures suggests that further research is needed to investigate their potential for conveying additional information that is difficult to convey in static graphics alone, for example, the random nature of events in health risks.

It is also important to note that our task in this study involved the comparison of two risks. This is a common issue in assessing health risks and making decisions accordingly. However, this may have introduced problems by dividing participants' attention between two areas of the screen, both of which were moving simultaneously. This may reflect the fact that people find focusing on competing animations difficult [17].

Previous work suggests an interaction between domain knowledge and the effects of animation. For example, animation helped more advanced students learn to transform graphs of simple mathematical functions into more complex functions but hindered novice students [24]. Our study, on the other hand, suggests that the performance of people with lower numeracy, who might be expected to have more trouble with animation, did not differ across conditions, whereas those with higher numeracy showed a decline in performance with the addition of animation (while still maintaining higher overall rates of accuracy and knowledge). We speculate that this effect depends on whether the animation builds on prior knowledge (as may, for example, be the case with animated displays of physics problems) or distracts from people's ability to perform required tasks. Recent work has shown that more numerate people tend to count icons in displays such as ours and derive their sense of risk magnitude from that process [40]. In our experiments, the attention-grabbing nature of the animation may have prevented higher-numeracy participants from applying this learning strategy, thereby degrading their performance.

It is also worth clarifying the distinction between the animated risk graphics tested here and interactive graphing tasks in which the message recipient has to alter the visual display to show or uncover risk information. Researchers have tested the impact of having people adjust bar graphs [26] or pictographs [28] to display provided risk statistics. However, these studies have had decidedly mixed effects, with the pictograph study finding that the interactive task significantly decreased people's ability to identify a dominant treatment option [28]. Another recent study used an exploratory task in which participants clicked in a matrix until they uncovered a risk event. This task elicited more emotional responses than static graphics, increasing qualitative statements of concern about large risks or relief about small ones. However, a subsequent experimental study found no overall effect of interactive versus static graphic type on risk estimates or risk feelings, though it did reduce disparities attributable to differences in numeracy [27]. Such mixed findings are mirrored in research on other forms of interactivity in health education such as video games [41] and immersive 3-dimensional environments [42].

There are also considerable parallels between our findings regarding the potentially distracting effects of animation and recent evidence in other data communication contexts that less can be more—that is, that simpler, more focused information presentations can result in improved understanding. For example, people are better able to identify preferred hospitals out of a set when tabular presentations of data excluded decision-irrelevant information [43]. Similarly, understanding of cancer recurrence risks and decisions about adjuvant therapies can be improved by removing information about irrelevant options [12], excluding redundant mortality statistics [44], and presenting relevant information one piece at a time [34]. Both these studies and our present investigation serve as reminders that people's ability to process multiple things at once is quite limited, and thus risk communications need to ensure that the user's attention is drawn narrowly and specifically to the most important piece of data or visual cue. In terms of animation, both our study and Tversky and colleagues' [23] review imply that static graphics may be preferable to animated versions as long as the static displays fully present the most decision-relevant information.

Conclusions

If the goal of a health risk communication is to ensure that patients understand the magnitude of risk and are able to make appropriate comparisons between two risks, our work suggests that the use of animation to provide motion cues in computer-administered risk graphics is fraught with peril. We tested 8 combinations of 3 core animations that we believed might support better understanding or satisfaction, but our results showed that these animations were at best unhelpful and often significantly detrimental. Static pictographs that grouped event icons at the bottom of the array consistently resulted in optimal treatment choices, higher knowledge accuracy, and better graph evaluation ratings. This finding adds to the growing literature supporting their use as best practice in many patient education contexts.

Computer-based communications are likely to be the mode of choice for many future efforts to educate patients about health risks that require preventive action and medical treatment decisions. Such technologies offer many new types of visual (and auditory) cues that could be used to reinforce risk information, and there are often pressures to use the latest “bells and whistles” in such applications. Our research, however, sounds a cautionary note.

Ultimately, effective patient decision making requires specific types of understanding. At a minimum, patients need to realize that a risk could occur and be able to identify which actions are more or less likely to lead to preferred outcomes. While we remain hopeful that certain types of animation might be useful in specific risk communication contexts (eg, by using building animations to show accumulation of risk over time), our present efforts did not support quality decision making. More research is clearly needed to evaluate different types of motion cues and to identify which animations lead to better results versus the features of those that do not. In the meantime, we reiterate the finding from our previous work in interactive graphics [28]: decision aid designers should proceed with caution when considering the use of flashy risk graphics.

Acknowledgments

This research was supported by a grant from the Foundation for Informed Medical Decision Making (IIG 0126-1). Dr Zikmund-Fisher is supported by a career development award from the American Cancer Society (MRSG-06-130-01-CPPB). The funding agreements ensured the authors’ independence in designing the study, interpreting the data, and publishing the report. The authors thank Mark Swanson for programming the Flash-based animations for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MP4 movie showing the V4 (Grouped, Built) animation of the risk of mouth and throat problems.

[MP4 File (MP4 Video), 333KB - jmir_v14i4e106_app1.mp4]

Multimedia Appendix 2

MP4 movie showing the V6 (Scatter, Built, Settles) animation of the risk of mouth and throat problems.

[MP4 File (MP4 Video), 349KB - jmir_v14i4e106_app2.mp4]

Multimedia Appendix 3

MP4 movie showing the V9 (Scatter, User Shuffles) animation of the risk of mouth and throat problems.

[MP4 File (MP4 Video), 452KB - jmir_v14i4e106_app3.mp4]

References

1. Waters EA, Weinstein ND, Colditz GA, Emmons K. Formats for improving risk communication in medical tradeoff decisions. J Health Commun 2006 Mar;11(2):167-182. [doi: 10.1080/10810730500526695] [Medline: 16537286]
2. Gigerenzer G. The psychology of good judgment: frequency formats and simple algorithms. Med Decis Making 1996;16(3):273-280. [Medline: 8818126]
3. Schapira MM, Nattinger AB, McHorney CA. Frequency or probability? A qualitative study of risk communication formats used in health care. Med Decis Making 2001 Dec;21(6):459-467. [Medline: 11760103]
4. Miron-Shatz T, Hanoch Y, Graef D, Sagi M. Presentation format affects comprehension and risk assessment: the case of prenatal screening. J Health Commun 2009;14(5):439-450. [doi: 10.1080/10810730903032986] [Medline: 19657924]
5. Cuite CL, Weinstein ND, Emmons K, Colditz G. A test of numeric formats for communicating risk probabilities. Med Decis Making 2008;28(3):377-384. [doi: 10.1177/0272989X08315246] [Medline: 18480036]
6. Zikmund-Fisher BJ. Time to retire the 1-in-X risk format. Med Decis Making 2011 Oct;31(5):703-704. [doi: 10.1177/0272989X11418238] [Medline: 21921147]
7. Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. BMJ 1995 Oct 21;311(7012):1056-9; discussion 1059. [Medline: 7580661]
8. Loewenstein GF, Weber EU, Hsee CK, Welch N. Risk as feelings. Psychol Bull 2001 Mar;127(2):267-286. [Medline: 11316014]

9. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Risky feelings: why a 6% risk of cancer does not always feel like 6%. Patient Educ Couns 2010 Dec;81 Suppl:S87-S93. [doi: 10.1016/j.pec.2010.07.041] [Medline: 20739135]
10. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. Patient Educ Couns 2008 Dec;73(3):448-455. [doi: 10.1016/j.pec.2008.07.023] [Medline: 18755566]
11. Feldman-Stewart D, Brundage MD, Zotov V. Further insight into the perception of quantitative information: judgments of gist in treatment decisions. Med Decis Making 2007;27(1):34-43. [doi: 10.1177/0272989X06297101] [Medline: 17237451]
12. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Improving understanding of adjuvant therapy options by using simpler risk graphics. Cancer 2008 Dec 15;113(12):3382-3390 [FREE Full text] [doi: 10.1002/cncr.23959] [Medline: 19012353]
13. Waters EA, Weinstein ND, Colditz GA, Emmons KM. Reducing aversion to side effects in preventive medical treatment decisions. J Exp Psychol Appl 2007 Mar;13(1):11-21. [doi: 10.1037/1076-898X.13.1.11] [Medline: 17385998]
14. Garcia-Retamero R, Galesic M. Who profits from visual aids: overcoming challenges in people's understanding of risks [corrected]. Soc Sci Med 2010 Apr;70(7):1019-1025. [doi: 10.1016/j.socscimed.2009.11.031] [Medline: 20116159]
15. Stone ER, Sieck WR, Bull BE, Yates JF, Parks SC, Rush CJ. Foreground:background salience: explaining the effects of graphical displays on risk avoidance. Organ Behav Hum Decis Process 2003 Jan;90(1):19-36. [doi: 10.1016/S0749-5978(03)00003-7]
16. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol 2009 Mar;28(2):210-216. [doi: 10.1037/a0014474] [Medline: 19290713]
17. Sutcliffe A, Namoune A. Getting the message across: visual attention, aesthetic design and what users remember. In: Proceedings of the 7th ACM conference on Designing interactive systems. 2008 Presented at: 7th ACM conference on Designing interactive systems; Feb 25-27, 2008; Cape Town, South Africa. [doi: 10.1145/1394445.1394447]
18. Hong W, Thong JYL, Tam KY. Does animation attract online users' attention? The effects of Flash on information search performance and perceptions. Inf Syst Res 2004 Mar;15(1):60-86. [doi: 10.1287/isre.1040.0017]
19. Mayer RE, Moreno R. Animation as an aid to multimedia learning. Educ Psychol Rev 2002 Mar;14(1):87-99. [doi: 10.1023/A:1013184611077]
20. Ozcelik E, Arslan-Ari I, Cagiltay K. Why does signaling enhance multimedia learning? Evidence from eye movements. Comput Human Behav 2010 Jan;26(1):110-117. [doi: 10.1016/j.chb.2009.09.001]
21. Tabbers HK, Martens RL, van Merriënboer JJ. Multimedia instructions and cognitive load theory: effects of modality and cueing. Br J Educ Psychol 2004 Mar;74(Pt 1):71-81. [doi: 10.1348/000709904322848824] [Medline: 15096299]
22. Ancker JS, Weber EU, Kukafka R. Effect of arrangement of stick figures on estimates of proportion in risk graphics. Med Decis Making 2011;31(1):143-150. [doi: 10.1177/0272989X10369006] [Medline: 20671209]
23. Tversky B, Morrison JB, Betrancourt M. Animation: can it facilitate? Int J Hum Comput Stud 2002 Oct;57(4):247-262. [doi: 10.1006/ijhc.1017]
24. Kalyuga S. Relative effectiveness of animated and static diagrams: an effect of learner prior knowledge. Comput Human Behav 2008 May;24(3):852-861. [doi: 10.1016/j.chb.2007.02.018]
25. Han PK, Klein WM, Killam B, Lehman T, Massett H, Freedman AN. Representing randomness in the communication of individualized cancer risk estimates: effects on cancer risk perceptions, worry, and subjective uncertainty about risk. Patient Educ Couns 2012 Jan;86(1):106-113. [doi: 10.1016/j.pec.2011.01.033] [Medline: 21377311]
26. Natter HM, Berry DC. Effects of active information processing on the understanding of risk information. Appl Cogn Psychol 2005;19:123-135. [doi: 10.1002/acp.1068]
27. Ancker JS, Weber EU, Kukafka R. Effects of game-like interactive graphics on risk perceptions and decisions. Med Decis Making 2011;31(1):130-142. [doi: 10.1177/0272989X10364847] [Medline: 20393103]
28. Zikmund-Fisher BJ, Dickson M, Witteman HO. Cool but counterproductive: interactive, Web-based risk communications can backfire. J Med Internet Res 2011;13(3):e60 [FREE Full text] [doi: 10.2196/jmir.1665] [Medline: 21868349]
29. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001 Feb;21(1):37-44. [Medline: 11206945]
30. Nelson W, Reyna VF, Fagerlin A, Lipkus I, Peters E. Clinical implications of numeracy: theory and practice. Ann Behav Med 2008 Jun;35(3):261-274. [doi: 10.1007/s12160-008-9037-8] [Medline: 18677452]
31. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Aff (Millwood) 2007;26(3):741-748. [doi: 10.1377/hlthaff.26.3.741] [Medline: 17485752]
32. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
33. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making 2007;27(5):663-671. [doi: 10.1177/0272989X07303824] [Medline: 17652180]
34. Zikmund-Fisher BJ, Angott AM, Ubel PA. The benefits of discussing adjuvant therapies one at a time instead of all at once. Breast Cancer Res Treat 2011 Aug;129(1):79-87. [doi: 10.1007/s10549-010-1193-4] [Medline: 20945090]

35. Cacioppo JT, Petty RE, Kao CF. The efficient assessment of need for cognition. J Pers Assess 1984 Jun;48(3):306-307. [doi: 10.1207/s15327752jpa4803_13] [Medline: 16367530]
36. StataCorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.
37. Galesic M, Garcia-Retamero R. Graph literacy: a cross-cultural comparison. Med Decis Making 2011 Jun;31(3):444-457. [doi: 10.1177/0272989X10373805] [Medline: 20671213]
38. Gaissmaier W, Wegwarth O, Skopec D, Müller AS, Broschinski S, Politi MC. Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. Health Psychol 2012 May;31(3):286-296. [doi: 10.1037/a0024850] [Medline: 21842998]
39. Mayer RE, Hegarty M, Mayer S, Campbell J. When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. J Exp Psychol Appl 2005 Dec;11(4):256-265. [doi: 10.1037/1076-898X.11.4.256] [Medline: 16393035]
40. Hess R, Visschers VHM, Siegrist M. Risk communication with pictographs: the role of numeracy and graph processing. Judgm Decis Mak 2011;6(3):263-274 [FREE Full text]
41. Annetta LA, Minogue J, Holmes SY, Cheng M. Investigating the impact of video games on high school students' engagement and learning about genetics. Comput Educ 2009 Jul;53(1):74-85. [doi: 10.1016/j.compedu.2008.12.020]
42. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
43. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr;64(2):169-190. [doi: 10.1177/10775587070640020301] [Medline: 17406019]
44. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. A demonstration of "less can be more" in risk graphics. Med Decis Making 2010;30(6):661-671. [doi: 10.1177/0272989X10364244] [Medline: 20375419]

Abbreviations

SNS: Subjective Numeracy Scale

Edited by G Eysenbach; submitted 20.12.11; peer-reviewed by M Feufel; comments to author 17.01.12; revised version received 02.03.12; accepted 24.05.12; published 25.07.12.

Please cite as:
Zikmund-Fisher BJ, Witteman HO, Fuhrel-Forbis A, Exe NL, Kahn VC, Dickson M
Animated Graphics for Comparing Two Risks: A Cautionary Tale
J Med Internet Res 2012;14(4):e106
URL: http://www.jmir.org/2012/4/e106/
doi:10.2196/jmir.2030
PMID:22832208

Review

Are Personal Health Records Safe? A Review of Free Web-Accessible Personal Health Record Privacy Policies

Inmaculada Carrión Señor[1], BSC; José Luis Fernández-Alemán[1], PhD; Ambrosio Toval[1], PhD

Research Group of Software Engineering, Department of Informatics and Systems, Faculty of Computer Science, Regional Campus of International Excellence “Campus Mare Nostrum”, University of Murcia, Murcia, Spain

Corresponding Author:
Inmaculada Carrión Señor, BSC
Research Group of Software Engineering
Department of Informatics and Systems, Faculty of Computer Science, Regional Campus of International Excellence “Campus Mare Nostrum”
University of Murcia
Facultad de Informática
Campus de Espinardo
Murcia, 30100
Spain
Phone: 34 868884621
Fax: 34 868884151
Email: mariainmaculada.carrion@um.es

Abstract

Background: Several obstacles prevent the adoption and use of personal health record (PHR) systems, including users' concerns regarding the privacy and security of their personal health information.

Objective: To analyze the privacy and security characteristics of PHR privacy policies. It is hoped that identification of the strengths and weaknesses of the PHR systems will be useful for PHR users, health care professionals, decision makers, and designers.

Methods: We conducted a systematic review using the principal databases related to health and computer science to discover the Web-based and free PHR systems mentioned in published articles. The privacy policy of each PHR system selected was reviewed to extract its main privacy and security characteristics.

Results: The search of databases and the myPHR website provided a total of 52 PHR systems, of which 24 met our inclusion criteria. Of these, 17 (71%) allowed users to manage their data and to control access to their health care information. Only 9 (38%) PHR systems permitted users to check who had accessed their data. The majority of PHR systems used information related to the users' accesses to monitor and analyze system use, 12 (50%) of them aggregated user information to publish trends, and 20 (83%) used diverse types of security measures. Finally, 15 (63%) PHR systems were based on regulations or principles such as the US Health Insurance Portability and Accountability Act (HIPAA) and the Health on the Net Foundation Code of Conduct (HONcode).

Conclusions: Most privacy policies of PHR systems do not provide an in-depth description of the security measures that they use. Moreover, compliance with standards and regulations in PHR systems is still low.

(J Med Internet Res 2012;14(4):e114) doi:10.2196/jmir.1904

KEYWORDS

Personal health records; privacy; computer security; standards; HIPAA; Health Insurance Portability and Accountability Act

Introduction

In many countries, it is tedious for patients to obtain copies of their official health records from health care providers, which makes it difficult for patients to seek second opinions or control their own information [1]. Moreover, records that patients create themselves tend not to be included in the official patient record. A personal health record (PHR) system can be maintained by patients and their families, can be shared with clinicians, and can support the maintenance of accurate and complete health records [1].

A PHR is "an electronic record of an individual's health information by which the individual controls access to the information and may have the ability to manage, track, and participate in his or her own health care" [2]. A PHR should include all relevant information about the user's life, including the following items: problem list, procedures, major illnesses, allergy data, home-monitored data, family history, social history and lifestyle, immunizations, medications, laboratory tests, and genetic information [3-5].

A PHR can take multiple forms: an independent software application running on a single computer; a Web service belonging to a single organization; a general Web service as a platform with which to collect different types of health information; or a USB-based PHR [6,7]. Maintaining data privacy is difficult in both PHRs and electronic health records (EHRs) [1], to the extent that, for instance, administrative staff could access information without the patient's explicit consent [8]. Consumer concerns regarding PHR systems were found to be focused on two major areas: privacy and security [9]. A total of 91% of surveyed Americans stated that they were very worried about the privacy and security of their health information [7,9]. The aim of this review is to answer the following research question: What security and privacy features do PHR systems have? We carried out an in-depth analysis of many significant issues related to the security and privacy features of PHR privacy policies. The data collected were contrasted by analyzing the privacy aspects of 50% of PHR systems.

Methods

The methods used to carry out the review were guided by a protocol. Iterative decisions concerning data collection, fields for extraction, analysis, and other relevant aspects of the survey were discussed in meetings that were attended and documented by the authors.

Review and Protocol

This review followed the quality reporting guidelines set out in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement [10].

Eligibility Criteria

We used the following inclusion criteria (IC): (1) IC1: free PHR systems, (2) IC2: PHR systems with a Web-based format, and (3) IC3: patient-centered PHR systems with a privacy policy.

Based on the International Organization for Standardization (ISO) standard ISO/TR 12773 (*Business Requirements for Health Summary Records*), a PHR is defined as an electronic, universally available, lifelong resource of health information maintained by individuals, as opposed to an EHR, which is a repository of health information gathered across the longitudinal electronic record of the patient. This information is generated by one or more encounters in any care delivery setting [11]. Among the current variety of PHR support technologies, we focused our study on Web-based, free PHR systems. Free PHR systems can be used by anyone and are easiest to access (IC1). Web-based PHRs have certain benefits with regard to the use of the Internet (IC2) [11]. Moreover, the US Institute of Medicine recommended that "access to care should be provided over the Internet, by telephone, and by other means in addition to in-person visits" [12], while the 2003 Health Information National Trends Survey indicated that consumers use the Internet to access health information more often than they obtain this information from their health care professionals [13]. In addition, the number of users who use the Internet to access and manage their PHR is increasing [14-18]. Finally, according to the ISO, the owner of the record in a PHR system can be the health care organization, provider, or patient [11]. We also stipulated that the PHR systems included in the review should be patient-centered applications—that is, according to the definition of a PHR in the Health Insurance Portability and Accountability Act (HIPAA) [2], the information should be totally or partially managed by the patient (IC3). We analyzed this type of PHR system because they are more flexible and useful than non-patient-centered PHR systems, although they can have more privacy and security problems.

Information Sources

We used two information sources: the myPHR website and scientific databases. The myPHR website was created by the American Health Information Management Association and contains information related to the use and creation of PHRs. To the best of our knowledge, this website provides the most comprehensive list of PHR systems that a user can find and has also been used to select PHR systems in multisource sampling [19]. Although our primary source was myPHR, we identified other PHR systems by reading articles extracted from the Medline, ACM Digital Library, IEEE Xplore Digital Library, and ScienceDirect databases, which we searched between February and April 2011. A systematic review was then used to review the articles indexed in these databases.

PHR System Selection

The PHR system selection process was organized in the following six phases:

1. The search for PHR systems from the myPHR website.

2. The search for PHR systems from scientific databases. This phase was performed by means of a systematic review with the following search string: ("PHR providers" OR "Microsoft HealthVault" OR "Google Health"), which we adapted to database search engines. We next explored the articles identified in order to find the names of Web-based PHR systems.

3. Exploration of the PHR systems found, and a selection based on eligibility criteria IC1 and IC2.

4. Exploration of the PHR websites identified in order to find each one's privacy policy and find out whether the PHR systems were patient-centered applications (eligibility criteria IC3).

5. A complete reading of each of the PHR privacy policies selected in the previous phase to extract their principal privacy and security characteristics.

The activities defined above were carried out by two authors independently. Disagreements were resolved by a third member of the team. The PHR system selection was developed in an iterative process of individual assessments until the interrater reliability was acceptable (0.9). In statistics, interrater reliability is the degree of agreement among raters, which gives a score with the level of consensus of the judges. We use the Cohen kappa coefficient for measuring this agreement. The Cohen kappa [20] coefficient is a statistical measure of interrater reliability for qualitative (categorical) items. A value of 0.9 indicates almost perfect agreement between the two privacy policy assessments performed by two authors.

Data Collection Process

We collected data by using a data extraction form. The PHR system privacy policies were used to extract the methods employed to maintain the privacy and security of the users' data. The privacy policy had to satisfy the security safeguards that are appropriate to the sensitivity of the information. They are used to protect personal information, according to Yee and Korba [21]. Note that Beldad et al [22] state that the omission of an assurance of security in a privacy statement may cause users to think that their personal data are susceptible to potential abuse, and this could discourage them from supplying the personal data needed to complete an online transaction. According to Earp et al [23], it is for this reason that online privacy statements often emphasize the application of security measures and the methods used for the collection of data.

Data Items

In this study, we analyzed security and privacy of PHR systems in reference to the ISO 13606 standard [24]. Security was analyzed in terms of availability, confidentiality, integrity, and accountability. According to the ISO 13606 standard (*Electronic Health Record Communication Part 4: Security*), *availability* refers to the "property of being accessible and useable upon demand by an authorized entity." This standard defines *confidentiality* as the "process that ensures that information is accessible only to those authorized to have access to it." *Integrity* refers to the duty to ensure that information is accurate and not modified in an unauthorized fashion. *Accountability* refers to a person's right to criticize or ask why something has occurred. The other topic analyzed in this study, privacy, has been defined as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [25]. The characteristics analyzed in the privacy policies allowed us to analyze how privacy, integrity, and confidentiality are maintained.

We designed a template for the data to be extracted from each PHR system. In total, 39 characteristics were analyzed and grouped into 12 categories, which we divided into *privacy*, *security*, and *standards and regulations*. Table 1 shows the category descriptions. Some of the characteristics are dependent on others. A complete list of the characteristics analyzed is described in Multimedia Appendix 1.

Table 1. Description of the assessed personal health record (PHR) system characteristics.

Category	Description
Privacy	
Privacy policy location	Considers whether user can easily access the privacy policy
Management and notification of privacy policy changes	Describes whether users are notified of changes in the privacy policy, and the means for doing so
Access management	Focuses on who shares the information, with whom it is shared, and types of permissions
Security: confidentiality and integrity	
Data management	Considers who manages the information, what information is managed, and where this information comes from
Data accessed without user's permission	Describes what data are shared without the user's explicit consent for secondary use of the data (eg, for marketing, policy)
Access audit	Informs whether the user can trace with whom his or her information has been shared
Access criteria	Establishes whether the user is authorized to access the particular resource and what actions she or he is permitted to take with respect to that resource in accordance with certain access criteria
Authentication	Describes the method used to prevent identity theft
Without cookies	Indicates whether the system uses cookies
Safeguards	Presents what security measures are deployed by the PHR system
Standards and regulations	
Standards or regulations	Describes whether the PHR system meets any standards or regulations

Each of these categories satisfied one or more of the eight principles concerning privacy policies by the Canadian Standards Association [21]. The categories, and the principles that they satisfy, are shown in Multimedia Appendix 1.

Quality Assessment

We evaluated each PHR system in relation to its characteristics. We then assigned three scores to each PHR system: total score (range 0 to 24), security score (range 0 to 14), and privacy score (range 0 to 8). The total score was obtained by adding 1 point for each characteristic that was satisfied. The security and privacy scores were obtained considering only the security and privacy characteristics, respectively, of the categories described above. To address the consistency of the rating system, we used triangulation [26] among the raters—that is, more than one researcher gathered and interpreted the security and privacy characteristics. We used a Cohen kappa coefficient of 0.95, which, according to Landis and Koch [20], indicates almost perfect agreement between two privacy policy assessments performed by two authors. In relation to content validity, we thoroughly reviewed the appropriate scientific literature to find recommendations and standards describing good practices for preparing privacy policies [21-23] to identify the items to be included. Experts then critically reviewed this list for relevance, comprehensibility, completeness, and level of detail.

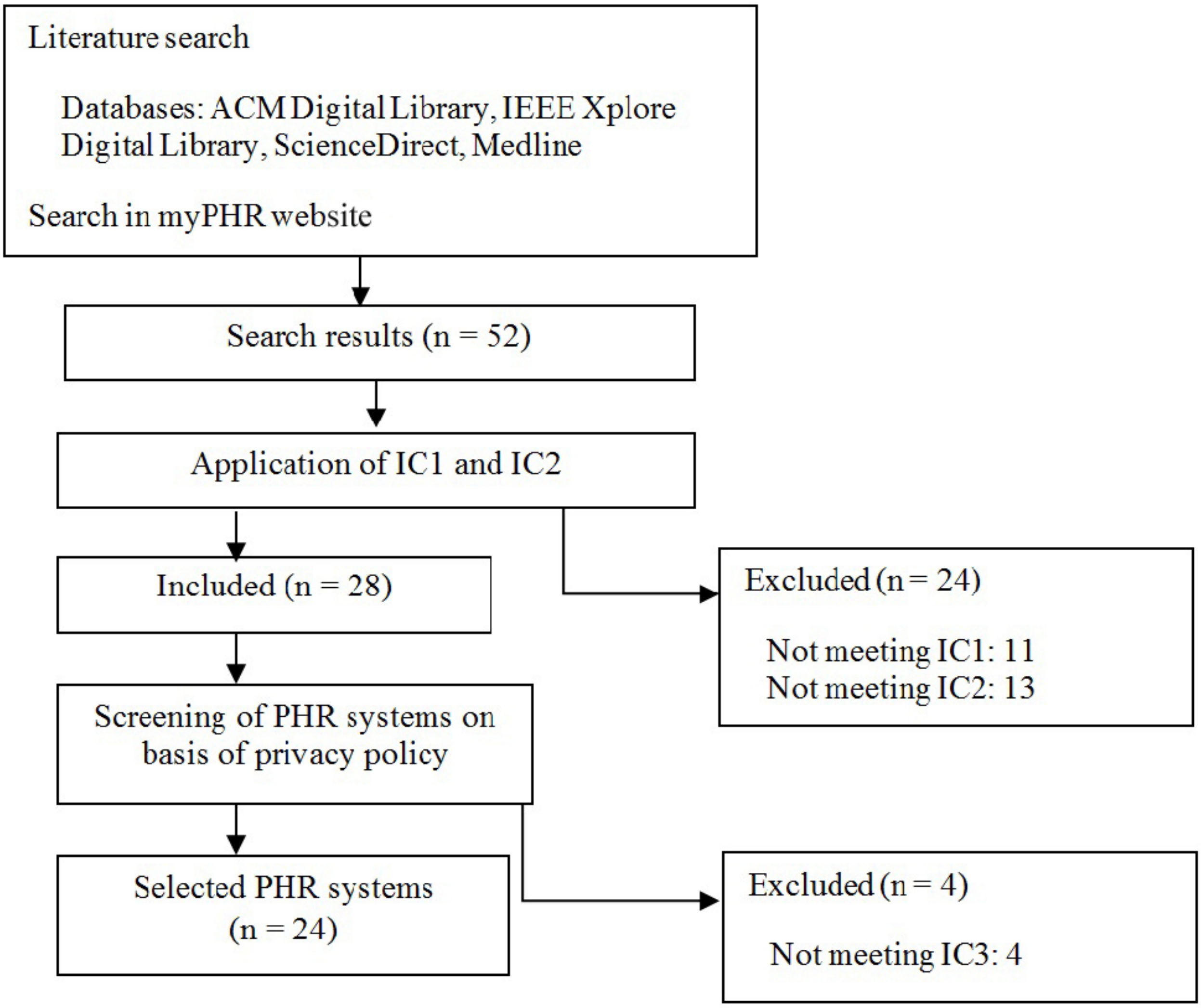
The test-retest [27] method was used to measure the reliability of the measuring procedure. The same test was performed on the same PHR systems after a month. We obtained a correlation of 0.96 between the scores in the two assessments.

Results

Study Selection

We identified 24 PHR systems in the review. The search of databases and the myPHR website provided a total of 52 PHR systems, but we discarded 11 because they did not satisfy IC1 and 13 because the did not satisfy IC2. The privacy policies of the remaining 28 PHR systems were examined, and 4 of these were discarded because they were not patient-centered PHR systems (IC3). Figure 1 shows a PRISMA flow diagram that summarizes this process. The PHR systems included in and discarded from the review are shown in Multimedia Appendix 2.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram. IC1–3 = inclusion criteria 1 to 3, PHR = Personal Health Record.



Study Characteristics

In this section, we describe the most important features of the PHR systems included in the review. Table 2 shows the percentage of PHR systems that satisfy each characteristic analyzed. Table 3 [28-51] shows the systems selected for the study and the three scores assigned to each: security score, privacy score, and total score. More detailed information about the PHR systems analyzed is provided in the tables shown in Multimedia Appendix 3. The percentages and the scores of the dependent characteristics were calculated in relation to the number of PHR systems that met the nondependent characteristic.

Table 2. Evaluation of personal health record (PHR) system characteristics and proportion of PHR systems (n = 24) satisfying each characteristic.

Characteristic	Depends on	n	%
Privacy policy location			
Accessible		23	96
Notification of changes to privacy policy			
Change notification	Accessible	14	61
Change notification on website	Change notification	12	86
Change notification directly	Change notification	3	21
Access management			
User grants access		17	71
User grants access to health care professionals	User grants access	10	59
User grants access to people with other roles	User grants access	3	18
Kinds of permissions		5	21
Access in case of emergency	User grants access	6	35
Data management			
User adds, modifies, removes, and updates information		20	83
Health care professionals update or add information		5	21
Family members’ data	User adds, modifies, removes and updates information	3	15
Connection with other PHRs	User adds, modifies, removes and updates information	4	20
Monitoring devices		2	8
Data accessed without user’s permission			
Not accessed or information related to the user’s accesses		6	25
Access audit			
Who has accessed it		9	38
With what aim	Who has accessed it	2	22
Access criteria			
Roles		13	54
Groups		0	0
Location		1	4
Time		2	8
Transaction type		0	0
Without cookies			
Yes		9	38
Authentication			
Something known		23	96
Something the user has		1	4
Biometric factors		0	0
Safeguards			
Physical security measures		15	63
Limited access		5	21
Electronic security measures		16	67
Encrypted data		12	50
Backup system		4	17

Characteristic	Depends on	n	%
Defined data security plan		1	4
Staff training		1	4
Privacy seal		4	17
Standard or regulations			
HIPAA[a] considered		10	42
HIPAA	HIPAA considered	6	60
HONcode[b]		7	29

[a] Health Insurance Portability and Accountability Act.

[b] Health on the Net Foundation Code of Conduct.

Table 3. The personal health record (PHR) systems and their assigned scores[a].

PHR and reference	Security score	Privacy score	Total score
Microsoft HealthVault [28]	14	7	23
Google Health [30]	10	7	18
NoMoreClipBoard [32]	8	6	16
HealthyCircles [34]	11	4	15
myHealthFolders [36]	10	5	15
RememberItNow! [38]	7	8	15
MiVIA [40]	8	4	14
Telemedical [42]	8	4	13
MedicAlert [44]	7	5	12
Juniper Health [46]	8	4	12
MediCompass [48]	6	3	12
myMediConnect [50]	8	3	12
Health Butler [29]	7	3	11
ZebraHealth [31]	8	1	11
My Doclopedia PHR [33]	5	5	11
Dr. I-Net [35]	7	3	11
Keas [37]	5	4	9
MedsFile.com [39]	6	3	9
PatientsLikeMe [41]	2	6	9
My HealtheVet [43]	6	1	9
dLife [45]	3	3	7
MyChart [47]	4	1	7
EMRy Stick [49]	5	2	7
iHealthRecord [51]	4	1	5

[a] Maximum possible scores: 14 (security score), 8 (privacy score), 24 (total score).

Privacy

The privacy policy document must be easily accessible to PHR systems users. This document was accessible or available in 23 of the 24 PHRs, with myMediConnect being the only PHR system to lack this characteristic. The details of its privacy policy were in the website's FAQ section [50]. Of the PHR systems analyzed, 14 indicated that their users are notified of changes to their privacy policy. Changes could be announced on the home page [28,30,32,33,37,38,40-42,44-46] or via email [39]. A total of 17 of the PHR systems allowed users to grant and revoke access to their data, and 10 indicated that users could grant access to their data to health care professionals. Among these, Google Health and Microsoft HealthVault [28,30] also allowed access to be granted to other system users or to certain services or applications (such as insurance companies or pharmacies).

Only 5 of the PHR systems reviewed defined kinds of permissions. The Google Health PHR system [30] determined two access types for services or applications: write-only access and read/write access. The RememberItNow! PHR system [38] defined three kinds of accesses: write, read, and administrator. Microsoft HealthVault [28] established access levels for users and programs. The Healthy Circles PHR system [34] defined read permission and read/write permission. PatientsLikeMe [41] allowed the contents to be public (anyone could access them) or visible (only PatientsLikeMe users could access them). Finally, only 6 PHRs considered data access in case of an emergency. This access could be total [34] or partial [28].

Confidentiality and Integrity

PHRs contain information users' personal data, which are managed by the user in 20 of the PHRs reviewed. However, MyChart indicated that its users could not manage their own data [47]. Users could only notify the associated health care providers of incorrect data, but not modify them. MyChart was responsible for managing the data. The remainder of the PHR systems did not indicate whether users could manage their data.

A total of 12 PHR systems used aggregated information about users to publish trends or to improve their services [29,30,32-34,37,38,40,41,44,46,48]. Of the PHR systems reviewed, 3 could access users' identifiable data without their consent [39,45,47].

One mechanism that allowed users to verify whether data confidentiality and integrity were maintained is access audit. In this respect, 9 of the PHR systems permitted users to check who had accessed their data [28,30,32,34,36,38,47,49], and 2 of them allowed users to verify what changes were made [28,30].

PHR systems also presented security measures to maintain data integrity and guarantee confidentiality. Of the PHR systems reviewed, 20 indicated whether they used physical or electronic security measures: 15 of them used physical security measures in their servers. On the other hand, we found 12 PHR systems that used encryption to protect the data during transmission [28,30,32,35,36,38,40,43,46,50], and 4 also stored the data encrypted [35,36,43,46]. And 1, ZebraHealth [31], stated that they regularly reviewed and revised data security plans as required by the evolution of technological and security needs. Some PHR systems even had a privacy seal: Microsoft HealthVault, Healthy Circles, Juniper Health, and dLife were certified by TRUSTe [52].

To avoid unauthorized access of users' records, an authentication system is required. The most widespread authentication system was the combination of a user ID with a password, which is something the user knows [28-51]. Some PHR systems combined this with the use of an activation code that had been given to users previously [37,47,48,51]. Only 1 PHR used something the user has for authentication. To access MedsFile.com [39], users had to enter the personal identification number on their access card.

As for the access criteria, the most common one was role-based access control [29,31,32,34,36,39,40,42-44,47,48,51]. PHR systems allowed patients, health care providers, insurances, companies, etc, to access records. Access criteria based on location were applied by 1 PHR [40]. This PHR changed the data shown, such as the list of health care providers, depending on the country from which the user accessed the system. Moreover, 2 PHRs enabled users to establish a period of validity for permissions, which were revoked once this period expired [28,38].

Standards and Regulations

Some legislation and statements are satisfied by or related to the PHR system analyzed. A total of 6 systems complied with HIPAA, while another 4 indicated in their privacy policy that they were not covered by HIPAA, although some of their procedures were inspired by this regulation. Finally, another 7 PHR systems complied with the Health on the Net Foundation Code of Conduct (HONcode) principles for trustworthy health information. HONcode is the oldest and most-used ethical and trustworthiness code for medical and health-related information available on the Internet.

Verification of the Information Contained in Privacy Policies

We triangulated [26] sources of data (chosen at random) to raise the level of confidence in our results and to ensure that the data we collected would enable us to draw valid conclusions. Two authors analyzed the privacy aspects of 11 systems by logging in to the PHR systems' Web portals and verifying whether their privacy policy satisfied the characteristics we had defined. Unfortunately, we were able to verify only a subset of the characteristics analyzed because we could not verify some of them, such as whether the physical measures were really being applied, from the websites. The results obtained were cross-checked against our two initial assessments of the PHR privacy policies. As Table 4 shows, the level of agreement is between high and perfect in three-quarters of the cases [20]. However, the differences we found are not significant because they only lay in some privacy functionalities that were not mentioned in the privacy policies.

Table 4. Kappa coefficients for level of agreement in cross-checks of privacy policy assessment.

Personal health record system	Kappa coefficient	Agreement level
Dr. I-Net	0.42	Low
EMRy Stick	0.77	High
HealthButler	0.79	High
HealthyCircles	0.82	Almost perfect
Juniper Health	0.77	High
Microsoft HealthVault	1	Perfect
My DoclopediaPHR	0.9	Almost perfect
myHealthFolders	0.81	Almost perfect
myMediConnect	0.55	Medium
NoMoreClipBoard	0.62	High
RememberItNow!	0.71	High
Telemedical	0.38	Low

Discussion

The main characteristics of the PHR systems reviewed are summarized below. These characteristics answer our research question of what security and privacy features PRH systems have.

What Security and Privacy Features do PHR Systems Have?

Privacy

In general, most of the PHR systems we reviewed had a document called a privacy policy. This document contains the information related to how the user's information is managed by the system. The user should be able to access this document [53]. Moreover, users must be notified of changes to the privacy policy, given the importance of this document. To fulfill this requirement, most of the PHR systems published an advertisement on their website, which obliges the user to check the PHR website to verify whether the privacy policy has changed. Some authors [54] believe that direct notification of any change is a better solution. One characteristic not found in the PHRs we reviewed is that of notifying users when their data have been exposed. Some regulations, such as the Directive on Privacy and Electronic Communications of the European Union [55], indicate that users have the right to be notified of any personal information disclosure. Most states in the United States also have data breach notification laws [56], which PHR systems must satisfy. These require a data custodian to report a data breach to the individuals affected, state attorneys general, the media, consumer reporting agencies, or other government agencies. One means to ensure that users trust their data security is to indicate that the PHR website is certified by a certification authority. The PHR systems we reviewed used the TRUSTe [52] certification, which guarantees that the security requirements included in the privacy policies are supported by the website.

With regard to PHR access management, 71% of PHR systems allowed users to grant and revoke access to their data. This characteristic is particularly important because users require more flexible ways of sharing data, allowing the user to choose who can access their data, which data they can access, and at what level of access [57]. A problematic issue is the access to users' data in case of emergency—that is, when users cannot explicitly grant access. We found that 35% of PHRs considered this case and provided some type of mechanism to permit the appropriate health care professionals (previously authorized by the user) to access the user's data. Some PHR systems, such as Microsoft HealthVault, allowed users to select what information could be shared and with whom in case of emergency. Nevertheless, emergency access increases the risk of data breaches. Some national laws assume implicit patient consent in an emergency situation [58], which does not guarantee the privacy of patients' data. Moreover, this unusual access adds an extra complexity level to the access control model [58]. On the other hand, not all users are very inclined to share their data in a health emergency. Users with good or excellent health are less likely to share their data during this kind of situation [59].

Finally, ownership of the PHR is an important issue to consider. In Europe, although the PHR can store patient information from a health care provider, the patient owns only the copy stored in the PHR, not the information stored with the provider [60]. This is, for instance, the case in the Dutch system. Such a system allows users to remove data from their PHR, but they cannot remove data from a hospital EHR. Other approaches allowed users to access but not modify their PHR, such as HealthSpace [57]. This may make the PHR of less value to patients and physicians, as no information flows back, but it does provide more security. In the United States, there is the case of My HealtheVet, which is a PHR system developed by the Department of Veterans Affairs. According My HealtheVet's privacy policy, although the content is managed by the Department, the PHR is the property of the veteran and she or he can also manage the information [61].

Confidentiality and Integrity

We examined patient-centered PHR systems in this review, and they allow users to manage their data. In other words, users can add, modify, remove, and update their health data in 83% of cases, according to our review. Connecting the PHR to the EHR would lead to more comprehensive data management by patients [62]. However some physicians have expressed their concern about giving patients so much control over their records, because the information stored in PHRs might be less accurate if patients do not know what exactly is included in them, in comparison with non-patient-centered PHRs [63]. Moreover, if a PHR is hacked—and the patient's data are modified—then, physicians cannot be sure of the correctness of the data [63]. When information comes from several sources, greater privacy and security risks emerge. However, determining the most appropriate strategy remains an open question: to have multiple reliable sources of information, or to have the patient be the only information source.

Few PHRs permit users to check who accessed their data. This aspect should be improved because, according to HIPAA's Privacy Rule and Security Rule and to ISO 13606, users should be aware of how their information has been shared.

We found that 3 (13%) of PHR systems used information related to users' accesses and identified user information to monitor system use without the user's explicit consent. Since the users' privacy should be guaranteed, their identifiable information should not be accessed without their consent [64]. Half of the PHR systems used de-identified or aggregated user information. However, it is very difficult to retrieve sufficient information when aggregated data are used in order to ensure that patients cannot be identified, so some risk of re-identification will usually remain [8,65]. A further issue is that PHR privacy policies did not indicate what information they aggregated. PHR designers could consider studies such as that of Sweeney, who designed a model called k-anonymity, and the accompanying policies that allow the individual's information to be protected, because this cannot be distinguished from, at least, k - 1 other individuals' information [66]. With regard to the information de-identification process, HIPAA indicates that there are two ways to do this: a formal determination by a qualified statistician, or the removal of specified identifiers of the individual and of the individual's relatives, household members, and employers. Removal of identifiers is adequate only if the entity covered has no actual knowledge that the remaining information could be used to identify the individual. In any case, one of these two means is required [2].

The PHR systems must take physical and electronic measures to protect user information [67]. Of the PHR systems we analyzed, in their privacy policies, 63% indicated their physical measures and 67% explicitly stated their electronic security measures; however, only 4 (17%) stated that the data were encrypted both for transmission over the network and for storage. The most widely used encryption scheme for communications was secure socket layer. However, encryption is only part of the solution to protect data. There are also other threats, such as virus-infected systems, against which the PHR systems must be protected. Although there are no well-documented examples of EHR/PHR systems linked to security breaches [68], designers should consider threats to Web applications at least when they deploy their PHR system. In 2008, over 63% of all documented vulnerabilities affected Web applications [69].

Important for security vulnerability is authentication [70]. All the PHR systems we analyzed used only one authentication method, the use of something the user knows or has. However, two of the following three methods are recommended for inclusion in an identification system: something a person knows, such as login ID; something a person has, such as an access card; or something that identifies a person, such as biometrics. Therefore, designers should incorporate another authentication system to strengthen authentication [71]. Moreover, the use of passwords as an authentication mechanism is exposed to multiple types of attacks, such as electronic monitoring of network traffic to capture information, or unauthorized access to the password file.

Finally, 38% of the PHR systems used cookies to remember that the user had already logged in. Using cookies increases the likelihood of identity attacks because the cookie's authentication data can be intercepted by a hacker to gain access to the user's health data [70].

Standards and Regulations

Finally, less than half of the PHR systems we reviewed were based on standards or regulations, and this shows that there is no guarantee that the privacy and security of patients' data is ensured. The most frequently referenced regulation is HIPAA, used in the United States. HIPAA is a federal law that protects health information and ensures that patients have access to their own medical records while assigning new responsibilities to those in charge of protecting this information. Although PHR systems are not required to meet HIPAA by law, users might believe that their data are better protected if the PHR satisfies HIPAA [72].

Limitations

This study had several limitations. Although we conducted a comprehensive literature search on numerous databases using a variety of pertinent search terms, certain PHR systems may have been overlooked due to the lack of indexing in the searched databases. In addition, we recognize that several key PHR systems that were included in the original sample of 51 were excluded as a result of selection criteria. Moreover, we may have excluded some PHR systems if we did not find their privacy policies on their website.

Since this study only analyzed the security and privacy characteristics of PHR systems, it lacks information about the users. Our results cannot easily be generalized to populations, since PHR systems are not equally used by people of different age groups.

The scope of this study did not include analysis of real functionality of PHR systems, and some PHR systems may not satisfy their own privacy policies, so incorrect data may have affected the results of the study. However, this limitation is diminished because we cross-checked the results against an evaluation of actual functionality of 50% of the PHRs.

Another limitation of our study is related to third-party access to the PHR. This characteristic turns PHR systems into a more flexible tool, although it would be necessary to analyze the privacy policies of these parties.

Conclusions

In general, PHR systems allow users to manage their personal health data and to control who has access to them. However, there is a debate regarding the degree to which individuals should be able to control this access, and the forms that this control may take: some PHR systems allow their users only read-only access, while others offer individuals total control [73,74].

The strengths and weaknesses in the privacy and security of PHR systems will be useful for PHR users, health care professionals, decision makers, and system builders. In accordance with the privacy policies, PHR systems do not provide an in-depth description of the security measures used. The designs of privacy policies also need to be improved to include more detailed information related to security measures, and PHR system designers should focus their efforts on increasing the quality of security measures at all stages of system development [75].

The use of standards and regulations by PHR systems is still low. The majority of companies that design PHR systems are not covered by HIPAA [7]. This may be one of the reasons why users do not use PHR systems [72].

Finally, the development of third-party applications that add new functionality to PHR systems is increasing. An example of this is Microsoft HealthVault, which has more than 50 third-party applications [28]. This connection to other applications, such as PHR systems, could also cause important security breaches.

Acknowledgments

This work has been partially financed by the Spanish Ministry of Science and Innovation, project PANGEA, TIN2009-13718-C02-02.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics analyzed and principles that they satisfy.

[PDF File (Adobe PDF File), 47KB - jmir_v14i4e114_app1.pdf]

Multimedia Appendix 2

List of personal health record systems excluded and included in the study.

[PDF File (Adobe PDF File), 29KB - jmir_v14i4e114_app2.pdf]

Multimedia Appendix 3

Characteristics of personal health record systems included in the review.

[PDF File (Adobe PDF File), 75KB - jmir_v14i4e114_app3.pdf]

References

1. Liu LS, Shih PC, Hayes GR. Barriers to the adoption and use of personal health record systems. In: Proceedings of the iConference. New York, NY: ACM; 2011 Presented at: iConference; Feb 8-11, 2011; Seattle, WA, USA p. 363-370. [doi: 10.1145/1940761.1940811]
2. US Department of Health and Human Services, Office for Civil Rights. 2008 Dec 15. Personal Health Records and the HIPAA Privacy Rule URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/special/healthit/phrs.pdf [accessed 2011-08-05] [WebCite Cache ID 60iXgrSXM]
3. Fernandez-Luque L, Karlsen R, Bonander J. Review of extracting information from the Social Web for health personalization. J Med Internet Res 2011 Jan;13(1):e15 [FREE Full text] [doi: 10.2196/jmir.1432] [Medline: 21278049]
4. 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 Apr;13(2):121-126 [FREE Full text] [doi: 10.1197/jamia.M2025] [Medline: 16357345]
5. Adida B, Kohane IS. GenePING: secure, scalable management of personal genomic data. BMC Genomics 2006 Apr;7:93 [FREE Full text] [doi: 10.1186/1471-2164-7-93] [Medline: 16638151]
6. Bonander J, Gates S. Public health in an era of personal health records: opportunities for innovation and new partnerships. J Med Internet Res 2010 Aug;12(3):e33 [FREE Full text] [doi: 10.2196/jmir.1346] [Medline: 20699216]

7. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008 Dec;15(6):729-736 [FREE Full text] [doi: 10.1197/jamia.M2547] [Medline: 18756002]
8. Anderson R, Brown I, Dowty T, Inglesant P, Heath W, Sasse A. Joseph Rowntree Reform Trust Ltd. 2009. Database State URL: http://www.cl.cam.ac.uk/~rja14/Papers/database-state.pdf [accessed 2011-12-02] [WebCite Cache ID 63dH8MiQ0]
9. Connecting For Health: The Personal Health Working Group Final Report. New York, NY: Markle Foundation; 2003. URL: http://www.policyarchive.org/handle/10207/bitstreams/15473.pdf [accessed 2012-08-06] [WebCite Cache ID 69iJt9UgM]
10. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-34. [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
11. Caligtan CA, Dykes PC. Electronic health records and personal health records. Semin Oncol Nurs 2011 Aug;27(3):218-228. [doi: 10.1016/j.soncn.2011.04.007] [Medline: 21783013]
12. Institute of Medicine. 2001 Mar 5. Crossing the Quality Chasm: A New Health System for the 21st Century URL: http://www.iom.edu/~/media/Files/Report%20Files/2001/Crossing-the-Quality-Chasm/Quality%20Chasm%202001%20%20report%20brief.pdf [accessed 2012-08-06] [WebCite Cache ID 69iPFk2Zq]
13. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
14. Sittig DF. Personal health records on the internet: a snapshot of the pioneers at the end of the 20th Century. Int J Med Inform 2002 Apr;65(1):1-6. [Medline: 11904243]
15. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
16. Kim MI, Johnson KB. Personal health records: evaluation of functionality and utility. J Am Med Inform Assoc 2002 Apr;9(2):171-180 [FREE Full text] [Medline: 11861632]
17. Simons WW, Mandl KD, Kohane IS. The PING personally controlled electronic medical record system: technical architecture. J Am Med Inform Assoc 2005 Feb;12(1):47-54 [FREE Full text] [doi: 10.1197/jamia.M1592] [Medline: 15492031]
18. Ball MJ, Smith C, Bakalar RS. Personal health records: empowering consumers. J Healthc Inf Manag 2007;21(1):76-86. [Medline: 17299929]
19. Hulse NC, Wood GM, Haug PJ, Williams MS. Deriving consumer-facing disease concepts for family health histories using multi-source sampling. J Biomed Inform 2010 Oct;43(5):716-724. [doi: 10.1016/j.jbi.2010.04.003] [Medline: 20382264]
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
21. Yee G, Korba L. Personal privacy policies. In: Vacca JR, editor. Computer and Information Security Handbook. Amsterdam: Morgan Kaufmann; 2009:487-505.
22. Beldad A, Jong M, Steehouder M. Reading the least read? Indicators of users' intention to consult privacy statements on municipal websites. Gov Inf Q 2010 Jul;27(3):238-244. [doi: 10.1016/j.giq.2010.01.004]
23. Earp JB, Anton A, Aiman-Smith L, Stufflebeam WH. Examining internet privacy policies within the context of user privacy values. IEEE Trans Eng Manag 2005;52(2):227. [doi: 10.1109/TEM.2005.844927]
24. International Organization for Standardization. 2011. ISO/TS 13606-4:2009: Health informatics -- Electronic Health Record Communication -- Part 4: Security URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50121 [accessed 2011-08-05] [WebCite Cache ID 60iYu0R7f]
25. Westin AF. Privacy and Freedom. London: Bodley Head; 1970.
26. Denzin NK. Sociological Methods: A Sourcebook. New Brunswick, NJ: Aldine Transaction; 2006.
27. Carmines EG, Zeller RA. Reliability and Validity Assessment. Thousand Oaks, CA: Sage Publications; 1979.
28. Microsoft. 2012. Microsoft HealthVault URL: http://www.microsoft.com/en-us/healthvault [accessed 2011-08-05] [WebCite Cache ID 60iZtNKO9]
29. Health Butler, LLC. 2011. HealthButler URL: http://healthbutler.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZYlaTs]
30. Google. 2011. Google Health URL: https://www.google.com/health/ [accessed 2011-08-05] [WebCite Cache ID 60ialazPN]
31. ZebraHealth, Inc. 2011. ZebraHealth URL: https://www.zebrahealth.com/ [accessed 2011-08-05] [WebCite Cache ID 60iaQFShX]
32. NoMoreClipboard.com. 2009. NoMoreClipBoard URL: http://www.nomoreclipboard.com/ [accessed 2011-08-05] [WebCite Cache ID 60iaC3pH9]
33. Doclopedia. My Doclopedia PHR URL: https://www.doclopedia.com/Login.aspx [accessed 2011-08-05] [WebCite Cache ID 60iZvVD0s]
34. HealthyCircles. 2011. HealthyCircles URL: https://www.healthycircles.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZb4z21]
35. Dr. I-Net Corp. Dr. I-Net URL: http://www.drinet.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZTcn9T]

36. What is myHealthFolders? MyHealthFolders, LLC# 2011 URL: https://myhealthfolders.com/ [accessed 2011-08-05] [WebCite Cache ID 60ia7bGZC]
37. Keas, Inc#. 2011. Keas URL: https://keas.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZjrP1e]
38. RememberItNow! LLC. RememberItNow! URL: http://www.rememberitnow.com/ [accessed 2011-08-05] [WebCite Cache ID 60iaEGjsk]
39. MedsFile. 2010. medsfile URL: http://www.medsfile.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZrMEgp]
40. Community Health Resource & Development Center. 2010. MiVIA URL: https://www.mivia.org/ [accessed 2011-08-05] [WebCite Cache ID 60iaMiN6d]
41. PatientsLikeMe. 2011. PatientsLikeMe URL: http://www.patientslikeme.com/ [accessed 2011-12-02] [WebCite Cache ID 63dHNkcBq]
42. Dr. Carr. 2011. Telemedical URL: http://www.telemedical.com/ [accessed 2011-08-05] [WebCite Cache ID 60iaKHr6W]
43. US Department of Veterans Affairs. 2012. My HealtheVet URL: http://www.myhealth.va.gov/ [accessed 2011-08-05] [WebCite Cache ID 60iZz644p]
44. MedicAlert Foundation. 2011. MedicAlert URL: http://www.medicalert.org/ [accessed 2011-08-05] [WebCite Cache ID 60iZmQ7p5]
45. LifeMed Media, Inc. 2011. dLife URL: http://www.dlife.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZOypVa]
46. Toucan Health Systems, LLC. 2011. Juniper Health URL: https://juniperhealth.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZhdziv]
47. Epic Systems Corporation. 2010. MyChart URL: https://www.mychartlink.com/mychart/ [accessed 2011-08-05] [WebCite Cache ID 60ia5GhT5]
48. iMetrikus, Inc. 2011. MediCompass Care Coordination Platform URL: https://www.medicompass.com/mcweb/default.aspx [accessed 2011-08-05] [WebCite Cache ID 60iZoo4n3]
49. MDGuidelines ePHR. Welcome to EmryStick. com URL: http://phr.emrystick.com/ [accessed 2011-08-05] [WebCite Cache ID 60iZVtRvV]
50. MediConnect Global Inc. 2011. myMediConnect URL: http://www.mymediconnect.net/index.php [accessed 2011-08-05] [WebCite Cache ID 60iZKF7Qb]
51. Medfusion. 2007. Weldome to iHealthRecord by Medfusion URL: http://medfusion.net/ihealthrecord/ [accessed 2011-08-05] [WebCite Cache ID 60iZdbCcs]
52. TRUSTe. TRUSTe URL: http://www.truste.com/ [accessed 2011-08-05] [WebCite Cache ID 60iadtEia]
53. Martino L, Ahuja S. Privacy policies of personal health records: an evaluation of their effectiveness in protecting patient information. In: Proceedings of the 1st ACM International Health Informatics Symposium. New York, NY, USA; 2010 Presented at: 1st ACM International Health Informatics Symposium (IHI '10); Nov 11-12, 2010; Arlington, VA, USA p. 191-200. [doi: 10.1145/1882992.1883020]
54. Carrión I, Fernández-Alemán JL, Jayne C, Palmer-Brown D, Toval A, Carrillo de Gea JM. Evaluation and neuronal network-based classification of the PHRs privacy policies. In: Proceedings of the 45th Hawaii International Conference on System Sciences. 2012 Presented at: 45th Hawaii International Conference on System Sciences; Jan 4-7, 2012; Wailea, Maui, HI. USA p. 2840-2849. [doi: 10.1109/HICSS.2012.257]
55. Official Journal of the European Communities. 2002 Jul 31. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2001 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) URL: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:201:0037:0047:en:PDF [accessed 2012-02-16] [WebCite Cache ID 65VNHZNQy]
56. Lesemann D. Once more unto the breach: an analysis of legal, technological and policy issues involving data breach notification statutes. Akron Intellect Prop J 2010;4:203-238 [FREE Full text]
57. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 2010;341:c5814 [FREE Full text] [Medline: 21081595]
58. van der Linden H, Kalra D, Hasman A, Talmon J. Inter-organizational future proof EHR systems. A review of the security and privacy related issues. Int J Med Inform 2009 Mar;78(3):141-160. [doi: 10.1016/j.ijmedinf.2008.06.013] [Medline: 18760661]
59. Weitzman ER, Kaci L, Mandl KD. Sharing medical data for health research: the early personal health record experience. J Med Internet Res 2010;12(2):e14 [FREE Full text] [doi: 10.2196/jmir.1356] [Medline: 20501431]
60. van't Noordende G. Security in the Dutch electronic patient record system. In: Proceedings. New York, NY, USA: ACM; 2010 Presented at: Second annual workshop on Security and Privacy in Medical and Home-Care Systems; Oct 4-8, 2010; Chicago, IL, USA p. 21-32. [doi: 10.1145/1866914.1866918]
61. Nazi KM, Hogan TP, Wagner TH, McInnes DK, Smith BM, Haggstrom D, et al. Embracing a health services research perspective on personal health records: lessons learned from the VA My HealtheVet system. J Gen Intern Med 2010 Jan;25 Suppl 1:62-67. [doi: 10.1007/s11606-009-1114-6] [Medline: 20077154]
62. Maloney FL, Wright A. USB-based Personal Health Records: an analysis of features and functionality. Int J Med Inform 2010 Feb;79(2):97-111. [doi: 10.1016/j.ijmedinf.2009.11.005] [Medline: 20053582]

63. Yau GL, Williams AS, Brown JB. Family physicians' perspectives on personal health records: qualitative study. Can Fam Physician 2011 May;57(5):e178-e184 [FREE Full text] [Medline: 21642732]
64. El Emam K, Jabbouri S, Sams S, Drouet Y, Power M. Evaluating common de-identification heuristics for personal health information. J Med Internet Res 2006 Nov;8(4):e28 [FREE Full text] [doi: 10.2196/jmir.8.4.e28] [Medline: 17213047]
65. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PLoS One 2011;6(12):e28071 [FREE Full text] [doi: 10.1371/journal.pone.0028071] [Medline: 22164229]
66. Sweeney L. k-Anonymity: a model for protecting privacy. Int J Uncertain Fuzziness Knowl Based Syst 2002;10(5):557-570 [FREE Full text] [doi: 10.1142/S0218488502001648]
67. Carrión I, Fernández-Alemán JL, Toval A. Usable privacy and security in personal health records. Lect Notes Comput Sci 2011;6949:36-43. [doi: 10.1007/978-3-642-23768-3_3]
68. Greenhalgh T, Stramer K, Bratan T, Byrne E, Russell J, Hinder S, et al. The Devil's in the Detail: Final Report of the Independent Evaluation of the Summary Care Record and HealthSpace programmes. London: University College London; 2010. URL: http://www.ucl.ac.uk/news/scriesummary.pdf [accessed 2011-12-12] [WebCite Cache ID 63snHmtym]
69. Huynh T, Miller J. An empirical investigation into open source web applications' implementation vulnerabilities. Empir Softw Eng 2010 Oct;15(5):556-576. [doi: 10.1007/s10664-010-9131-y]
70. National Vulnerability Database. DHS National Cyber Security Division/US-CERT. CWE: Common Weakness Enumeration URL: http://nvd.nist.gov/cwe.cfm [accessed 2011-08-05] [WebCite Cache ID 60iaZ4JzW]
71. Park MA. Embedding security into visual programming courses. In: Proceedings of the Information Security Curriculum Development Conference. New York, NY, USA: ACM; 2011 Presented at: Information Security Curriculum Development Conference; Sep 30-Oct 1, 2011; Kennesaw, GA, USA p. 84-93. [doi: 10.1145/2047456.2047469]
72. Srinivasan A. Keeping online personal records private: security and privacy considerations for Web-based PHR systems. J AHIMA 2006 Jan;77(1):62-3, 68. [Medline: 16475740]
73. Bourgeois FC, Taylor PL, Emans SJ, Nigrin DJ, Mandl KD. Whose personal control? Creating private, personally controlled health records for pediatric and adolescent patients. J Am Med Inform Assoc 2008 Dec;15(6):737-743 [FREE Full text] [doi: 10.1197/jamia.M2865] [Medline: 18755989]
74. Steinbrook R. Personally controlled online health data--the next big thing in medical care? N Engl J Med 2008 Apr 17;358(16):1653-1656. [doi: 10.1056/NEJMp0801736] [Medline: 18420496]
75. Fernández-Medina E, Piattini M. Designing secure databases. Inf Softw Technol 2005;47(7):463-477. [doi: 10.1016/j.infsof.2004.09.013]

Abbreviations

EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
HONcode: Health on the Net Foundation Code of Conduct
ISO: International Organization for Standardization
PHR: personal health record
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

Edited by K El Emam; submitted 12.08.11; peer-reviewed by S Porter, B Malin, G van't Noordende; comments to author 17.10.11; revised version received 03.01.12; accepted 30.05.12; published 23.08.12.

Please cite as:
Carrión Señor I, Fernández-Alemán JL, Toval A
Are Personal Health Records Safe? A Review of Free Web-Accessible Personal Health Record Privacy Policies
J Med Internet Res 2012;14(4):e114
URL: http://www.jmir.org/2012/4/e114/
doi:10.2196/jmir.1904
PMID:22917868

Original Paper

Adoption, Acceptability, and Accuracy of an Online Clinical Trial Matching Website for Breast Cancer

Ellyn Cohen[1], PhD; Jeff Belkora[1], PhD; Joanne Tyler[2], MA; Joan Schreiner[2†], MS; Mary Jo Deering[3], PhD; Lakshmi Grama[4], MA, MLS; Brenda Duggan[3], RN; Julie Illi[1], MSN; Julia Pederson[1], BA; Aprajita Anand[1], MPH; Alexandra Teng[1], BA; Erin McCreary[1], BA; Dan Moore[5], PhD; Debu Tripathy[1], MD; Michael Hogarth[6], MD; Morton Lieberman[7], PhD; John Park[1], MD; Laura Esserman[1], MBA, MD

[1]Carol Franc Buck Breast Care Center, University of California San Francisco, San Francisco, CA, United States
[2]Patient Advocate, San Francisco, CA, United States
[3]Center for Biomedical Informatics and Information Technology, National Cancer Institute, Rockville, MD, United States
[4]Office of Communications and Education, National Cancer Institute, Rockville, MD, United States
[5]California Pacific Medical Center Research Institute, San Francisco, CA, United States
[6]Department of Pathology and Laboratory Medicine, University of California Davis, Davis, CA, United States
[7]Department of Psychiatry, University of California San Francisco, San Francisco, CA, United States
[†]deceased

Corresponding Author:
Ellyn Cohen, PhD
Carol Franc Buck Breast Care Center
University of California San Francisco
3450 California Street
San Francisco, CA, 94118
United States
Phone: 1 415 476 0264
Fax: 1 415 476 0262
Email: elly.cohen@ucsfmedctr.org

Abstract

Background: Less than 5% of breast cancer patients participate in clinical trials. To increase patients’ awareness and access to trials, we created BreastCancerTrials.org, a clinical trial matching website. BreastCancerTrials.org matched patients to trials based on their self-reported breast cancer history. It also provided a messaging platform through which patients could self-refer themselves to participating research sites.

Objective: To assess adoption by research sites, acceptability to patients, and patients’ accuracy in providing information to BreastCancerTrials.org.

Methods: We approached 13 research sites in Northern California to list their trials on BreastCancerTrials.org. For adoption, we examined the willingness of contacted research sites to collaborate with BreastCancerTrials.org. For acceptability, we analyzed usage statistics of visitors who completed the BreastCancerTrials.org health history questionnaire in the first 14 months after launch and surveyed users who visited the website during its first year about their experience. For accuracy, we compared the self-reported health history of 20 patients against their medical records. The health history questionnaire was divided into four sections: About Me, personal information including date of birth and sex; My Health as of Today, current status including cancer stage, menopausal status, and sites with evidence of disease; My Cancer, diagnostic information such as hormone and human epidermal growth factor receptor 2 status; and My Treatment, an itemized record of past treatment including responses to therapy.

Results: A total of 12 sites contributed 55 trials. Regarding acceptability, 733 visitors registered on the website; 428 reported their health history; and 407 matched to at least one trial. Of 375 patients who were sent a survey, 75 responded (20%); 23 of the 75 (31%) contacted a research site, 12 of the 23 (52%) were eligible for a trial, and 5 of the 12 (42%) reported enrolling. As for accuracy, 20 clinic visitors reported 1456 health history items, 1324 of which matched their clinic record (90.93%).

Conclusions: BreastCancerTrials.org was adopted by research sites. Patients found it acceptable and were able to provide accurate information for trial matching. Based on our findings, we launched an upgraded version of BreastCancerTrials.org as a national service in October 2008.

(J Med Internet Res 2012;14(4):e97) doi:10.2196/jmir.1855

KEYWORDS

Breast cancer; clinical trials; cancer information

Introduction

Most treatments used for breast cancer today are based on the results of clinical trials. Among the first pivotal trials that changed practice was NSABP B-06 [1], which demonstrated that lumpectomy with lymph node removal and radiation, compared with mastectomy, is as effective at reducing distant recurrence and prolonging survival. Clinical trials have enabled the field to progress by providing better options and more clarity to patients on the risks and benefits of cytotoxic chemotherapy, targeted agents, and local–regional strategies. However, less than 5% of women with breast cancer participate in clinical trials, limiting the pace at which researchers can test new and alternative treatment and prevention strategies [2].

Joining a clinical trial requires awareness of opportunities, alignment of opportunities with the individual’s goals for treatment, and access to research sites that are conducting trials. Patients’ lack of awareness about trials is a major barrier to accrual and the timely completion of clinical trials [3,4]. Most patients learn about trials from their physicians and are more apt to participate if physicians support their patients’ involvement. However, many physicians are unaware of opportunities, do not have the time to review inclusion requirements, or make erroneous assumptions about their patients’ potential eligibility [5,6].

Various initiatives are under way to address barriers to clinical trial participation. Several organizations are conducting awareness campaigns for the general public, as well as underserved populations not typically represented in clinical trials. In the United States, these include the Coalition of Cancer Cooperative Groups, C-Change, the Lance Armstrong Foundation, the Education Network to Advance Cancer Clinical Trials, the Center for Information and Study on Clinical Research Participation, and Eliminating Disparities in Clinical Trials. Additionally, several groups have developed prototype information systems for matching patients’ histories to trial eligibility criteria at the point of care [7-11].

Simultaneously, patients are turning to the Internet for health care-related information [12]. Several resources are available to assist such individuals in finding a clinical trial as described in a 2008 review by the US National Cancer Institute [13]. Patients can search for, evaluate, and locate clinical trials on government-sponsored websites (www.clinicaltrials.gov or www.cancer.gov); on a site sponsored by the nonprofit Coalition of Cancer Cooperative Groups (www.cancertrialshelp.org); and on for-profit websites (eg, www.emergingmed.com) [14-17].

One potential problem with these websites is that they rely on research site locations and contact information listed in government databases. As a consequence, one study pointed out that these sources sometimes feature “out-of-date” research site information, with “incorrect contact information and trial listings” [18].

In 1999, two women with breast cancer, Joan Schreiner and Joanne Tyler, conceived of an independent, nongovernmental, nonprofit clinical trial matching service, available over the Internet, which would enable breast cancer patients to find trials personalized to their situation. With their input, clinical investigators from the Center of Excellence for Breast Cancer Care at the University of California, San Francisco (UCSF) and health and information specialists from the National Cancer Institute developed the concept into a website named BreastCancerTrials.org. By design, BreastCancerTrials.org worked with research sites (ie, the hospitals and clinics that enroll patients) to maintain an up-to-date list of open trials. BreastCancerTrials.org also facilitated a process by which individuals communicate directly with research sites about trials of interest. This design addressed the need to align opportunities with individual conditions and goals, and for patients to access a research site that is implementing a desired trial protocol.

BreastCancerTrials.org launched in 2005 as a regional research study featuring trials in the San Francisco Bay Area. We were generally interested in whether research sites would adopt or collaborate with BreastCancerTrials.org, whether patients would accept the website requirements for matching, and whether patients could accurately report detailed personal health information.

Methods

Study Questions

We formulated 3 specific study questions. First, we asked whether research sites would adopt and collaborate with a clinical trial matching service that requires them to update information about their trials and accept messages from patients. Second, we asked whether patients would complete a detailed online health history, match to a trial, contact a trial site, enroll in a trial, and be satisfied with their experience. Third, we asked whether patients would enter correct information into the BreastCancerTrials.org health history questionnaire.

Study Design

We answered the study questions through a descriptive case study. We launched BreastCancerTrials.org as the intervention, and monitored the behaviors of research sites and patients through data collection mechanisms described below. We sought and obtained approval from the UCSF Committee on Human Research and administered an approved outreach, recruitment, and consent process for both research site investigators and patients who participated in the study.

The 3 study questions involved different dimensions of BreastCancerTrials.org. We therefore present our methods, results, and conclusions in a parallel structure. Specifically, for each section of the ensuing report, we have separate paragraphs for the adoption, acceptability, and accuracy questions.

Intervention

Overall, the intervention consisted of research sites' and patients' use of BreastCancerTrials.org. (Figure 1).

Research sites and patients interacted with different components of the website, so the intervention varied as described below.

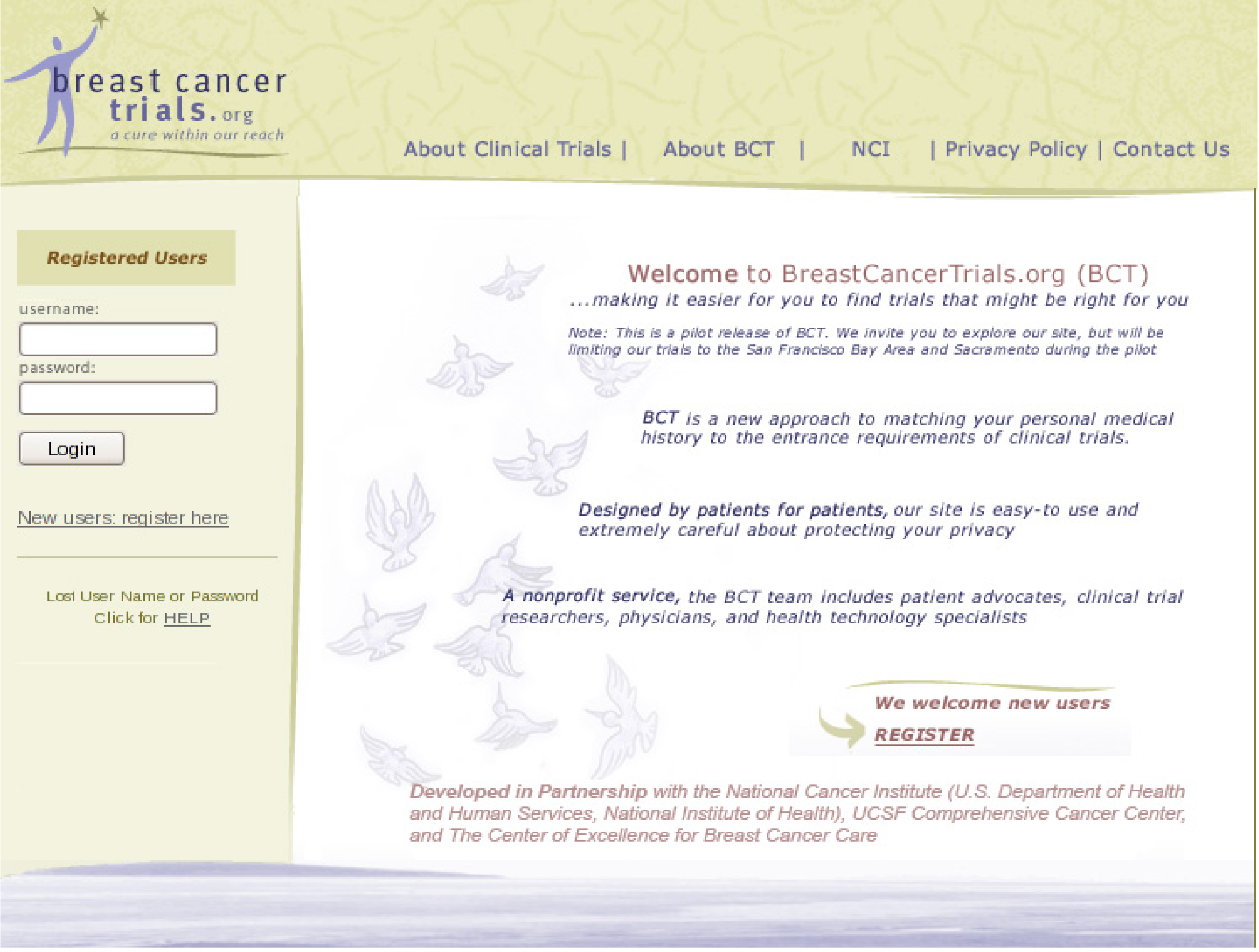
For the question of research site adoption, to appear in BreastCancerTrials.org, research sites had to submit their protocols and contact information and review our coding of the eligibility criteria into a machine-readable format. We solicited study protocols prior to the launch of BreastCancerTrials.org and on a monthly basis afterward. We coded protocols and alerted research sites when their trials were uploaded to the website. At this point site personnel were invited to review coding and provide feedback.

Regarding acceptability to patients, patients using BreastCancerTrials.org were matched to relevant trials based on the completion of a self-reported health history questionnaire. The questionnaire functioned as an intervention in our study, since it was a component of a health service for clinical trial matching.

The health history questionnaire was divided into four sections, three of which captured a fixed number of items: About Me, 12 items, including patients' date of birth and sex; My Health as of Today, 15 items, including current cancer stage, menopausal status, and sites with evidence of disease (Figure 2); and My Cancer, 13 items, including estrogen, progesterone, and human epidermal growth factor receptor 2 (HER2/neu) status. In the fourth section, My Treatment, patients entered their past surgeries as well as radiation, chemo-, hormonal, targeted/biological, and bisphosphonate therapies. After selecting a treatment, patients also provided additional information, for example, whether the treatment was received in the neoadjuvant, adjuvant, or metastatic setting and their response to each therapy. The number of items for My Treatment had no upper limit because respondents could add as many treatments as they had experienced including, for example, multiple cycles of chemotherapy. The matching engine, caMatch, was developed in collaboration with the National Cancer Institute, based on a prototype built at UCSF. It compared patients' health history questionnaires with trial eligibility requirements in the coded protocols. The matching engine provided a report of relevant trials along with the contact information for the enrollment coordinator at each trial site. An online messaging platform was included as part of the matching service. It allowed patients to contact a research site about their interest in a specific matched trial and to invite a coordinator to view their online health history. Patients' consent was required for BreastCancerTrials.org to make the history available to a specific coordinator for online viewing; patients' health summaries were never sent to coordinators via email.

In addressing the accuracy question, the intervention consisted of asking UCSF breast cancer patients to register on the website and complete the health history questionnaire. Subsequently we evaluated the accuracy of each patient-reported health history by comparing it with the patient's official medical record, as described in the outcomes section below.

Figure 1. Homepage of BreastCancerTrials.org.



Welcome to BreastCancerTrials.org (BCT)

...making it easier for you to find trials that might be right for you

Note: This is a pilot release of BCT. We invite you to explore our site, but will be limiting our trials to the San Francisco Bay Area and Sacramento during the pilot

BCT is a new approach to matching your personal medical history to the entrance requirements of clinical trials.

Designed by patients for patients, our site is easy-to use and extremely careful about protecting your privacy

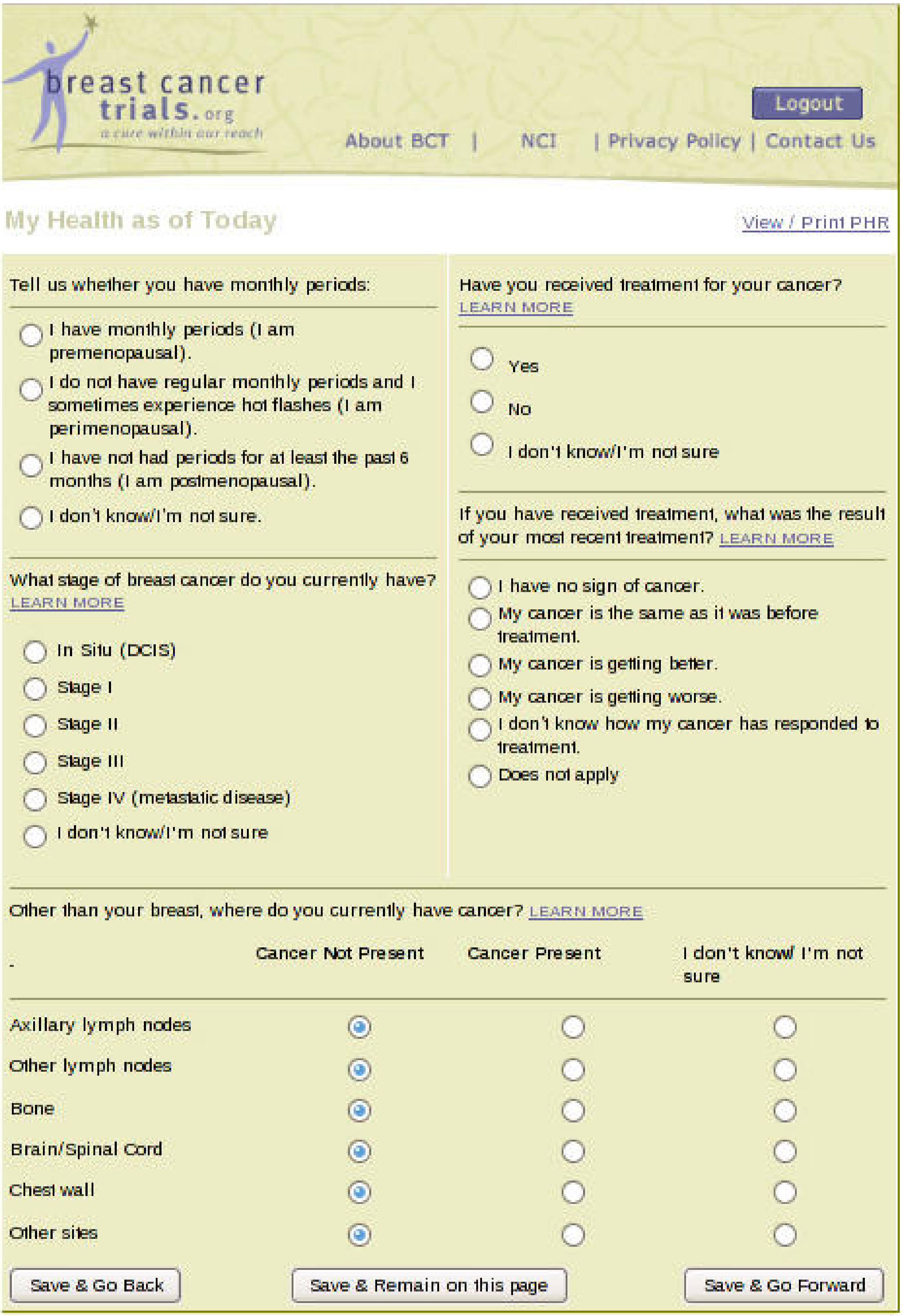
A nonprofit service, the BCT team includes patient advocates, clinical trial researchers, physicians, and health technology specialists

We welcome new users

REGISTER

Developed in Partnership with the National Cancer Institute (U.S. Department of Health and Human Services, National Institute of Health), UCSF Comprehensive Cancer Center, and The Center of Excellence for Breast Cancer Care

Figure 2. A section from the My Health portion of the BreastCancerTrials.org health history questionnaire.



My Health as of Today

Tell us whether you have monthly periods:

- ○ I have monthly periods (I am premenopausal).
- ○ I do not have regular monthly periods and I sometimes experience hot flashes (I am perimenopausal).
- ○ I have not had periods for at least the past 6 months (I am postmenopausal).
- ○ I don't know/I'm not sure.

What stage of breast cancer do you currently have? LEARN MORE

- ○ In Situ (DCIS)
- ○ Stage I
- ○ Stage II
- ○ Stage III
- ○ Stage IV (metastatic disease)
- ○ I don't know/I'm not sure

Have you received treatment for your cancer? LEARN MORE

- ○ Yes
- ○ No
- ○ I don't know/I'm not sure

If you have received treatment, what was the result of your most recent treatment? LEARN MORE

- ○ I have no sign of cancer.
- ○ My cancer is the same as it was before treatment.
- ○ My cancer is getting better.
- ○ My cancer is getting worse.
- ○ I don't know how my cancer has responded to treatment.
- ○ Does not apply

Other than your breast, where do you currently have cancer? LEARN MORE

-	Cancer Not Present	Cancer Present	I don't know/ I'm not sure
Axillary lymph nodes	●	○	○
Other lymph nodes	●	○	○
Bone	●	○	○
Brain/Spinal Cord	●	○	○
Chest wall	●	○	○
Other sites	●	○	○

Sample

For the question about adoption by research sites, from February 2005 to May 2005, author EC identified a convenience sample of 13 institutions in the 9-county San Francisco Bay Area and Sacramento, where she had found investigators who were conducting breast cancer clinical trials. The research sites included academic, private practice, and managed care settings. EC invited these research sites to submit descriptions of their open trials, including eligibility criteria, prior to the regional launch of BreastCancerTrials.org and on a monthly basis thereafter.

For the question about acceptability to patients, we included all people who registered on the BreastCancerTrials.org website in the first 14 months after launch (between June 2005 and September 2006) in our analysis of website usage. Between July 1, 2005 and June 30, 2006, we sent emails inviting website registrants who had completed the health history questionnaire in the first 12 months after launch to respond to our survey asking for user ratings of the website.

For the question about patients' accuracy, between April 25, 2006 and March 29, 2007, we invited a convenience sample of patients who had completed at least one episode of treatment at the UCSF Carol Franc Buck Breast Care Center. Through a process approved by our institutional review board, a research assistant periodically reminded the 12 attending physicians at the Breast Care Center of the opportunity to refer patients to this study. The physicians served as referral sources but were not otherwise involved in the study. The research assistant went to the clinic whenever patients referred by their physicians were being seen. The research assistant approached such patients until we had accrued our target of 20 who had completed the study requirements by filling out the health history questionnaire online.

Measures and Data Collection Procedures

We measured research site adoption by counting the number of sites that contributed trials to the regional launch of BreastCancerTrials.org, and by tallying the number of trials each research site submitted and kept up-to-date in the BreastCancerTrials.org repository of trials.

We measured acceptability to patients of BreastCancerTrials.org by observing the number of (1) patients registering, (2) consenting to the study and starting the health history questionnaire, and (3) completing the health history questionnaire. We measured patients' reactions to BreastCancerTrials.org by administering a follow-up, anonymous survey with 13 items, 6 of which were relevant to our study questions. We therefore analyzed responses to the following 6 summative evaluation items. A total of 3 yes/no items probed whether the patient contacted a research site after matching to a trial on the website; whether they were eligible for the trial on further review by the research site; and whether they ultimately enrolled in the trial. For these binary items, we counted the number and proportion of respondents answering affirmatively. A further 3 items addressed overall satisfaction; likelihood to recommend to friends or colleagues; and ease of navigation [19]. These were rated on a 10-point scale where 1 anchored the most negative response and 10 the most positive. We calculated the mean responses to each of these items individually. The remaining 7 survey questions were not directly related to our study questions and we did not analyze them for this study.

We measured accuracy of patient-reported health history by comparing patients' responses to the BreastCancerTrials.org health history questionnaire with the health history reflected in the patients' medical charts. Patients filled out the health history questionnaire via the BreastCancerTrials.org website, using their own personal records and recollections. Meanwhile, a study coordinator (author AA or JP) abstracted data from each respondent's medical chart and created a second, chart-based instance of the patient's health history questionnaire. We treated this chart-based version of the health history as the reference standard (sometimes known as the gold standard). We calculated accuracy by forming a ratio. Where patients had either answered a question incorrectly or omitted an item that was reported in the chart, we defined the denominator as the total number of items reported in the chart-based health history, and defined the numerator as the number of responses where the patient and chart agreed. Where patients provided items that were not in the chart, we added the number of such items to the denominator, and kept the numerator as the number of responses where the patient and chart agreed. We omitted from the accuracy analysis 12 items in the About Me section of the survey, as we considered the patient's current responses to these demographics questions definitive regardless of what the chart indicated. We also omitted 2 items from the My Cancer section, neither of which was used for matching. These were "Are tissue samples from your breast cancer available for further testing?" and the optional question "If you were positive for HER2/neu, what method was used to test it?"

Analysis

We analyzed quantitative measures using descriptive statistics including counts, means, medians, and standard deviations. We analyzed qualitative measures, such as open-ended survey responses, by reading for themes and then discussing them among study authors until we arrived at a consensus interpretation.

Results

Regarding research site adoption, we approached 13 San Francisco and Sacramento-area health care organizations. At each research site, we invited a clinical trials manager or physician in charge to coordinate submission of active breast cancer study protocols and institutional review board approval documents to our research team. Of the 13 research sites, 12 completed the requirements for participation, by providing sufficient information for our research team to code their trials using our structured data entry forms. Of the 12 research sites, 2 were academic medical centers, 6 were community hospitals, 1 was a participant in the National Cancer Institute-funded Community Clinical Oncology Program, 1 was a health maintenance organization, and 2 were private oncology practices (Table 1). These 12 research sites contributed all of their breast cancer protocols, representing 55 studies during the study period.

The BreastCancerTrials.org team coded the trials in the BreastCancerTrials.org database and sent them to the research sites for review and approval. The research sites reported no disagreements with the BreastCancerTrials.org codes, so the team activated the trials on the production website. A total of 11 research sites remained in the study through its completion; 1 withdrew after the first year because the clinical trial manager submitting active study documents felt this required too much time.

With regard to patients using the site and matching to clinical trials, the population for this study question consisted of 733 patients registering on BreastCancerTrials.org between June 2005 and September 2006. Registration involved providing a name and email address, at which point patients could navigate the website and consent (or not) to use the matching service. Of 733 registrants, 614 (83.8%) consented to use the matching service, of whom 491 (80.0%) also reported demographics (Table 2).

We tracked website usage as an indication of BreastCancerTrials.org's acceptability to patients. Of the 614 patients who registered and consented, 428 (69.7%) completed the minimum health history elements required to match to a trial. Of these 428, 407 (95.1%) matched to at least one trial. Of the 407 matched visitors, 70 (17%) connected with a research site through BreastCancerTrials.org's Message Center.

Between July 1, 2005 and June 30, 2006, we sent email surveys to 375 website users who had completed the health history questionnaire during the first 12 months after launch, and invited them to respond anonymously. We found that 75 of 375 (20%) responded, and 23 of the respondents (31%) reported contacting a research site. Among the 23 who contacted a research site, 12 (52%) reported being told they were eligible for a trial, and 5 of these 12 (42%) reported enrolling in a trial.

Table 1. Research sites providing trials at the launch of BreastCancerTrials.org.

Practice setting	Research site
Academic	University of California, San Francisco
	University of California, Davis
Community hospital	Sutter East Affiliated Hospitals (Sacramento)
	Sutter West Affiliated Hospitals (Alta Bates Summit Medical Center, California Pacific Medical Center, Mills-Peninsula Health Services)
	John Muir Health
Community Clinical Oncology Program	Bay Area Tumor Institute
Health maintenance organization	Northern California Kaiser Permanente
Private practice	Camino Alto (Peninsula)
	California Cancer Care (Marin)

Table 2. Demographics characteristics of respondents to acceptability and accuracy study questions.

Demographic characteristic		Acceptability analysis				Accuracy analysis	
		Website visitors reporting demographics (n = 491)		Online satisfaction survey respondents (n = 75)		Clinic-based health history questionnaire respondents (n = 20)	
		n	%	n	%	n	%
Age range (years)							
	<45	132	26.9%	15	20%	5	25%
	45–54	178	36.3%	23	31%	8	40%
	55–64	140	28.5%	17	23%	6	30%
	65+	39	8%	3	4%	1	5%
	No response	2	<1%	17	23%		
Education							
	Graduate/professional school	200	40.7%	31	41%	14	70%
	College graduate	163	33.2%	16	21%	6	30%
	<4 years of college	128	26.1%	7	9%	0	0%
	No response			21	28%		
Race/ethnicity							
	White	413	84.1%	49	65%	19	95%
	Hispanic/Latino	34	7%	2	3%	0	0%
	Black/African American	16	3%	2	3%	0	0%
	Asian/Pacific Islander	14	3%	2	3%	1	5%
	Other	14	3%	3	4%	0	0%
	No response			17	23%		

We present demographics reported by 75 survey respondents in Table 2 (middle columns). Respondents rated satisfaction with BreastCancerTrials.org at a mean level of 7 out of a maximum of 10. Respondents rated willingness to recommend BreastCancerTrials.org at a mean level of 7 out of 10. Finally, respondents rated the ease with which they completed the health history questionnaire at a mean level of 8 out of 10.

With regard to study questions about patients' accuracy, 26 patients consented out of 57 approached (46%), 20 of whom completed the study requirements, for a response rate of 20 out of 57 (35%) among all approached, or a completion rate of 20 out of 26 (77%) among consenting patients. We invited them to fill out the health history questionnaire so we could compare their responses with the information in their medical record. Table 2 summarizes the demographic profiles of the final sample (last columns). These patients provided a total of 1456 items, matching the chart for 1324. Therefore, the overall accuracy rate was 90.93%. For the standard items in the My Health and the My Cancer sections, the accuracy rate was 469 out of 520 (90.2%). For the variable items in the My Treatment section, the accuracy rate was 855 out of 936 (91.4%).

On an item-by-item level, accuracy ranged from 65% to 100% (see Table 3). On the low end of the range, 13 of the 20 (65%) respondents to the standard items correctly reported their progesterone status. At the high end of the range, respondents were 100% accurate in reporting their status with respect to pregnancy, current well-being, hypertension, cardiac arrhythmia, disease in other sites, local recurrence, type of bisphosphonate therapy taken, and type of biologic therapy taken.

Table 3. Accuracy of patients' responses to health history questionnaire items, compared with a study coordinator's abstraction of the data from the patients' medical charts.

Item by website section	Total responses per Item	Items for which patient's response matches chart	Accuracy rate
My Health as of Today			
Menopausal status	20	19	95%
Stage of current cancer	20	15	75%
Treatment received? (yes/no)	20	20	100%
Response to last treatment	20	15	75%
Pregnant or nursing	20	20	100%
Current well-being	20	20	100%
Hypertension	20	20	100%
Cardiac arrhythmia	20	20	100%
Other medical problems	20	17	85%
Disease in axillary lymph nodes	20	16	80%
Disease in nonaxillary lymph nodes	20	19	95%
Disease in bone	20	19	95%
Disease in brain or spinal cord	20	19	95%
Disease in chest wall	20	17	85%
Disease in other sites	20	20	100%
Subtotals	300	276	92.0%
My Cancer			
Date of breast cancer diagnosis	20	19	95%
Stage at diagnosis	20	18	90%
Type of cancer	20	17	85%
Estrogen receptor status	20	18	90%
Progesterone receptor status	20	13	65%
Positive lymph nodes	20	17	85%
Sentinel node biopsy	20	17	85%
Local recurrence	20	20	100%
Metastasis	20	19	95%
Inflammatory breast cancer	20	19	95%
HER2/neu[a] status	20	16	80%
Subtotals	220	193	87.7%
My Treatment			
Treatment modality (eg, surgery, radiation, chemotherapy)	96	84	88%
Type of surgery (eg, mastectomy)	150	137	91.3%
Location of radiation therapy (eg, breast, bone)	63	57	90%
Type of hormone therapy (eg, tamoxifen)	91	90	99%
Type of bisphosphonate therapy (eg, Zometa[b])	12	12	100%
Type of biologic therapy (eg, Herceptin[c])	8	8	100%
Type of chemotherapy (eg, paclitaxel)	340	328	96.5%
Setting (neoadjuvant, adjuvant, metastatic)	88	81	92%

Item by website section	Total responses per Item	Items for which patient's response matches chart	Accuracy rate
Overall response (eg, no cancer, cancer progressed)	88	58	66%
Subtotals	936	855	91.4%
Total	1456	1324	90.93%

[a] Human epidermal growth factor receptor 2.

[b] Generic name zoledronic acid.

[c] Generic name trastuzumab.

While patients were accurate overall, we call attention to items for which patients had difficulty matching the chart. A total of 5 patients did not match the chart with respect to the stage of current cancer; 5 patients were not accurate in responding to response to last treatment; and 5 patients were inaccurate in responding to items about disease in lymph nodes, 4 of whom were wrong about axillary and 1 about nonaxillary lymph nodes.

Discussion

Interpretation and Analysis

Regarding the adoption by research sites, in the 3 months prior to launch, we attracted 12 of 13 research sites, 11 of which remained active in the study 1 year later. We had some concerns about whether the research sites would see collaborating with BreastCancerTrials.org as an added burden. Our experience suggests that research sites were willing to add BreastCancerTrials.org as an additional channel through which to reach potential study participants. We believe that in an environment of relatively low study participation, the value of adding participants is so high that research sites are motivated to spend additional time and other resources on exposing potential study candidates to the opportunity.

Regarding the acceptability to patients, almost all visitors (407/428, 95.1%) who provided sufficient information for matching did in fact match to at least one trial. Website visitors were satisfied, and 70 contacted a research site. These data support our hypothesis that as more people seek health information online [12], many could be attracted to websites dealing with clinical trials.

The demographics of participants who used the BreastCancerTrials.org matching tool were highly educated and predominantly white, not surprising given the demographics of people using the Internet to find health information in 2005 [12]. As more diverse populations use the Internet to access health information, further studies should explore whether they too are open to using BreastCancerTrials.org or similar clinical trial matching websites.

As for the accuracy of patients' responses, the BreastCancerTrials.org health history questionnaire used in this study was notable for its level of detail. We presented a minimum of 40 items before patients even began filling out My Treatment, which could also require dozens of items depending on the complexity of treatment history. Our assumption that patients would be capable of answering detailed questions about their cancer treatment was validated by the high degree of accuracy when their answers were compared with their medical records. For example, 1 patient with metastases entered 10 chemotherapy regimens correctly.

We included both estrogen and progesterone receptor status in the health history questionnaire. However, during the study period, physicians in our clinic sometimes discussed estrogen or hormone receptor status without specifying the progesterone receptor status. We believe this may explain the discrepancy between patients' ability to self-report their estrogen and progesterone receptor status. Given the emergence of triple-negative breast cancer as a distinct clinical entity, designers of clinical trial matching systems should include progesterone receptor status in their matching criteria and prompt patients to ask their physicians about it if necessary.

Of 20 patients, 5 (25%) were not accurate about their tumor response to prior treatment. In the adjuvant setting where there is no longer evidence of disease, it is easy to understand patients' confusion with this question, as the possible answers focused on how the "cancer" responded to treatment. As we disseminated the website nationally, we recast the question. On all forms, we ask whether therapy was completed and, if not, why it was stopped. The choices for why it was stopped were "Stopped treatment due to side effects," "Tumor occurred, recurred, or did not shrink with therapy," and "I don't' know." We standardized the question to fit across patient types and therapies.

We learned that, while 18 of 20 respondents matched their chart regarding stage at diagnosis, only 15 matched the chart when responding to "What stage of breast cancer do you have as of today?" One patient said in an unsolicited phone call to the BreastCancerTrials.org offices, "I've completed treatment. I don't have cancer now." We have since deleted this question from the BreastCancerTrials.org online health history as being confusing for people whose breast cancer tumor has been removed by local therapy.

Connections to the Literature

We now turn to a discussion of how our findings complement or conflict with previous reports in the published literature.

With regard to our first study question, on adoption by research sites, we have not found published evidence regarding whether multiple research sites will list their trials in a common clinical trial matching website. Our positive finding extends the results of a study of clinical trial matching at a single location [18]. In that study, the University of Pennsylvania collaborated with EmergingMed to provide clinical trial matching for trials open at the University of Pennsylvania's cancer center.

Regarding acceptability to patients, we did not know whether privacy issues or complexity would inhibit patients from providing the required health history data to BreastCancerTrials.org. A previously published usability study of BreastCancerTrials.org suggested that both could be barriers and tested design changes that addressed user concerns [20,21]. Our finding that 428 patients over 14 months reported their health history is consistent with the University of Pennsylvania's experience, where they found that 627 patients provided sufficient information to match to clinical trials over a 15-month period [18]. Whereas the University of Pennsylvania system featured trials from one location across many cancer types, our system featured breast cancer trials only, but from many research site locations.

Regarding patient accuracy, we found 91% overall accuracy about details of the breast cancer diagnosis in a sample of patients undergoing breast cancer treatment at our medical center. No studies have reported patients' accuracy for the other clinical trial matching systems. Other studies have found high overall accuracy of patients' self-reports regarding the existence of a prior breast cancer and its treatment [22]. However, some patient subgroups are less accurate [22]. Also, patients may be less accurate in recollecting some aspects of their breast cancer history such as the stage of their cancer [23].

The American Society of Clinical Oncology has recommended that all patients be provided with a treatment summary as part of standard care for survivors [24]. This kind of summary would provide women with the data required for accurate clinical trial matching. In the future, the ability to capture these data in an electronic fashion would potentially enable an automated way for patients and providers to identify appropriate trials as a routine of practice.

Limitations

This evaluation did not address some of the known threats to internal and external validity. For example, the proportion of research sites willing to provide their trial data may be different from a nationally representative sample.

In studying acceptability to patients, we invited only individuals who completed our health history to take the survey. With a response rate of 20%, we don't know whether the majority of nonrespondents felt differently about their experience with BreastCancerTrials.org. In particular, although we learned that 5 respondents enrolled in a clinical trial, the overall enrollment rate among BreastCancerTrials.org users may have been higher or lower depending on whether respondents differed from nonrespondents. Sensitivity analysis suggests the enrollment rate may have varied from 1% (5/375 surveyed) to 7% (5/75 responding). In addition, our survey solicited only anonymous responses, so we could not follow up on responses. Thus, we learned that 12 of 23 people who contacted a research site were eligible for a study, but could not ascertain why the remaining 11 were ineligible.

We conducted the accuracy evaluation with a small sample of patients undergoing treatment at our academic medical center, not with website registrants. Our accuracy results may not reflect the accuracy of actual visitors to the website. We observed a 46% consent rate and a 77% completion rate for an overall 35% response rate relative to the patients we approached. This means that the majority of patients we approached declined to complete the health history questionnaire, limiting our ability to generalize our accuracy results. The accuracy results could also be misleading in that we treated the chart as being definitively correct. In fact, charts may be incomplete or erroneous.

For all of the study questions, our samples were highly educated relative to the general population. The clinic sample in the accuracy study was even better educated than the website registrants overall or than website registrants responding to our follow-up acceptability survey. This may limit the generalizability of our findings, as studies have shown that low-income and elderly populations face challenges to using electronic health tools [25].

Conclusions

We conclude that BreastCancerTrials.org is a promising vehicle for connecting patients with opportunities to participate in clinical trials. Research sites were willing to register their trials, and over half of the patients who found BreastCancerTrials.org registered and completed a health history questionnaire. Patients also contacted research sites and some joined trials. The degree of accuracy in patients' responses was high, which may be due to the fact that all of the patients had completed a 4-year college degree and 14 out of 20 (70%) reported postgraduate training. Still, based on our analysis of response patterns, we have been able to improve the way in which health history questions are asked. Researchers should continue to explore the question of accuracy with more diverse populations. People with lower health literacy may require assistance in providing accurate information to BreastCancerTrials.org or similar clinical trial matching tools.

Our study contributes new knowledge to the literature, including evidence that research sites will provide details of their open trials to third-party clinical trial matching systems; and patients can and will accurately fill out a detailed health history questionnaire as a means to matching to clinical trials online. Overall, this evaluation suggests that BreastCancerTrials.org could contribute to the public health agenda of making people more aware of clinical trials with a view to ultimately increasing enrollment.

As a result of this study, we secured funding to disseminate BreastCancerTrials.org so that it includes research sites across the United States. This is a reportable outcome in terms of implementation and maintenance of a research-initiated innovation. The nationwide service operates at www.BreastCancerTrials.org as a program of Quantum Leap Healthcare Collaborative, a nonprofit corporation [26]. Quantum Leap has successfully solicited philanthropic donations to maintain the clinical trial matching service and is exploring additional sources of recurring revenue to sustain it.

In preparing for the nationwide launch, the BreastCancerTrials.org team made improvements to the software based on lessons from this study. New features included the ability of users to select a health history form customized to their situation (having a new diagnosis vs metastatic vs posttreatment survivor); to use BreastCancerTrials.org as an anonymous guest; to register for a Trial Alert Service; and to browse through BreastCancerTrials.org listings without creating a profile. The trial coding interface and matching engine (caMatch) now broker more complex eligibility criteria. Although trials are still entered by BreastCancerTrials.org staff, the trial coding process is now more efficient. The nationwide service launched in October 2008 [27] and currently lists over 500 trials. Patients who wish to use this or other clinical trial matching websites should keep copies of their medical records so that they can enter a profile that will facilitate their gaining access to appropriate opportunities to participate in research studies. Initiatives aimed at providing patients with treatment summaries, especially if integrated with electronic health records, will make it easier for patients to use online resources such as BreastCancerTrials.org.

Acknowledgments

This work was supported by the California Breast Cancer Research Program (9DB-0070), the Department of the Army (DAMD 17-03-1-10481), The Cancer League, and Mount Zion Health Systems.

We very much appreciate the contributions of the following individuals: Charles Griffen, JJ Mauer, and Satish Patel, from Ekagra Software; and Vinona Bhatia, Michael Kamerick, Michelle Melisko, Margot Pass, Donny Tjandro, Weimin Shen, and Stephen Wong from UCSF. We also express our deep gratitude to Sue Dubman and Ken Buetow from the National Cancer Institute Center for Biomedical Informatics and Information Technology, whose guidance and support made this project possible. Lastly, we wish to thank the patients, investigators, and clinical research associates who participated in this study.

Dr Deering is now with the Office of the National Coordinator for Health Information Technology, US Department of Health and Human Services; Julia Pederson is attending Stanford University School of Medicine, Stanford, CA; Alexandra Teng and Erin McCreary are attending UCSF School of Medicine, San Francisco, CA; Dr Tripathy is now with the Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; Ms Schreiner is deceased.

Conflicts of Interest

None declared.

References

1. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med 2002 Oct 17;347(16):1233-1241 [FREE Full text] [doi: 10.1056/NEJMoa022152] [Medline: 12393820]
2. Friedman MA, Cain DF. National Cancer Institute sponsored cooperative clinical trials. Cancer 1990 May 15;65(10 Suppl):2376-2382. [Medline: 2334876]
3. Mills EJ, Seely D, Rachlis B, Griffith L, Wu P, Wilson K, et al. Barriers to participation in clinical trials of cancer: a meta-analysis and systematic review of patient-reported factors. Lancet Oncol 2006 Feb;7(2):141-148. [doi: 10.1016/S1470-2045(06)70576-9] [Medline: 16455478]
4. Sung NS, Crowley WF, Genel M, Salber P, Sandy L, Sherwood LM, et al. Central challenges facing the national clinical research enterprise. JAMA 2003 Mar 12;289(10):1278-1287. [Medline: 12633190]
5. Comis RL, Miller JD, Colaizzi DD, Kimmel LG. Physician-related factors involved in patient decisions to enroll onto cancer clinical trials. J Oncol Pract 2009 Mar;5(2):50-56. [doi: 10.1200/JOP.0922001] [Medline: 20856718]
6. Lara PN, Higdon R, Lim N, Kwan K, Tanaka M, Lau DH, et al. Prospective evaluation of cancer clinical trial accrual patterns: identifying potential barriers to enrollment. J Clin Oncol 2001 Mar 15;19(6):1728-1733. [Medline: 11251003]
7. Fink E, Kokku PK, Nikiforou S, Hall LO, Goldgof DB, Krischer JP. Selection of patients for clinical trials: an interactive web-based system. Artif Intell Med 2004 Jul;31(3):241-254. [doi: 10.1016/j.artmed.2004.01.017] [Medline: 15302090]
8. Ash N, Ogunyemi O, Zeng Q, Ohno-Machado L. Finding appropriate clinical trials: evaluating encoded eligibility criteria with incomplete data. Proc AMIA Symp 2001:27-31. [Medline: 11825151]
9. Ohno-Machado L, Wang SJ, Mar P, Boxwala AA. Decision support for clinical trial eligibility determination in breast cancer. Proc AMIA Symp 1999:340-344. [Medline: 10566377]
10. Breitfeld PP, Weisburd M, Overhage JM, Sledge G, Tierney WM. Pilot study of a point-of-use decision support tool for cancer clinical trials eligibility. J Am Med Inform Assoc 1999;6(6):466-477 [FREE Full text] [Medline: 10579605]
11. Breitfeld PP, Ullrich F, Anderson J, Crist WM. Web-based decision support for clinical trial eligibility determination in an international clinical trials network. Control Clin Trials 2003 Dec;24(6):702-710. [Medline: 14662275]
12. Fox S. Online health search 2006. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. URL: http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf [accessed 2012-06-07] [WebCite Cache ID 5xnSUoYXn]

13. Atkinson NL, Saperstein SL, Massett HA, Leonard CR, Grama L, Manrow R. Using the Internet to search for cancer clinical trials: a comparative audit of clinical trial search tools. Contemp Clin Trials 2008 Jul;29(4):555-564. [doi: 10.1016/j.cct.2008.01.007] [Medline: 18346942]
14. US National Institutes of Health. ClinicalTrials.gov URL: http://www.clinicaltrials.gov/ [accessed 2011-10-25] [WebCite Cache ID 62hbO18Y1]
15. National Cancer Institute, US National Institutes of Health. Clinical Trials Search URL: http://www.cancer.gov/clinicaltrials/search [accessed 2011-10-25] [WebCite Cache ID 62hbldK1V]
16. Coalition of Cancer Cooperative Groups. CancerTrialsHelp.org URL: http://www.cancertrialshelp.org/ [accessed 2011-10-25] [WebCite Cache ID 62hbpO58B]
17. EmergingMed. 2011. Find Cancer Clinical Trial Options URL: http://www.emergingmed.com/ [accessed 2011-10-25] [WebCite Cache ID 62hbrXYXt]
18. Metz JM, Coyle C, Hudson C, Hampshire M. An Internet-based cancer clinical trials matching resource. J Med Internet Res 2005 Jul 1;7(3):e24 [FREE Full text] [doi: 10.2196/jmir.7.3.e24] [Medline: 15998615]
19. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982;5(3):233-237. [Medline: 10259963]
20. Atkinson NL, Massett HA, Mylks C, Hanna B, Deering MJ, Hesse BW. User-centered research on breast cancer patient needs and preferences of an Internet-based clinical trial matching system. J Med Internet Res 2007;9(2):e13 [FREE Full text] [doi: 10.2196/jmir.9.2.e13] [Medline: 17513284]
21. Atkinson NL, Massett HA, Mylks C, McCormack LA, Kish-Doto J, Hesse BW, et al. Assessing the impact of user-centered research on a clinical trial eHealth tool via counterbalanced research design. J Am Med Inform Assoc 2011;18(1):24-31 [FREE Full text] [doi: 10.1136/jamia.2010.006122] [Medline: 21169619]
22. Liu Y, Diamant AL, Thind A, Maly RC. Validity of self-reports of breast cancer treatment in low-income, medically underserved women with breast cancer. Breast Cancer Res Treat 2010 Feb;119(3):745-751. [doi: 10.1007/s10549-009-0447-5] [Medline: 19551500]
23. Phillips KA, Milne RL, Buys S, Friedlander ML, Ward JH, McCredie MR, et al. Agreement between self-reported breast cancer treatment and medical records in a population-based Breast Cancer Family Registry. J Clin Oncol 2005 Jul 20;23(21):4679-4686. [doi: 10.1200/JCO.2005.03.002] [Medline: 15851764]
24. Ganz PA, Hahn EE. Implementing a survivorship care plan for patients with breast cancer. J Clin Oncol 2008 Feb 10;26(5):759-767. [doi: 10.1200/JCO.2007.14.2851] [Medline: 18258984]
25. 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]
26. BreastCancerTrials.org. 2011. Find a Trial That's Right for You URL: https://www.breastcancertrials.org/bct_nation/;jsessionid=D697799A8A8826908706519BB41F3772 [accessed 2011-10-25] [WebCite Cache ID 62hbu9tRz]
27. Cohen EJ, Belkora JK, McCreary E, Colen S, Maple T, Bechtold T, et al. BreastCancerTrials.org: Early evaluation of a nationwide clinical trial matching service. MA: org: early evaluation of a nationwide clinical trial matching service. conf-loc:; 2010 Presented at: Thirty-Third Annual CTRC-AACR San Antonio Breast Cancer Symposium; Dec 8-12, 2010; San Antonio, TX, USA URL: http://www.posters2view.com/sabcs10/viewp.php?nu=P5-07-01

Abbreviations

HER2/neu: human epidermal growth factor receptor 2
USCF: University of California, San Francisco

Edited by G Eysenbach; submitted 02.06.11; peer-reviewed by J Marquard, S Saperstein; comments to author 13.09.11; revised version received 27.10.11; accepted 07.04.12; published 11.07.12.

Please cite as:
Cohen E, Belkora J, Tyler J, Schreiner J, Deering MJ, Grama L, Duggan B, Illi J, Pederson J, Anand A, Teng A, McCreary E, Moore D, Tripathy D, Hogarth M, Lieberman M, Park J, Esserman L
Adoption, Acceptability, and Accuracy of an Online Clinical Trial Matching Website for Breast Cancer
J Med Internet Res 2012;14(4):e97
URL: http://www.jmir.org/2012/4/e97/
doi:10.2196/jmir.1855
PMID:22784878

Original Paper

Impact of the Presence of Medical Equipment in Images on Viewers' Perceptions of the Trustworthiness of an Individual On-Screen

Moyez Jiwa[1], MBChB, MD, MMedSc, MRCGP, FRACGP; Stephan Millett[2], BEcon, BA(Hons), PhD, MACE; Xingqiong Meng[3], PhD, BMed, MAppE; Vivien M Hewitt[4], BSc(Hons), GradDipLib, GradDipForSci

[1]Curtin Health Innovation Research Institute, Curtin University, Perth, Western Australia, Australia
[2]Research & Development, Administration & Finance, Curtin University, Perth, Western Australia, Australia
[3]School of Public Health, Curtin University, Perth, Western Australia, Australia
[4]Fremantle Hospital, Fremantle, Western Australia, Australia

Corresponding Author:
Moyez Jiwa, MBChB, MD, MMedSc, MRCGP, FRACGP
Curtin Health Innovation Research Institute
Curtin University
GPO Box U1987
Technology Park
Perth, Western Australia, 6845
Australia
Phone: 61 61892661768
Fax: 61 61892662608
Email: m.jiwa@curtin.edu.au

Abstract

Background: It is now common practice for doctors to consult patients by means other than face-to-face, often appearing before the patient on a computer screen. Also, many websites are using depictions of health professionals to increase the credibility of their services. Being trustworthy is an essential attribute for successful ehealth services. Little is known about which depicted accessories make a health professional appear more trustworthy.

Objective: To estimate the odds of an individual on-screen being rated trustworthy when viewed in a static image holding or wearing specific items of medical equipment.

Methods: We surveyed consecutive people attending community pharmacies to collect prescriptions in Western Australia. Respondents were presented with a series of 10 photographs, generated at random, of a man with varying numbers and combinations of medical equipment: stethoscope, reflex hammer, surgical scrubs, otoscope, and pen. They were then invited to rate the man as honest, trustworthy, honorable, moral, ethical, or genuine, or a combination of these, on the Source Credibility Scale.

Results: A total of 168 of 250 people gave informed consent, for a participation rate of 67.2%. There were 102 female and 66 male respondents. Of the 168 respondents, 96 (57%) were born in Australia and 102 (60.7%) were attending medical practices with more than one general practitioner. The mean age of respondents was 47 (SD 16) years (range 26–92 years). When only 1 item was present in an image, the stethoscope was associated with the highest odds for the person being considered honest (odds ratio [OR] 2.6, 95% confidence interval [CI] 1.6–4.3), trustworthy (OR 2.3, 95% CI 1.4–3.8), honorable (OR 2.7, 95% CI 1.6–4.5), moral (OR 2.4 95% CI 1.4–4.1), ethical (OR 2.6, 95% CI 1.5–4.6), and genuine (OR 1.8, 95% CI 1.0–3.1). The presence of a stethoscope increased the odds of the person being rated in a positive light in all photographs in which it was included.

Conclusions: When an individual is portrayed in a static image, concurrent presentation of 3 or more items of medical equipment, and especially a stethoscope, is likely to exert a positive influence on the viewers' perceptions of the qualities of the person depicted.

(J Med Internet Res 2012;14(4):e100) doi:10.2196/jmir.1986

KEYWORDS

Icons; semiotics; stethoscope; doctors; trustworthiness

Introduction

When people are concerned about their physical or psychological well-being, they may consult a doctor. When a patient consults a doctor face-to-face, all five senses affect the experience: sight, hearing, touch, smell (eg, medications, disinfectants, and tissues), and sometimes taste (eg, medications, and equipment used to examine the mouth or throat) [1-5]. In consultations where the doctor and patient are not even in the same room, there are limits to the extent to which all the senses can be engaged [6]. It is important that a doctor (or online ehealth service or website) be perceived as trustworthy, especially in circumstances in which there are limits to the engagement of the senses as occurs, for example, when people interact with doctors via a computer or when a doctor is portrayed in a static image in support of a message on a website.

When someone is experiencing symptoms, the condition does not require the presence of objective organic pathological analysis. Indeed, some people experience relief from symptoms despite being treated with a placebo or inactive drug [7,8]. It has been demonstrated in conditions that are not the result of organic pathology that a doctor has powers to relieve disease in a manner analogous to that of a placebo [9]. In common with many placebos, the physical appearance of the doctor may account for some of the therapeutic response. As McCroskey stated, “No message is received independently of its source or presumed source” [10]. Various writers have agreed that there is a dimension to the perception of an individual that can be referenced as trustworthiness (character, sagacity, safety, and honesty) [11].

The literature records many studies in which credibility and persuasive communication have been tested with specific reference to the physical appearance of doctors. In an experiment conducted by computer scientists, advice was shown to be much more persuasive when presented as coming from doctors as depicted in static pictures [12]. Similarly, previously published studies have reported that patients prefer traditional physicians’ attire consisting of a white coat and professional dress [13]; other studies indicate that patients are equally satisfied with their physicians regardless of casual or business attire [14]. In a more recent experiment, researchers have concluded that doctors working with a primarily older population may find their patients prefer their physicians to wear white coats, whereas a large majority of parents do not expect their child's doctor to wear a white coat and that the a parent’s trust is not compromised by less-traditional physicians’ attire [15].

In addition to what the doctor is wearing, we must also consider what he or she is seen to be holding, or using. For example, the stethoscope has been the cornerstone of medical diagnostics for nearly 200 years. This monaural device improved physicians’ ability to hear clues to their patients’ underlying pathology and thus apply the appropriate treatment in many cases. It has been eulogized by many commentators, such as the following:

> *I contend that the stethoscope best symbolizes the practice of medicine. Whether absentmindedly worn around the neck like an amulet or coiled gunslinger-style in the pocket, ever ready for the quick draw, the stethoscope is much more than a tool that allows us to eavesdrop on the workings of the body. Indeed, it embodies the essence of doctoring: using science and technology in concert with the human skill of listening to determine what ails a patient.* [16]

The stethoscope is one of several medical instruments that are recognized by laypeople as belonging in a doctor’s office and that have significance in the consultation that extends beyond the instrument’s physical functionality. Similarly, reflex hammers and other items used in clinical examination have developed an iconic status for doctors [17]. In this study we hypothesized that in the absence of any other information, an individual appearing in a static image may be rated as more trustworthy when he is viewed holding or wearing specific items of medical equipment. Further, we hypothesized that some items of medical equipment would have a greater influence than others on the perception of trustworthiness.

Methods

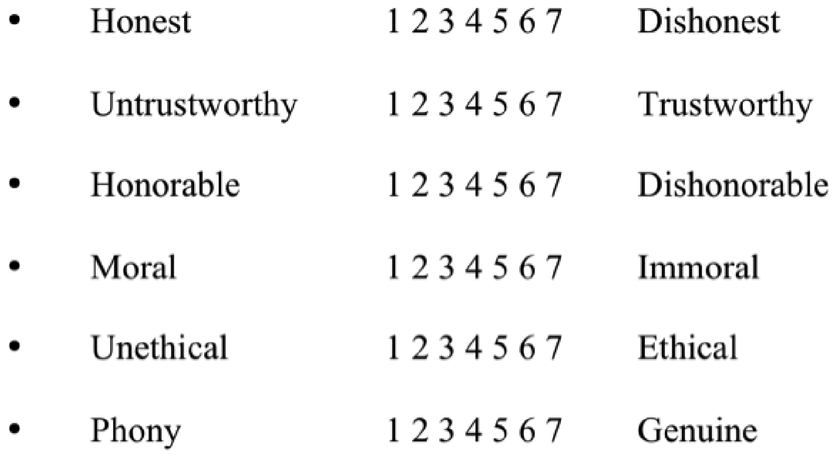
We obtained human research ethics approval from Curtin University (approval number RD-23-10) before commencing the study. All respondents provided consent before data collection. Consecutive people waiting in six community pharmacies to collect prescriptions for drugs prescribed by a doctor were presented with a series of photographs on an iPad. Each photograph presented an image of a man with a neutral facial expression who was wearing a casual shirt. Some of the images contained medical equipment. The respondents were not told anything about the person in the photograph. The photographs were presented as a series of 10 consecutive images containing an increasing number of items of medical equipment ranging from 0 to 5. Each series of 10 was unique to each participant and drawn at random from a pool of images. The equipment (collectively called icons here) consisted of a stethoscope, otoscope, reflex hammer, surgical scrubs, and pen. We chose these icons from the list of items that have been recommended as necessary in a general practitioner’s bag [18]. The icons were presented in increasing numbers and in the following order, with the collection of icons chosen at random: (1) no icons (1 photo, shown first), (2) 1 icon (2 photos, each containing 1 icon selected at random), (3) 2 icons (2 photos, each containing 2 icons selected at random), (4) 3 icons (2 photos, each containing 3 icons selected at random), (5) 4 icons (2 photos, each containing 4 icons selected at random), and (6) 5 icons (1 photo, shown last).

Figure 1 and Multimedia Appendix 1, Multimedia Appendix 2, Multimedia Appendix 3, Multimedia Appendix 4, Multimedia Appendix 5, and Multimedia Appendix 6 illustrate examples of the photos.

The participants were invited to rate the individual using the trustworthiness measure of the Source Credibility Scale (SCS) [11]. The SCS is an 18-item survey comprising three separate measures that capture respondents’ perceptions of an individual’s “competence, trustworthiness, and goodwill/caring.” The three measures in the SCS each represent a unique construct, the scores for which must be considered in isolation from the others. Each individual measure may validly be subject to regression analysis [11]. For this study and to keep the required sample size to a minimum, we invited the participants to respond only to the six questions from the SCS that make up the trustworthiness measure [11]. The internal consistency (alpha reliability) of the trustworthiness measure is .93. Each element was scored from 1 to 7 at either end of the scale, as shown in Figure 2.

Figure 1. Example of an image shown to the respondents.

Figure 2. Elements of the Source Credibility Scale.



•	Honest	1 2 3 4 5 6 7	Dishonest
•	Untrustworthy	1 2 3 4 5 6 7	Trustworthy
•	Honorable	1 2 3 4 5 6 7	Dishonorable
•	Moral	1 2 3 4 5 6 7	Immoral
•	Unethical	1 2 3 4 5 6 7	Ethical
•	Phony	1 2 3 4 5 6 7	Genuine

Sample Size

For logistic regression, to estimate the odds of a specific outcome within 80% confidence intervals at 5% significance, in this case scoring an image as trustworthy and modeling 11 explanatory variables, we required 138 respondents based on a moderate effect size [11,19].

Statistical Analysis

We used ordinal logistic regression to examine the influence on primary outcomes of icons presented in the photos. The influence was calculated as the odds of a person being rated honest, trustworthy, honorable, moral, ethical, or genuine, or a combination of these, after adjusting the respondents' characteristics (eg, sex, age group, and country of birth). Five predictor variables (the icons, ie, the stethoscope, otoscope, reflex hammer, surgical scrubs, and pen) were assessed after controlling for respondents' demographics and experience of consulting doctors. This consulting experience included the type of medical practice frequented (solo practitioner or group practice), number of general practitioner visits, and number of hospital visits.

Regression analysis was performed to assess the effect of the 5 icons and the different combinations of icons on rating for honest, trustworthy, etc. A variable that defined the 5 icons (model 1), the different combination of icons (2 icons, 3 icons, 4 icons: models 2–4), and all 5 icons (model 5) were used in the corresponding models. In each model, the group of vignettes (photographs) with no icon was treated as the reference group, and the groups of vignettes with 1, 2, 3, 4, or 5 icons were compared with the reference group respectively in the regression model. The lack of independence between individual participants, which causes intragroup correlation, was adjusted in the models through estimating the clustered robust standard errors. Statistical analysis was performed using Stata Statistical Software (IC 11.1; State Corp LP, College Station, TX, USA).

Results

We invited 250 people to participate, and 168 completed the survey, giving a participation rate of 67.2%. Demographic details are presented in Table 1. There were more female than male respondents, and more respondents attending medical practices with more than one general practitioner. Most respondents were born in Australia, and the mean age of respondents was 47 (SD 16) years (range 26–92 years).

Table 1. Demographic characteristics of respondents (n = 168).

Characteristic	n	%	*P* value[a]
Age (years)			.64
≤45	81	48%	
>45	87	52%	
Sex			.006
Male	66	39%	
Female	102	60.7%	
Country of birth			.06
Australia	96	57%	
Other	72	43%	
Type of medical practice frequented			.006
Clinic with more than one general practitioner	102	60.7%	
Solo practitioner	66	39%	

[a] *P* values were derived from 2-sample test of proportion.

Table 2, Table 3, and Table 4 show that the stethoscope was more likely to be present in images where the person in the photograph was regarded as one or more of honest, trustworthy, honorable, moral, ethical, or genuine. This was particularly true in images that contained only one piece of equipment. However, the stethoscope was also more likely to be present in images credited with any of these characteristics when multiple pieces of equipment were presented.

Table 2. Data for person in photograph being perceived as honest and trustworthy when 1 or more icons[a] are present versus no icons.

Model	Honest		Trustworthy	
	OR[b]	95% CI[c]	OR	95% CI
1(1 icon)				
S	2.6	1.6–4.3***	2.3	1.4–3.8***
O	1.6	1.1–2.4*	1.4	0.9–2.3
RH	1.5	0.9–2.5	1.6	1.0–2.5*
SS	1.3	0.8–2.1	1.2	0.8–1.9
P	1.3	0.8–2.1	1.2	0.8–2.0
No icon[d]	1.0		1.0	
2 (2 icons)				
S+RH	3.9	2.1–7.4***	2.3	1.1–4.6*
S+SS	2.5	1.5–4.3***	1.3	0.7–2.5
S+P	2.5	1.3–4.8**	2.7	1.5–4.8***
O+RH	2.9	1.7–4.9***	1.8	0.9–3.4
O+P	1.8	0.8–3.7	2.3	1.2–4.3*
RH+SS	2.5	1.3–4.8**	1.4	0.6–3.1
RH+P	1.8	0.9–3.4	1.8	1.0–3.2[e]
SS+P	1.7	0.9–3.4	1.9	1.0–3.5[e]
O+P	2.4	1.0–5.5*	4.0	2.3–7.0***
No icon[d]	1.0		1.0	
3 (3 icons)				
S+O+RH	3.0	1.6–5.7***	3.1	1.6–5.8***
S+O+RH	4.0	2.3–6.8***	2.3	1.3–4.2**
S+O+P	4.0	2.1–7.6***	3.9	2.0–7.5***
S+RH+SS	2.6	1.3–5.1**	2.8	1.5–5.2***
S+RH+P	2.3	1.0–5.1*	2.2	1.1–4.5*
S+SS+P	3.4	1.8–6.4***	2.7	1.4–5.4**
O+SS+P	1.8	0.7–4.5	1.7	0.5–5.3
No icon[d]	1.0		1.0	
4 (4 icons)				
S+O+RH+SS	4.4	2.5–7.7***	2.6	1.4–4.9**
S+O+RH +P	4.0	2.5–6.4***	3.0	1.9–4.8***
S+O+SS+P	3.7	2.2–6.1***	3.6	2.3–5.7***
S+RH+SS+P	2.1	1.3–3.4**	2.8	1.8–4.6***
O+RH+SS+P	3.4	1.9–6.2***	1.6	0.8–3.2
No icon[d]	1.0		1.0	
5 (5 icons)	3.7	2.4–5.5***	3.7	2.5–5.5***
No icon[d]	1.0		1.0	

[a] Icons are as follows: S = stethoscope, O = otoscope, RH = reflex hammer, SS = surgical scrubs,*P* = pen.

[b] Odds ratio.

[c] Confidence interval.

[d] Photo with no icon group was the reference group for each of the models.

[e] 95% CI was rounded to 1.00 but the *P* value was >.05.
*$P < .05$, **$P < .01$, ***$P < .001$.

Table 3. Data for person in photograph being perceived as honorable and moral when 1 or more icons[a] are present versus no icons.

Model	Honorable		Moral	
	OR[b]	95% CI[c]	OR	95% CI
1(1 icon)				
S	2.7	1.6–4.5	2.4	1.4–4.1**
O	1.3	0.8–2.1	1.3	0.8–2.2
RH	1.6	1.0–2.6[e]	1.7	1.1–2.6*
SS	1.3	0.8–2.2	1.4	0.9–2.3
P	1.7	1.0–2.6*	1.1	0.7–1.8
No icon[d]	1.0		1.0	
2 (2 icons)				
S+RH	2.2	1.1–4.5*	2.1	1.1–4.1*
S+SS	2.5	1.5–4.2***	2.0	1.2–3.4**
S+P	1.7	0.9–3.3	2.2	1.3–4.0**
O+RH	2.4	1.3–4.5**	2.4	1.4–4.3**
O+P	1.9	0.9–3.7	1.5	0.7–3.5
RH+SS	2.3	1.1–4.7*	2.0	1.0–4.0[e]
RH+P	2.7	1.3–5.6**	2.7	1.4–5.0**
SS+P	2.7	1.4–5.1**	2.5	1.4–4.5**
O+P	1.7	0.7–4.0	1.9	0.8–4.3
No icon[d]	1.0		1.0	
3 (3 icons)				
S+O+RH	3.5	1.9–6.6***	3.3	1.8–6.1***
S+O+SS	3.0	1.7–5.5***	2.9	1.6–5.1***
S+O+P	2.4	1.2–5.1*	2.3	1.1–5.1*
S+RH+SS	3.0	1.5–6.0**	2.5	1.3–5.0**
S+RH+P	2.4	1.1–5.1*	2.1	1.0–4.3*
S+SS+P	4.2	2.3–7.6***	3.9	2.2–7.1***
O+SS+P	2.3	1.1–5.1*	2.0	1.0–4.0[e]
No icon[d]	1.0		1.0	
4 (4 icons)				
S+O+RH+SS	3.7	2.0–6.9***	3.6	2.0–6.5***
S+O+RH+P	3.3	2.0–5.5***	3.0	1.9–4.8***
S+O+SS+P	2.6	1.6–4.3***	2.5	1.5–4.2***
S+RH+SS+P	2.4	1.4–3.9***	1.8	1.1–3.1*
O+RH+SS+P	3.6	2.0–6.4***	3.4	2.1–5.7***
No icon[d]	1.0		1.0	
5 (5 icons)	3.1	2.1–4.6***	3.1	2.2–4.4***
No icon[d]	1.0		1.0	

[a] Icons are as follows: S = stethoscope, O = otoscope, RH = reflex hammer, SS = surgical scrubs,*P* = pen.

[b] Odds ratio.

[c] Confidence interval.

[d] Photo with no icon group was the reference group for each of the models.

[e] 95% CI was rounded to 1.00 but the *P* value was >.05.

*$P < .05$, **$P < .01$, ***$P < .001$.

Table 4. Data for person in photograph being perceived as ethical and genuine when 1 or more icons[a] are present versus no icons.

Model	Ethical		Genuine	
	OR[b]	95% CI[c]	OR	95% CI
1(1 icon)				
S	2.6	1.5–4.6***	1.8	1.0–3.1*
O	1.7	1.0–2.8[e]	1.4	0.9–2.3
RH	1.5	0.9–2.6	1.7	1.1–2.6*
SS	1.2	0.7–2.1	1.3	0.8–2.0
P	1.3	0.8–2.1	0.9	0.5–1.5
No icon[d]	1.0		1.0	
2 (2 icons)				
S+RH	2.8	1.4–5.7**	2.1	1.1–4.0*
S+SS	1.6	0.9–3.1	1.3	0.7–2.4
S+P	2.8	1.5–5.1***	2.5	1.4–4.1***
O+RH	2.1	1.1–4.3*	1.5	0.7–3.2
O+P	2.8	1.4–5.5**	2.3	1.2–4.3*
RH+SS	2.9	1.4–5.7**	2.7	1.4–5.2**
RH+P	2.6	1.4–4.5***	1.8	1.0–3.2*
SS+P	2.8	1.4–5.5**	2.0	1.0–3.9[e]
O+P	3.7	2.1–6.7***	3.5	1.6–7.6***
No icon[d]	1.0		1.0	
3 (3 icons)				
S+O+RH	3.3	1.8–5.9***	2.4	1.3–4.1**
S+O+SS	2.6	1.4–5.0**	2.1	1.1–3.8*
S+O+P	3.4	1.8–6.4***	3.1	1.6–5.7***
S+RH+SS	3.2	1.6–6.3***	2.9	1.6–5.3***
S+RH+P	2.1	0.9–4.6	1.7	0.7–4.1
S+SS+P	3.4	1.8–6.7***	2.5	1.2–5.4*
O+SS+P	2.3	1.0–5.4[e]	1.5	0.7–3.3
No icon[d]	1.0		1.0	
4 (4 icons)				
S+O+RH+SS	3.5	2.0–6.2***	2.9	1.6–5.1***
S+O+RH+P	2.9	1.7–5.0***	2.6	1.6–4.3***
S+O+SS+P	3.1	1.9–5.0***	2.6	1.6–4.3***
S+RH+SS+P	2.9	1.8–4.8***	2.5	1.6–3.9***
O+RH+SS+P	2.4	1.3–4.5**	1.8	1.0–3.5[e]
No icon[d]	1.0		1.0	
5 (5 icons)	4.1	2.8–5.9***	3.1	2.2–4.5***
No icon[d]	1.0		1.0	

[a] Icons are as follows: S = stethoscope, O = otoscope, RH = reflex hammer, SS = surgical scrubs,*P* = pen.

[b] Odds ratio.

[c] Confidence interval.

[d] Photo with no icon group was the reference group for each of the models.

[e] 95% CI was rounded to 1.00 but the *P* value was >.05.

**P* < .05, ***P* < .01, ****P* < .001.

Discussion

When only 1 icon was presented in the photograph, the stethoscope was associated with the highest odds for the person being considered honest, trustworthy, moral, honorable, ethical, or genuine, or a combination of these. The stethoscope evoked strong positive perceptions. There were no differences in scores on any characteristics for any combination of 2 icons. However, photographs with 3 or more icons had significantly higher scores for all characteristics than photographs with 0 or 1 icon. Literature on the impact of prominently displaying icons in medical consulting rooms concluded that they may determine the extent to which a doctor is perceived to be open to the ideas, concerns, and expectations of the patient in the consultation [16]. Similarly, our data suggest that items of medical equipment may influence the perceptions patients have of a person in a photograph.

Strengths and Limitations

The same images, albeit a limited random selection, were shown to all respondents. This is a possible limitation because the demand characteristics placed on the participants may have been enough to explain the increasing trustworthiness ratings with increasing number of medical devices in the photographs. The participants may have worked out that their ratings were supposed to be based on the changing elements of the pictures and that more devices should indicate greater trustworthiness. However, the relative trustworthiness ratings of types of devices might be free from demand characteristics, and we were able to estimate this from the data. The demographic characteristics of the respondents were similar to those of patients who generally visit general practitioners in Australia, and most were women over 45 years of age [20]. The data suggest that including certain icons in the static image of medical practitioners can improve the patient's perception of the doctor's trustworthiness. On the other hand, there were a number of limitations. We assumed everyone would recognize all items of medical equipment and the surgical scrubs. This could not be confirmed and informal feedback to the researchers conducting the interviews suggested that some of the respondents did not recognize the surgical scrubs or the reflex hammer in particular. Some items of medical equipment are no longer used exclusively by doctors and are now also commonly used by other health care professionals including nurses, physiotherapists, and paramedical staff. The data may not be generalizable because respondents were mostly female, with none aged younger than 18 years. All participants, or their relatives, had consulted a doctor recently and all were attending a community pharmacy. There were other signs in this study that we did not take into consideration and that respondents would have interpreted through community-accepted codes. These included facial expression, color of the person's shirt, his race (white European), and his gender. These signs would have been likely to affect respondents' assessment of trustworthiness. In this study, we did not test the icons with video footage—the pictures were static images and not animated as they would appear in an online consultation with a doctor. Finally, we also acknowledge the importance of context (which both patient and professional will bring to the consultation) with additional cues (identifying the professional as a doctor), which may either enhance or diminish the effect of the icon.

Future Research

In this study we introduced participants to a series of photographs of a person along with a variety of items associated with doctors. These items were used because they are commonly recognized symbols or icons of medical practice that patients use to make meaning of a medical consultation. In Peircean semiotics, each of the items we employed is more properly an index because the item is connected directly in some way to the functions we ascribe to doctors, and hence the signifier is not entirely arbitrary [21]. It could be argued that other symbols, not usually seen in medical practice, may also be associated with trustworthiness, but that does not necessarily detract from the finding that a perception of trustworthiness is correlated with the presence of certain items used in medical consultations. It is possible that this study has further established what is already known: stethoscopes in particular are associated with medicine, and doctors are considered trustworthy. At the least, however, this study demonstrates additionally that the presence of certain icons is strongly correlated with perceived trustworthiness. The study invites the further question of whether the presence of such icons can be a way to more quickly build a relationship of trust, which is considered an important element in therapeutic relationships [22]. In future studies for consultations in which the doctor and patient are not in the same room, we need to demonstrate that even if the equipment is not deployed, displaying this equipment on-screen can have a measurable impact on outcomes for patients. Similarly, the impact of medical equipment may also apply in circumstances in which doctors consult patients in person. This too needs to be further explored.

Conclusions

When doctors appear in static images, 3 or more icons of medical equipment may be helpfully included in the images, one of which should be a stethoscope. These icons are likely to have a positive influence on patients' perceptions of the trustworthiness of the practitioner.

Acknowledgments

We acknowledge the support of Zaminah Khadaroo, Dr Wendy Chan She-Ping Delfos, and Christopher Kueh in developing the materials for this study and in data collection. We thank Dr David Lim, Ms Lucy Wilkie, and Ms Deepa Sriram for support in preparing the manuscript.

Funding: Curtin University.

Conflicts of Interest

None declared.

Multimedia Appendix 1

No icon.

[JPG File, 1MB - jmir_v14i4e100_app1.jpg]

Multimedia Appendix 2

Otoscope.

[JPG File, 1MB - jmir_v14i4e100_app2.jpg]

Multimedia Appendix 3

Reflex hammer.

[JPG File, 1MB - jmir_v14i4e100_app3.jpg]

Multimedia Appendix 4

Pen.

[JPG File, 1MB - jmir_v14i4e100_app4.jpg]

Multimedia Appendix 5

Stethoscope.

[JPG File, 1MB - jmir_v14i4e100_app5.jpg]

Multimedia Appendix 6

Surgical scrubs.

[JPG File, 1MB - jmir_v14i4e100_app6.jpg]

References

1. Bleakley A, Farrow R, Gould D, Marshall R. Learning how to see: doctors making judgements in the visual domain. J Workplace Learn 2003;15(7-8):301-306. [doi: 10.1108/13665620310504765]
2. Fletcher C. Listening and talking to patients. I: The problem. Br Med J 1980 Sep 27;281(6244):845-847. [Medline: 7427478]
3. Bruhn JG. The doctor's touch: tactile communication in the doctor-patient relationship. South Med J 1978 Dec;71(12):1469-1473. [Medline: 364668]
4. Sanders G, Craddock C, Wagstaff I. Factors influencing default at a hospital colposcopy clinic. Qual Health Care 1992 Dec;1(4):236-240 [FREE Full text] [Medline: 10136870]
5. Cornfield LI. Candy coated tongue depressor. US patent 2,857,908. 1958 Oct 28. URL: http://tiny.cc/u52zgw [accessed 2012-06-28] [WebCite Cache ID 68we9HqMf]
6. Agha Z, Roter DL, Schapira RM. An evaluation of patient-physician communication style during telemedicine consultations. J Med Internet Res 2009;11(3):e36 [FREE Full text] [doi: 10.2196/jmir.1193] [Medline: 19793720]
7. Stone J, Carson A. Functional neurologic symptoms: assessment and management. Neurol Clin 2011 Feb;29(1):1-18, vii. [doi: 10.1016/j.ncl.2010.10.011] [Medline: 21172567]
8. Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, Singer JP, et al. Placebos without deception: a randomized controlled trial in irritable bowel syndrome. PLoS One 2010;5(12):e15591 [FREE Full text] [doi: 10.1371/journal.pone.0015591] [Medline: 21203519]

9. Brody H. The doctor as therapeutic agent: a placebo effect research agenda. In: Harrington A, editor. The Placebo Effect: An Interdisciplinary Exploration. Cambridge, MA: Harvard University Press; 1997:77-92.
10. McCroskey JC. An Introduction to Rhetorical Communication, 7th edition. Boston, MA: Allyn and Bacon; 1997.
11. McCroskey JC, Teven JJ. Goodwill: a reexamination of the construct and its measurement. Commun Monogr 1999;66:90-103. [doi: 10.1080/03637759909376464]
12. Nguyen H, Masthoff J. Is it me or is it what I say? Source image and persuasion. Lect Notes Comput Sci 2007;4744:231-242. [doi: 10.1007/978-3-540-77006-0_29]
13. Rehman SU, Nietert PJ, Cope DW, Kilpatrick AO. What to wear today? Effect of doctor's attire on the trust and confidence of patients. Am J Med 2005 Nov;118(11):1279-1286. [doi: 10.1016/j.amjmed.2005.04.026] [Medline: 16271913]
14. Fischer RL, Hansen CE, Hunter RL, Veloski JJ. Does physician attire influence patient satisfaction in an outpatient obstetrics and gynecology setting? Am J Obstet Gynecol 2007 Feb;196(2):186.e1-186.e5. [doi: 10.1016/j.ajog.2006.09.043] [Medline: 17306675]
15. Thomas MW, Burkhart CN, Lugo-Somolinos A, Morrell DS. Patients' perceptions of physician attire in dermatology clinics. Arch Dermatol 2011 Apr;147(4):505-506. [doi: 10.1001/archdermatol.2011.70] [Medline: 21482906]
16. Schuklenk U. Medical professionalism and ideological symbols in doctors' rooms. J Med Ethics 2006 Jan;32(1):1-2. [doi: 10.1136/jme.2005.014373] [Medline: 16373513]
17. Lynn LA, Bellini LM. Portable knowledge: a look inside white coat pockets. Ann Intern Med 1999 Feb 2;130(3):247-250. [Medline: 10049219]
18. Hiramanek N, O'Shea C, Lee C, Speechly C, Cavanagh K. What's in the doctor's bag? Aust Fam Physician 2004 Sep;33(9):714, 716-714, 720 [FREE Full text] [Medline: 15487365]
19. Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991;26:499-510. [doi: 10.1207/s15327906mbr2603_7]
20. Britt H, Miller GC, Charles J, Henderson J, Bayram C, Harrison C, et al. General Practice Activity in Australia 2007-08. General Practice Series No. 22. Cat. No. GEP 22. Canberra: Australian Institute of Health and Welfare; 2008.
21. Peirce CS. Collected Papers of Charles Sanders Peirce, Volumes 1-8. Edited by: C Hartshorne,*P* Weiss, and AW Burks. Cambridge, MA: Harvard University Press; 1965.
22. Pellegrino ED, Veatch RM, Langan JP. Ethics, Trust, and the Professions: Philosophical and Cultural Aspects. Washington, DC: Georgetown University Press; 1991.

Abbreviations

CI: confidence interval
OR: odds ratio
SCS: Source Credibility Scale

Edited by G Eysenbach; submitted 07.11.11; peer-reviewed by N Siriwardena, L Phillips; comments to author 15.01.12; revised version received 09.02.12; accepted 29.04.12; published 10.07.12.

Please cite as:
Jiwa M, Millett S, Meng X, Hewitt VM
Impact of the Presence of Medical Equipment in Images on Viewers' Perceptions of the Trustworthiness of an Individual On-Screen
J Med Internet Res 2012;14(4):e100
URL: http://www.jmir.org/2012/4/e100/
doi:10.2196/jmir.1986
PMID:22782078

Original Paper

P2P Watch: Personal Health Information Detection in Peer-to-Peer File-Sharing Networks

Marina Sokolova[1,2,3], MSc, PhD; Khaled El Emam[1,2], PhD; Luk Arbuckle[1], MSc; Emilio Neri[1], BEng; Sean Rose[4], MSc; Elizabeth Jonker[1], BA

[1]Electronic Health Information Laboratory, CHEO Research Institute, Ottawa, ON, Canada

[2]Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[3]Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[4]Privacy Analytics, Ottawa, ON, Canada

Corresponding Author:
Marina Sokolova, MSc, PhD
Electronic Health Information Laboratory
CHEO Research Institute
401 Smyth Rd
Ottawa, ON, K1H 8L1
Canada
Phone: 1 613 737 7600 ext 4104
Fax: 1 613 731 1374
Email: sokolova@uottawa.ca

Abstract

Background: Users of peer-to-peer (P2P) file-sharing networks risk the inadvertent disclosure of personal health information (PHI). In addition to potentially causing harm to the affected individuals, this can heighten the risk of data breaches for health information custodians. Automated PHI detection tools that crawl the P2P networks can identify PHI and alert custodians. While there has been previous work on the detection of personal information in electronic health records, there has been a dearth of research on the automated detection of PHI in heterogeneous user files.

Objective: To build a system that accurately detects PHI in files sent through P2P file-sharing networks. The system, which we call P2P Watch, uses a pipeline of text processing techniques to automatically detect PHI in files exchanged through P2P networks. P2P Watch processes unstructured texts regardless of the file format, document type, and content.

Methods: We developed P2P Watch to extract and analyze PHI in text files exchanged on P2P networks. We labeled texts as PHI if they contained identifiable information about a person (eg, name and date of birth) and specifics of the person’s health (eg, diagnosis, prescriptions, and medical procedures). We evaluated the system’s performance through its efficiency and effectiveness on 3924 files gathered from three P2P networks.

Results: P2P Watch successfully processed 3924 P2P files of unknown content. A manual examination of 1578 randomly selected files marked by the system as non-PHI confirmed that these files indeed did not contain PHI, making the false-negative detection rate equal to zero. Of 57 files marked by the system as PHI, all contained both personally identifiable information and health information: 11 files were PHI disclosures, and 46 files contained organizational materials such as unfilled insurance forms, job applications by medical professionals, and essays.

Conclusions: PHI can be successfully detected in free-form textual files exchanged through P2P networks. Once the files with PHI are detected, affected individuals or data custodians can be alerted to take remedial action.

(J Med Internet Res 2012;14(4):e95) doi:10.2196/jmir.1898

KEYWORDS

Privacy; personal health information; natural language processing, text data mining

Introduction

Evidence shows that files sent through peer-to-peer (P2P) file-sharing networks can disclose an individual's personal health information (PHI) to millions of network users. PHI refers to information about one's health that can be discussed in a clinical setting [1]. For example, in more than 3000 files exchanged on P2P networks, 5% contained either sensitive or sufficient information to commit medical identity theft, sometimes for thousands of individuals [2]. In another study, the authors semimanually examined 859 files gathered from two P2P networks and found that 8 (1%) files contained PHI [3]. Although the disclosure numbers look comparatively small, files on P2P networks are accessible to millions of network users. The same study also showed that the P2P network users may not even be aware that the files can be read by all the peers.

P2P files may be in various media (eg, visual, audio, and text), may address various topics (eg, fashion, tax report, and family life), and may be written in any language (eg, Spanish and French). To effectively deal with such challenges, automated PHI detection must perform well on multiple tasks, such as language identification, filtering out of damaged and virus files, text extraction, and hierarchical multiclass classification of documents. At the same time, the volumes of files exchanged through P2P networks and expectations of privacy of personal communications make manual PHI detection impossible. While there are several traditional PHI detection tools, they are not suitable for the large-volume analysis of heterogeneous documents. For example, some of these tools are designed to work with semistructured electronic health records and to find personal identifiers, such as patients' and doctors' names, insurance parameters, and hospital and clinic names [4-10].

In this paper, we describe and evaluate a new system—P2P Watch—that has been constructed specifically to crawl through P2P networks and automatically detect whether a retrieved file contains PHI. To be defined as PHI, a P2P file must contain information that would provide someone with the ability to identify a unique individual as well as health information on that individual, such as procedures or drugs. For example, generic statements such as "John Smith caught a cold" would be rejected as PHI according to our definition, unless they are reinforced by Smith's residential or work address and the prescription drugs he is taking.

We empirically evaluated our system on three networks: FastTrack, Gnutella, and eD2K. These were chosen due to their global popularity and high share of users. We harvested 3924 files and applied P2P Watch on the file contents. No author metadata was used in the file analysis.

We concentrate on PHI for Canadians. For example, our syntactic patterns that detect provincial health care numbers and the organization types are adjusted for Canada. Although P2P Watch focuses on Canadian PHI, at the same time, it recognizes geographic locations and zip codes in the United States because Canadians may have been born elsewhere or the PHI may concern a trip.

P2P Watch provides a mechanism for data custodians and individuals to determine whether information about their patients, employees, or themselves is being exposed. To minimize potential organizational and individual harm from such inappropriate disclosures, automated PHI detection tools can crawl through P2P networks looking for PHI. Once PHI is detected, affected individuals and data custodians can be alerted to take remedial action.

PHI disclosure on P2P networks is part of a wider trend of PHI presence on the Web [11-14]. PHI appears in electronic news, blogs by health care professionals and military personnel, Web-posted user messages, medical student papers, and personal letters [15-23]. For example, Doing-Harris and Zeng-Treitler [15] extracted health-related terms from messages posted on PatientsLikeMe.com. They manually evaluated the used vocabulary and found 651 health terms that were not yet included in a medical thesaurus. Another study analyzed user requests posted on an involuntary childlessness message board [16]. Blogs written by military servicemen were examined to find descriptions of clinically relevant combat exposure [17]. Lampos and Christianini [18] used Wikipedia's page on influenza and the UK's National Health Service website for automated extraction of influenza-like illness markers. They subsequently used the markers to find H1N1-related tweets but did not extract personally identifiable information (PII) about the users. Sokolova et al [19] presented a method of patient-based health information extraction from P2P files. They manually evaluated extraction accuracy on 2000 P2P files.

Interest in PHI published on the Web is ongoing [20]. Although openness and information sharing are beneficial to the population at large, users may not be aware of the secondary use of their information and consequent privacy issues [21]. A qualitative study of 123 user comments on the online community PatientsLikeMe was dedicated to analysis of sharing PHI among people with similar ailments [22]. Chou et al [23] identified younger users, those with poorer subjective health, and those with a personal cancer experience as more likely participants in online support groups and more willing to share their PHI.

So far, PHI detection tools have been developed and deployed by health care organizations in the context of de-identifying the organization's records, such as clinical discharge summaries, nurses' notes, and pathology reports. The main de-identification approach was to classify individual words as presenting personally identifiable information (PII) or not [4-10]. Such approaches require a substantial amount of labeled training data (eg, 1000 documents [7]) and consume considerable processing time.

In electronic health records, PHI detection can be boosted by the use of the personal information found in the structured part of the document or by pulling in structured information from the medical record database. Customized dictionaries present another source of accuracy in detecting PHI—these include local geographic names, health care organizations, and patient names [24]. These tools also can determine with certainty that there is health information in the documents they analyze and therefore focus only on the detection of PII. We provide evidence that such tools can fail to identify PHI in free-form textual files.

Methods

We designed and implemented P2P Watch, which automatically detects P2P network files that contain PHI.

System Architecture

The system was designed as a pipeline of seven components: (1) duplicate file removal, (2) media content removal, (3) text extractor, (4) language identifier, (5) publishable content identifier, (6) PII detector, and (7) patient-oriented health information detector. Components 1–4 identify and filter out irrelevant files through a shallow analysis, component 5 finds irrelevant files by applying partial content analysis, and components 6 and 7 identify relevant files within the remaining set. At each stage files may be discarded, resulting in fewer and fewer files making it through the pipeline.

Figure 1 presents the system design.

Duplicate Removal

The first task of the file processing was to find and remove multiple copies of the same file; such duplicates can happen, as the same file can be harvested from multiple users. For each pair of files, we compared their sizes (in kilobytes), titles, and first and last sentences. If all the parameters were the same, we tagged two files as duplicates and kept only one for further processing.

Media Content Removal

We assumed that any published text was not leaking PHI—for example, writings describing fictional characters, and magazine and newspaper articles that contain information that is already public. We used Amazon Web Services as a source database of publication titles [25]. Although the number of titles fluctuates almost daily, the database has 400,000 to 500,000 titles for books; recording companies such as Sony, EMI, and Universal have more than 250,000 music titles in the database. Files with exact matching titles were discarded. Exceptions were made for files with titles that included such words as *notification*, *affidavit*, *justice*, *discharge*, and *lab*. Our system did not discard these files and retained them for further processing. If there was no exact title matching, the file was passed on for further processing.

Text Extractor

The text extractor converted files into raw text by removing all formatting meta-information. It also discarded nontext files (eg, images and music), corrupted files, and viruses. A wide range of input file formats suggested the use of several tools, with each tool extracting text from specific formats. We used the open source program Antiword to extract text from Microsoft Word documents [26]. To extract text from PDF, RTF, HTML, or XML files, we used the open source programs MineText [27] and GetText [28].

Language Identifier

We differentiated between English-language texts and texts written in other languages. We applied a publicly available language identifier, TextCat Language Guesser, which can identify 69 languages [29]. For text, the tool outputs several possible languages. If English was the most likely language of the text, then it appeared at the beginning of the output. Our manual examination had shown that, in our sample data, the first English tag always correctly marked texts written in English. We discarded a file if English was not the most likely language of the text.

Publishable Content Removal

P2P Watch looked for files with nonpersonal content. It filtered out published and educational materials (eg, assignments and theses) and nonpersonal texts (eg, manuals and technical reports) that were not found by the title lookup. We also hypothesized that music lyrics, discussion of popular fictional characters or current political events, and advertisement would be unlikely candidates for leaking explicit, detailed PHI. We built a list of fictional characters (eg, Bart Simpson), celebrities (eg, Paris Hilton), and public figures (eg, George Bush). We considered that, by the nature of their occupations, celebrities and public figures would have a lot of information about them publicly known and therefore any PHI pertaining to these individuals would not be considered a breach. To perform this task, we built a list of terms that appeared in the preface of publishable and educational texts; the terms are listed in Multimedia Appendix 1. We used regular expressions to locate those terms and their variations in the first 200 words of the file. Table 1 lists word categories and examples.

Figure 1. Components of P2P Watch. P2P = peer-to-peer, PHI = personal health information.

Table 1. Publishable content identifiers.

Category	Example
Books	Ebook, ISBN
Education	Thesis, assignment
Retail	Tim Hortons
Periodical	Magazine, article
Fictional	Harry Porter
Politics	Nicolas Sarkozy

PII Detector

We considered that a person can be identified from first and last names, addresses, dates (which can be linked to identifiable events such as birth, death, and marriage), and organization names (eg, school, church, and professional association). We divided PII into three categories: person information (eg, names, family relations, and age-defining events), structured information (eg, dates, telephone numbers, and email), and geographic location (eg, street address and organization). For instance, telephone numbers are both geographic identifiers and numeric identifiers. Figure 2 illustrates the category relations.

For the personal name lookup, we acquired a female first name list, a male first name list, and a last name list. The lists contained both formal and informal name forms (eg, William, Bill, and Billy) and non-Anglo-Saxon names (eg, Meehai and Leila), as well as common misspellings (eg, Bll). The lists came with a commercial database marketing tool that performed best in an independent evaluation [30].

To reduce computationally expensive person name lookups, we first searched for patterns of family relations in the text (eg, “my daughter” and “an uncle of”), self-identification (eg, “my name” and “sincerely”), or life event (eg, “was born” and “died in”). Depending on the patterns, either the preceding or the following capitalized words were stored in a file’s name list. Having a file’s name list considerably accelerated the file processing. Further, when P2P Watch checked for a person name, it first checked with the file’s name list.

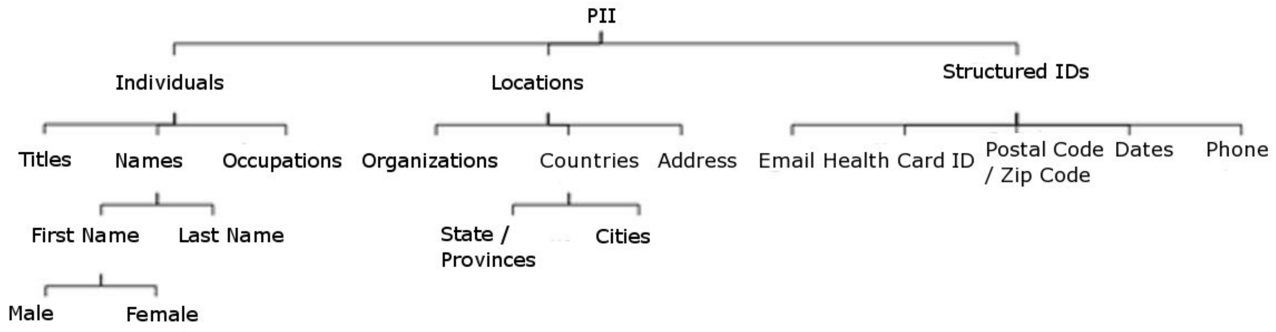
Additionally, PII included standardized information as follows: (1) telephone numbers: we looked for complete and incomplete formats, used in North America, (2) health insurance numbers: we looked for health insurance numbers assigned by each province (Canada), (3) dates: we restricted the dates to the 20th and 21st centuries, as earlier dates are unlikely to be related to health information of living human beings; also, dates had to be specific: “March 9th, 1999” indicated a specific date, whereas “March was chilly” did not, (4) email address: for example, john@canada.ca, john AT Canada DOT ca, (5) postal codes in Canada and zip code in the United States. These five categories were retrieved by manually built soft regular expressions.

Apart from geographic locations in Canada and the United States, we used some international information and introduced different granularity for different geographic categories. First was country: all the UN-recognized countries and their capitals (eg, France and Paris; Liberia and Monrovia) and self-proclaimed entities (eg, Eritrea and Abkhazia). Second was place. In the United States, we used state name, state capital, and—to cover the biggest single population unit—the largest city in the state; for example, we had Illinois, Springfield, and Chicago. In Canada, we used province, provincial capital, largest cities, and tourist attractions (eg, Alberta, Edmonton, Calgary, Banff), and the same for territories. In Europe, Latin America, Asia, Africa, and Australia, we used major cities that are not national capitals. The list is given in Multimedia Appendix 2.

The Canadian Judicial Council further considers certain organizations to be part of PII [31], such as public institutions (eg, schools and churches), care providers (eg, aid societies and foster homes), the names of support organizations (eg, women’s and senior’s support centers), and other location identifiers (eg, educational institutions and military bases). The organization names were expressed in many language forms. We modeled language patterns (eg, “lived in” and “come from”), organization types (eg, schools, military units, and churches), and the target population (eg, youth, women, and seniors).

To be marked as a text with PII, the file had to contain a geographic identifier and two other personal identifiers, such as a person’s first name and last name, or a person’s first name and date of birth. All the files marked as PII were passed on to the last component.

Figure 2. Personally identifiable information (PII) categories and their subcategories. ID = identification.



Patient-Oriented Health Information Detector

Disease names (eg, arthritis and mumps) and symptoms (eg, chest pain and headache), and procedures (eg, heart surgery and x-rays) most directly convey health information that is usually discussed in a clinical setting. To build a list of corresponding terms, we used the International Classification of Diseases [32] and the Medical Dictionary for Regulatory Activities (MedDRA) [33]. Drug names, too, may allow one to infer a specific medical, behavioral, or psychological condition or ailment of another individual. We used the Canadian Drug Product Database (Active and Inactive), which contains the names of drugs approved for use in Canada and previously available drugs [34]. To accommodate extraction of various drug names, we obtained a list of generic drug names and the trade names associated with them [35]. However, the resources listed above leave some gaps in the detection of health information. The most noticeable are acronyms (eg, ICU), specialties of health care providers (therapist, surgeon), and some condition names (blood pressure, tube fed). To fill the gaps, we manually searched Webster's New World Medical Dictionary [36]. We minimized the above-mentioned resources by removing unrelated categories (animal diseases, animal drugs). Then the remaining texts were converted to lowercase, punctuation marks and numbers were removed, and stop words (eg, of and when) were eliminated. The list of the resulting keywords is given in Multimedia Appendix 3.

More details on the method can be found in Sokolova et al [19].

Empirical Evaluation

Project approval from the Research Ethics Board of Children's Hospital of Eastern Ontario was obtained prior to retrieving data from the P2P networks.

Files Analyzed

The files were gathered from April 2008 to June 2009 from three networks. We selected FastTrack, Gnutella, and eD2K networks due to their global popularity and high share of users [2]. To automatically search for and download P2P files, we modified the publicly available Shareaza P2P client [37], which is a software package allowing one to connect to multiple P2P networks simultaneously. In-house modifications to Shareaza included changes to the search function and an increase of logging capabilities.

We focused on capturing the most popular document formats: Microsoft Word (.doc), raw text (.txt), Rich Text Format (.rtf), Excel (.xls), PowerPoint (.ppt), Portable Document Format (.pdf), Extensible Markup Language (.xml), and HyperText Markup Language (.html). The search function was modified to automatically search for those formats and automatically retrieve the files. Automatic searches were conducted by the code at 15 minute intervals.

In total, we have gathered 3924 files. The data were sent for processing as is, without preliminary normalization: we preserved all the initial spelling, capitalization, grammar, and so on.

Performance Evaluation

We evaluate the system's efficiency through the time (seconds) it took each component to process the related files. Technical specifications of our equipment were as follows: Windows Server 2003 (Microsoft Corporation, Redmond, WA, USA), 3.20 GHz Intel Core i3 processor (Intel Corporation, Santa Clara, CA, USA), 4 GB RAM, and 500 GB SATA hard disk.

Effectiveness Evaluation

We evaluated the effectiveness of P2P Watch by the number of PHI files found among the files it discarded as non-PHI (false-negative PHI files) and by the number of PHI files it marked as PHI (true-positive PHI files).

Exact estimation of true-positive PHI files was feasible due to the small output we expected. However, the exact estimation of false-negative PHI was not feasible due to the large volume of data. To estimate the presence of false-negative PHI files in the discarded files, we employed a sampling technique. The sampling technique is described below.

Comparison With Other PHI Detection Tools

Currently deployed PHI tools are designed to analyze electronic health records produced by selected health care organizations [4-10]. The tools work on the assumption that (1) there are PHI words within each document, and (2) these words belong to a restricted number of categories (eg, local doctors, local hospitals). Most of the tools are proprietary and cannot be easily assessed for comparative evaluation [4-7,9,10].

The open source PHI detection tool De-id is popular among researchers [8,24]. To ensure a fair tool comparison and adherence to a main content assumption of PHI presence in the text, we applied De-id on P2P files that P2P Watch had been identified as PHI.

Sample Size to Detect PHI in Discarded Files: False-Negative Evaluation

Because there were 3924 files in total and we expected most of them not to contain PHI, a manual examination of all of the discarded files would have been exceedingly time consuming. To compute the false-negative rate, we estimated the number of PHI files that could appear in a random sample of P2P files.

We wanted to determine sample sizes in advance. To obtain a conservative estimate, we decided on multiple sampling, where an individual sample is randomly drawn from a separate group of discarded files. Based on the P2P Watch architecture, we identified the three main groups of discarded files: group 1—files discarded by Amazon search, text extractor, and language identifier; group 2—files discarded by the content filter; group 3—files discarded by the PII detector and health information detector.

We used a binomial distribution as a model for P2P file data, assuming that either a file contains PHI or it does not [38,39]. Equation 1 in Figure 3 shows the probability of detecting at least one file with PHI, assuming a binomial distribution, where θ is the rate of PHI we wished to detect, and n is the number of independent samples. Equation 1 can be rearranged as Equation 2 (Figure 3).

Previous studies showed similarity in the ratio of PHI files among reviewed P2P files: approximately 1% of all the files contained PHI. Therefore, we used this estimate in Equation 2 to define the sample size for groups 1–3. For each group of discarded files, to have a 95% chance of detecting PHI when the underlying rate of files with PHI is at least 1%, we needed to sample at least 300 files. For the complete data, we wanted to sample at least 900 files.

Figure 3. Equation (1, rearranged in 2) for determining the probability of detecting at least one file containing personal health information.

$$P(X>0)=1-P(X=0)=1-(1-\theta)^{n}, \quad (1)$$

$$n=\frac{\ln(1-P(X>0))}{\ln(1-\theta)}. \quad (2)$$

Special Protocols

Three special protocols were put in place for this study. First, we expected some files to contain inappropriate or obscene material (eg, pornography). We therefore did not explicitly look through image files (file extensions .gif, .jpg, .psd, .tif, and .bmp). Second, if we discovered any illegal materials (eg, child pornography), we passed that information on to the police. Third, if there were cases of disclosure of particularly sensitive personal information or PHI for a large number of individuals, then we reported them to the appropriate federal or provincial privacy commissioner for follow-up.

Results

We applied P2P Watch for PHI detection in 3924 files exchanged on the three P2P networks. The total data set size was 9887 MB. The total processing time was 4132.41 seconds.

Figure 4 illustrates changes in the number of files processed by each component during our empirical evaluation.

Figure 4. File processing steps by P2P Watch. PHI = personal health information, PII = personally identifiable information.

Observations

We made several general observations during the analysis as follows.

Duplicate File Removal

We considerably reduced processing time by first testing whether a file was a duplicate of an already gathered file. Removal of duplicates does not affect the quality of PHI detection, as P2P Watch should process the original file. We discarded 514 files (eg, Quick Recipe And Meal Ideas, 36 Christmas Carols & Songs), mostly books and technical manuals; the size of discarded files was 786 MB.

Media Content Removal

We discarded 286 files, which contained content that was found in the Amazon.com database (eg, Iron Crypt of the Heretics; The Haunted Lighthouse; New York, New York; ABBA's The Winner Takes It All; and The Abominable Snowman); the size of discarded files was 1234 MB. We requested the exact match of the file title with the database entry; hence, some files with published content were passed through this filter.

Text Extractor

We worked with different file formats to extract text and discard images, virus files, and corrupted files. At this stage, 673 files were discarded, mostly mp3 and rm files with music and video contents; the size of the discarded files was 3077 MB.

Language Identifier

We identified English-language texts among the remaining files; 882 non-English texts were discarded, with a total size of 10 MB. To avoid misrepresentation of North America as a polyglot continent, we checked the discarded files and found that many of them had explicitly erotic content.

Publishable Content Identifier

Based on processing of the first 200 words, 880 files were discarded, with a total size of 171 MB. Books composed the majority of the discarded files, with the remaining part comprising cover letters, resumes, homework, and forms.

Personally Identifiable Information Detector

Based on the search for PII within a complete file, 550 files were discarded. Most of these texts were in a small publishable form such as articles, opinion pieces, local community letters, forms, and job application packages. The size of discarded files was 3.9 MB.

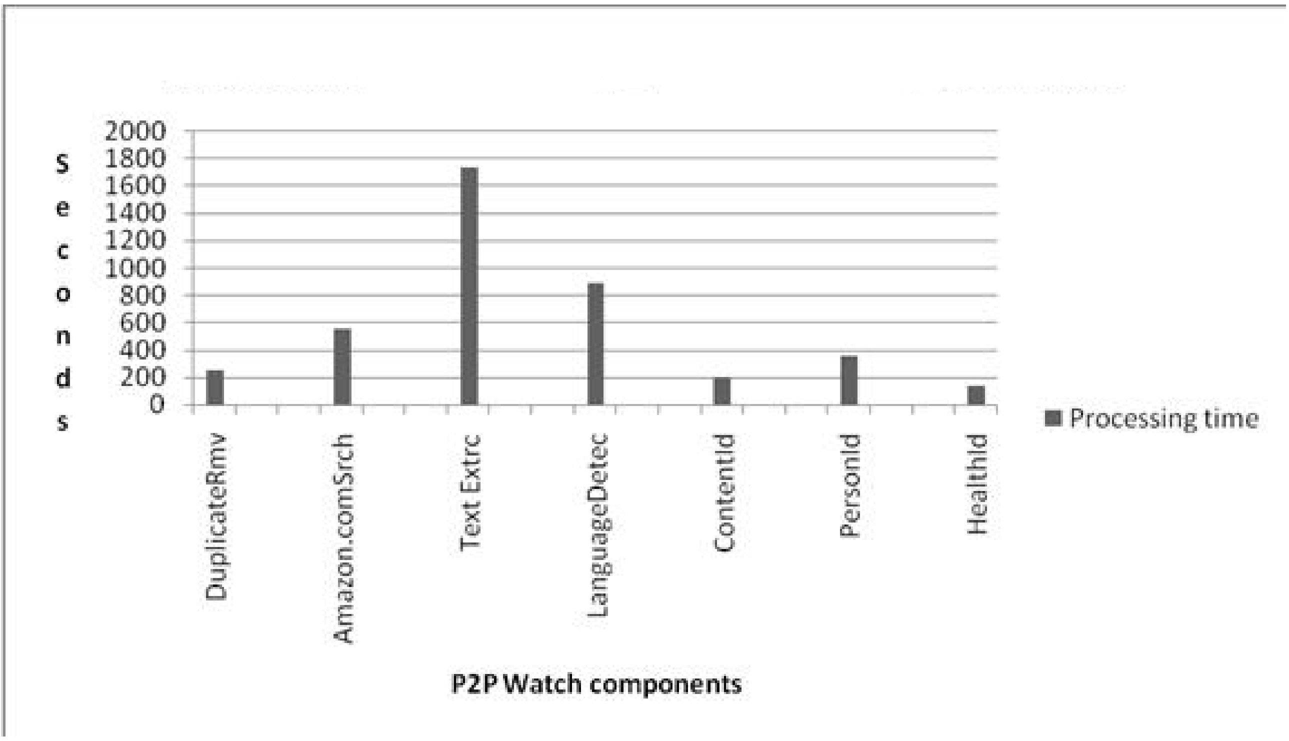
Patient-Oriented Health Information Detector

The last component identified 57 files as potential PHI and discarded 240 files. Among the discarded files, most promoted different types of consumer goods and services (credit services, fitness and skin care, gadgets, etc); some tax and insurance forms were discarded as well.

Efficiency Evaluation

We timed the performance of each component. The most time was used by the text extractor, the first component to process the complete file text. The most efficient components were duplicate removal and Amazon.com search, which processed the most files in the least time. Figure 5 illustrates time spent by each component on the data processing.

Figure 5. Processing time (seconds) of P2P Watch components.



Effectiveness Evaluation: True-Positive Rate

P2P Watch flagged 57 files as containing PHI. Manual analysis confirmed that all the files contained PII and PHI. However, we distinguished between true PHI and pseudo-PHI. Among these 57 files, 11 contained health information about an identifiable individual: they were indeed PHI (true PHI). PHI appeared in various types of documents—for example, a note to a temporary guardian and a lawyer's note were shown to contain PHI.

Other marked files contained both PII and PHI, but were not PHI (pseudo-PHI). A few examples of pseudo-PHI discovered through a manual check of the P2P files contained information that the information owner had intentionally allowed to be part of a public audience. In one instance, the data owner stated "When I was four, I ended up in the hospital for playing with medicine" as part of a book report on Curious George. Another pseudo-PHI file was a nursing student's assignment, which contained initials for a patient at a hospital and a room number, but not enough information to identify the patient without access to more detailed hospital records. The curriculum vitae of a medical professional and a medical insurance form are other examples of such files.

Effectiveness Evaluation: False-Negative Rate

We ran P2P Watch without the duplicate removal filter. We invited two independent evaluators to read all the sampled files and mark them as PHI or not. One author (EN) participated in the evaluation of some files.

We first randomly selected three samples of 300 discarded files, where one sample consisted of files discarded by Amazon search, text extractor, and language identifier; another sample consisted of files discarded by the content filter; and the third sample consisted of files discarded by the PII detector and health information detector.

None of sampled 900 files contained PHI. We, however, were concerned with the possibility of PHI in files discarded by the PII detector and health information detector. We therefore chose to manually check all the remaining 678 files in this stratum, and found none containing PHI.

As a result, we manually checked 1578 discarded files and found that none of them contained PHI.

Comparison With De-id

The open source tool De-id [8] gave us an opportunity to test the applicability of existing tools to finding PHI in P2P files. We applied De-id to find PHI in the 57 files that P2P Watch marked as PHI. By that time, we knew that the files indeed contain PII and PHI. We had to use the text extractor component, as De-id works only with text format.

De-id crashed on 3 files. The tool did not recognize any of remaining 54 files as PHI. The tool mislabeled many critical terms (eg, *risk* and *blood* were both marked as ambiguous last names, and *disorder* and *depression* were not recognized as health related). De-id took an average of 11 seconds to process a short file.

This empirical evidence showed that major components have to be added to De-id: a text extractor, and new PII and PHI detection components.

Discussion

In this study, we have introduced P2P Watch, which detects PHI in files shared by users of P2P networks. Albeit the proportion of PHI files among P2P files is rather small, the overall problem is big, as by some estimates 50% of files downloaded and 80% of files uploaded on the Internet are through P2P networks[40]. However, even one PHI file can do much harm, especially if it contains an exact pointer to a publicly available data base. At the same time, we empirically showed that traditional de-identification tools are not designed to detect PHI in P2P files.

P2P Watch is capable of working within the complex environment of P2P networks. It detects PHI in files in which context, content, and format type vary. Within the data set of 3924 P2P files, the system detected 11 PHI files. Our manual evaluation of 1578 files, marked by the system as non-PHI, confirmed that these files indeed did not contain PHI. The sampling results showed that P2P Watch was very unlikely to miss PHI files.

For successful PHI detection in P2P networks, it is essential that the detection system process large volumes of heterogeneous data input in a timely manner and can withstand substantial irrelevant information. A reliable solution is based on two factors [41]. First, a high confidence that the limited number of analyzed texts will not exclude any possible PHI texts is needed; this can be achieved through filtering out only the files guaranteed not to contain PHI. Second, a speedy detection process is needed that prevents a prolonged presence of PHI texts on the network; this can be achieved through minimizing time of filtering with respect to performed text analysis.

P2P Watch efficiently reduces the time of PHI exposure. Our detection strategy is up-front shallow text processing, whose goal is to quickly process the vast majority of input files, followed by a thorough analysis of a small number of selected texts. This thorough analysis phase used electronic and hard-copy dictionaries of health care terms, an ontology of medical terms, and lists of personal names. We supplemented those sources with in-house built gazetteers (topical lists of geographic information) and lists of organization types.P2P Watch reserves comprehensive text analysis for a small number of selected files, while performing fast and accurate shallow processing for the vast majority of files. This is the principal difference from previously built systems, which process all the input files equally. The difference may be explained by the fact that previous systems were designed to detect patient's PII in electronic medical records, whereas our P2P Watch searches for both PII and PHI in previously unseen documents. Once a file is flagged as containing PHI, the individuals affected can be alerted. A search for a data custodian's name within the flagged files would indicate which custodian to alert, for example.

Several possible expansions of the functionality of P2P Watch are being considered for future work. Our current detection is limited to text written in English. Expanding P2P Watch capacities to other language such as French and Spanish would capture more PHI leaks. Furthermore, we want to build separate components to identify files that contain PII and HI, but are not PHI (pseudo-PHI). Resumes, recipes, incomplete health forms (eg, insurance), and public health announcements are examples of files that were falsely labeled as PHI. The idea would be to detect these types of documents and automatically exclude them early in the P2P Watch pipeline. A challenge with forms is that the empty forms may contain pseudo-PHI (eg, fields for human immunodeficiency test results). Special analysis of such forms is required to determine which content is part of the form and which is completed by the user.

Another future avenue is to add localized detection of PHI in the United States; this expansion may involve building new customized lists of organization and trademark names. In the future, a deeper analysis phase, perhaps coreference resolution, could be done, potentially increasing the precision of the whole detection process.

Acknowledgments

We thank Ontario Centres of Excellence and the Natural Sciences and Engineering Research Council of Canada for their financial support. We also thank Terry Copeck for his work on the text extractor filter. We thank Sadrul Chowdhury for his work on the system. We thank the manuscript reviewers and editors for thorough and helpful comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Publishable information keywords.

[TXT File, 1KB - jmir_v14i4e95_app1.txt]

Multimedia Appendix 2

Geographic locations.

[TXT File, 6KB - jmir_v14i4e95_app2.txt]

Multimedia Appendix 3

Health information keywords.

[TXT File, 951KB - jmir_v14i4e95_app3.txt]

References

1. Hersh W. Information Retrieval: A HealthBiomedical Perspective. 3rd edition. Hannah K, Ball M. editors. Health Informatics. New York, NY: Springer; 2009.
2. Johnson E. Data hemorrhages in the health-care sector. Lect Notes Comput Sci 2009;5628:71-89. [doi: 10.1007/978-3-642-03549-4_5]
3. El Emam K, Neri E, Jonker E, Sokolova M, Peyton L, Neisa A, et al. The inadvertent disclosure of personal health information through peer-to-peer file sharing programs. J Am Med Inform Assoc 2010;17(2):148-158 [FREE Full text] [doi: 10.1136/jamia.2009.000232] [Medline: 20190057]
4. Uzuner O, Sibanda TC, Luo Y, Szolovits P. A de-identifier for medical discharge summaries. Artif Intell Med 2008 Jan;42(1):13-35. [doi: 10.1016/j.artmed.2007.10.001] [Medline: 18053696]
5. Uzuner O, Luo Y, Szolovits P. Evaluating the state-of-the-art in automatic de-identification. J Am Med Inform Assoc 2007;14(5):550-563 [FREE Full text] [doi: 10.1197/jamia.M2444] [Medline: 17600094]
6. Aberdeen J, Bayer S, Yeniterzi R, Wellner B, Clark C, Hanauer D, et al. The MITRE Identification Scrubber Toolkit: design, training, and assessment. Int J Med Inform 2010 Dec;79(12):849-859. [doi: 10.1016/j.ijmedinf.2010.09.007] [Medline: 20951082]
7. Yeniterzi R, Aberdeen J, Bayer S, Wellner B, Hirschman L, Malin B. Effects of personal identifier resynthesis on clinical text de-identification. J Am Med Inform Assoc 2010;17(2):159-168 [FREE Full text] [doi: 10.1136/jamia.2009.002212] [Medline: 20190058]
8. Neamatullah I, Douglass MM, Lehman LW, Reisner A, Villarroel M, Long WJ, et al. Automated de-identification of free-text medical records. BMC Med Inform Decis Mak 2008;8:32 [FREE Full text] [doi: 10.1186/1472-6947-8-32] [Medline: 18652655]
9. Wellner B, Huyck M, Mardis S, Aberdeen J, Morgan A, Peshkin L, et al. Rapidly retargetable approaches to de-identification in medical records. J Am Med Inform Assoc 2007;14(5):564-573 [FREE Full text] [doi: 10.1197/jamia.M2435] [Medline: 17600096]
10. Meystre SM, Friedlin FJ, South BR, Shen S, Samore MH. Automatic de-identification of textual documents in the electronic health record: a review of recent research. BMC Med Res Methodol 2010;10:70 [FREE Full text] [doi: 10.1186/1471-2288-10-70] [Medline: 20678228]
11. Lagu T, Kaufman EJ, Asch DA, Armstrong K. Content of weblogs written by health professionals. J Gen Intern Med 2008 Oct;23(10):1642-1646. [doi: 10.1007/s11606-008-0726-6] [Medline: 18649110]
12. Buchanan T, Joinson A, Paine C, Reips UD. Looking for medical information on the Internet: self-disclosure, privacy and trust. Health Inf Internet 2007;58(1):8-9.
13. Wright KB, Bell SB, Wright KB, Bell SB. Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. J Health Psychol 2003 Jan;8(1):39-54. [doi: 10.1177/1359105303008001429] [Medline: 22113899]
14. Joinson A. Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. Eur J Soc Psychol 2001;31:177-192.
15. 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 [FREE Full text] [doi: 10.2196/jmir.1636] [Medline: 21586386]
16. Himmel W, Reincke U, Michelmann HW. Text mining and natural language processing approaches for automatic categorization of lay requests to web-based expert forums. J Med Internet Res 2009;11(3):e25 [FREE Full text] [doi: 10.2196/jmir.1123] [Medline: 19632978]
17. Konovalov S, Scotch M, Post L, Brandt C. Biomedical informatics techniques for processing and analyzing web blogs of military service members. J Med Internet Res 2010;12(4):e45 [FREE Full text] [doi: 10.2196/jmir.1538] [Medline: 20923755]
18. Lampos V, Christianini N. Tracking the flu pandemic by monitoring the social web. 2010 Presented at: 2nd Workshop on Cognitive Information Processing; Jun 14-16, 2010; Elba Island, Tuscany, Italy.
19. Sokolova M, El Emam K, Rose S, Chowdhury S, Neri E, Jonker E, et al. Personal health information leak prevention in heterogeneous texts. 2009 Presented at: Adaptation of Language Resources and Technology to New Domains; Sep 17, 2009; Borovets, Bulgaria URL: http://dl.acm.org/citation.cfm?id=1859157

20. Sokolova M, Schramm D. Building a patient-based ontology for mining user-written content. 2011 Presented at: Recent Advances in Natural Language Processing; Sep 12-14, 2011; Hissar, Bulgaria URL: http://aclweb.org/anthology-new/R/R11/
21. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
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. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Dec;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
24. Tu K, Klein-Geltink J, Mitiku TF, Mihai C, Martin J. De-identification of primary care electronic medical records free-text data in Ontario, Canada. BMC Med Inform Decis Mak 2010;10:35 [FREE Full text] [doi: 10.1186/1472-6947-10-35] [Medline: 20565894]
25. Amazon.com. 2010. Amazon Web Services URL: http://aws.amazon.com/ [accessed 2010-08-23] [WebCite Cache ID 5sCQ5qVCZ]
26. van Os A. Adri van Os. 2008 Apr 15. Antiword URL: http://www.winfield.demon.nl/ [accessed 2010-08-23] [WebCite Cache ID 5sCQHkVjX]
27. Text-Mining Tool.com. 2010. Text Mining Tool URL: http://text-mining-tool.com/ [accessed 2010-08-23] [WebCite Cache ID 5sCQOQOmM]
28. Kryloff Technologies, Inc. 2010. GetText URL: http://www.kryltech.com/freestf.htm [accessed 2010-08-23] [WebCite Cache ID 5sCQWKNpJ]
29. van Noord G. Cavnar WB, Trenkle JM. TextCat URL: http://www.let.rug.nl/~vannoord/TextCat/ [accessed 2010-08-23] [WebCite Cache ID 5sCQhnQNb]
30. El Emam K, Jabbouri S, Sams S, Drouet Y, Power M. Evaluating common de-identification heuristics for personal health information. J Med Internet Res 2006;8(4):e28 [FREE Full text] [doi: 10.2196/jmir.8.4.e28] [Medline: 17213047]
31. Judges' Technology Advisory Committee. Canadian Judicial Council. 2005 Mar. Use of Personal Information in Judgments and Recommended Protocol URL: http://cjc-ccm.gc.ca/cmslib/general/news_pub_techissues_UseProtocol_2005_en.pdf [accessed 2012-06-05] [WebCite Cache ID 68CIEk1BH]
32. World Health Organization. US Centers for Disease Control and Prevention. 2009. International Classification of Diseases, Ninth Revision (ICD-9). Sep 1 URL: http://www.cdc.gov/nchs/icd/icd9.htm [accessed 2010-08-23] [WebCite Cache ID 5sCQnajgl]
33. MedDRA MSSO. Northrop Grumman Corporation. 2010 Jul 29. Medical Dictionary for Regulatory Activities URL: http://www.meddramsso.com/index.asp [accessed 2010-08-23] [WebCite Cache ID 5sCR0vi82]
34. Health Canada. 2010 Apr 1. Drug Product Database URL: http://www.hc-sc.gc.ca/dhp-mps/prodpharma/databasdon/ [accessed 2010-08-23] [WebCite Cache ID 5sCRCKIZv]
35. Merck Sharp & Dohme Corp. 2008 Jul. Trade Names of Some Commonly Used Drugs URL: http://www.merck.com/mmpe/appendixes/ap2/ap2a.html [accessed 2010-08-23] [WebCite Cache ID 5sCRGueLP]
36. Anonymous. Webster's New World Medical Dictionary. New York, NY: Wiley; 2003.
37. Discordia Ltd. 2010. Shareaza URL: http://www.shareaza.com/ [accessed 2010-08-23] [WebCite Cache ID 5sCPqEOcd]
38. Christman M. Sampling of rare populations. In: Pfeffermann D, Rao CR, editors. Handbook of Statistics 29A: Sample Surveys: Design, Methods and Applications. New York, NY: Elsevier; 2009.
39. Venette RC, Moon RD, Hutchison WD. Strategies and statistics of sampling for rare individuals. Annu Rev Entomol 2002;47:143-174. [doi: 10.1146/annurev.ento.47.091201.145147] [Medline: 11729072]
40. Chen K, Shen H, Zhang H. Leveraging social networks for P2P content-based files sharing in mobile ad hoc networks. In: Proceedings. 2011 Presented at: 2011 IEEE 8th International Conference on Mobile Adhoc and Sensor Systems; Oct 17-22, 2011; Valencia, Spain. [doi: 10.1109/MASS.2011.24]
41. Sokolova M, El Emam K, Chowdhury S, Neri E, Rose S, Jonker E. Evaluation of rare event detection. Lect Notes Comput Sci 2010;6085:379-383. [doi: 10.1007/978-3-642-13059-5_51]

Abbreviations

P2P: peer-to-peer
PHI: personal health information
PII: personally identifiable information

Edited by G Eysenbach; submitted 27.07.11; peer-reviewed by D Perez-Rey, K Doing-Harris; comments to author 17.11.11; revised version received 05.03.12; accepted 23.04.12; published 09.07.12.

Please cite as:
Sokolova M, El Emam K, Arbuckle L, Neri E, Rose S, Jonker E
P2P Watch: Personal Health Information Detection in Peer-to-Peer File-Sharing Networks
J Med Internet Res 2012;14(4):e95
URL: http://www.jmir.org/2012/4/e95/
doi:10.2196/jmir.1898
PMID:22776692

Corrigenda and Addenda

Correction: Development and Validation of Filters for the Retrieval of Studies of Clinical Examination From Medline

Robert Badgett[1], MD; Nader Shaikh[2], MD, MPH; Mini Pi[3]; Nancy L Wilczynski[4], PhD; K. Ann McKibbon[4], PhD; Andrea M Ketchum[5], MLIS; R. Brian Haynes[4], MD, PhD

[1]University of Kansas School of Medicine at Wichita, Department of Internal Medicine and Department of Preventive Medicine and Public Health, Wichita, KS, United States

[2]Children’s Hospital of Pittsburgh of UPMC, General Academic Pediatrics, UPMC, Pittsburgh, PA, United States

[3]Carnegie Mellon University, Department of Biology, Pittsburgh, PA, United States

[4]McMaster University, Department of Clinical Epidemiology and Biostatistics, Hamilton, ON, Canada

[5]University of Pittsburgh, Health Sciences Library System, Pittsburgh, PA, United States

Corresponding Author:
Robert Badgett, MD
University of Kansas School of Medicine at Wichita
Department of Internal Medicine and Department of Preventive Medicine and Public Health
1010 N Kansas
Wichita, KS, 67214-3199
United States
Phone: 1 316 293 2604
Fax: 1 316 293 1878
Email: rbadgett@kumc.edu

Related Article:

Correction of: http://www.jmir.org/2011/4/e82/

(J Med Internet Res 2012;14(4):e108) doi:10.2196/jmir.2232

KEYWORDS

information retrieval; MEDLINE

The authors of the recent JMIR publication “Development and Validation of Filters for the Retrieval of Studies of Clinical Examination From Medline” (J Med Internet Res 2011;13(4):e82) [1] inadvertently misplaced a decimal point in the final row of Table 4 (Comparison of the performance of filters for clinical examination, diagnosis, and treatment). The precision of the original report of the Haynes 1994 filter should be 22% and not 0.22. This leads to the F-measure being 36 and the NNR being 4.5. In addition, the following line in the discussion is incorrect and has been removed: “In addition, our results are consistent with Haynes initial reports of locating trials in the early 1990s (19) before the National Library of Medicine collaborated with the Cochrane Collaboration and introduced 'randomized controlled trial' as a MeSH term (final row of Table 4)”.

The data in Table 4 was compiled to support our recommendation that the National Library of Medicine should create a publication type for “diagnostic accuracy study” to label studies that quantify sensitivity and specificity for diagnosis. We regret the error, but do not believe it weakens our recommendation or alters the results of our study.

The precision of the Haynes 1994 treatment filter was surprisingly high in their study and is due in part to the development and validation being limited to 10 high-impact journals [2]. Our study used 161 journals.

The online version of the JMIR paper has been corrected together with publication of this correction notice. A corrected version has been submitted by the publisher to PubMed Central, but incorrect versions may persist on other sites.

References

1. Shaikh N, Badgett RG, Pi M, Wilczynski NL, McKibbon KA, Ketchum AM, et al. Development and validation of filters for the retrieval of studies of clinical examination from Medline. J Med Internet Res 2011;13(4):e82 [FREE Full text] [doi: 10.2196/jmir.1826] [Medline: 22011384]
2. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994 Dec;1(6):447-458 [FREE Full text] [Medline: 7850570]

Edited by G Eysenbach; submitted 27.07.12; this is a non–peer-reviewed article;accepted 27.07.12; published 03.08.12.

Please cite as:
Badgett R, Shaikh N, Pi M, Wilczynski NL, McKibbon KA, Ketchum AM, Haynes RB
Correction: Development and Validation of Filters for the Retrieval of Studies of Clinical Examination From Medline
J Med Internet Res 2012;14(4):e108
URL: http://www.jmir.org/2012/4/e108/
doi:10.2196/jmir.2232
PMID:22864147

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org