
Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 9 (2007), Issue 3 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Review

- Internet-Based Physical Activity Interventions: A Systematic Review of the Literature (e26)
Marleen van den Berg, Johannes Schoones, Theodora Vliet Vlieland. 3

Tutorial

- Reliability and Validity Issues Related to Interactive Tailored Patient Assessments: A Case Study (e22)
Cornelia Ruland, Suzanne Bakken, Jo Røislien. 20

Original Papers

- 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) (e20)
Ellen Beckjord, Lila Finney Rutten, Linda Squiers, Neeraj Arora, Lindsey Volckmann, Richard Moser, Bradford Hesse. 29
- Response Rate and Completeness of Questionnaires: A Randomized Study of Internet Versus Paper-and-Pencil Versions (e25)
Sissel Kongsved, Maja Basnov, Kurt Holm-Christensen, Niels Hjollund. 40
- Impact of a Statewide Internet-Based Tobacco Cessation Intervention (e28)
Jessie Saul, Barbara Schillo, Sharrilyn Evered, Michael Luxenberg, Annette Kavanaugh, Nathan Cobb, Lawrence An. 50
- Toward a Model for Field-Testing Patient Decision-Support Technologies: A Qualitative Field-Testing Study (e21)
Rhodri Evans, Glyn Elwyn, Adrian Edwards, Eila Watson, Joan Austoker, Richard Grol. 61
- Stakeholder Perspectives on the Development of a Virtual Clinic for Diabetes Care: Qualitative Study (e23)
Natalie Armstrong, Hilary Hearnshaw, John Powell, Jeremy Dale. 73
- An Evaluation of Personal Health Information Remnants in Second-Hand Personal Computer Disk Drives (e24)
Khaled El Emam, Emilio Neri, Elizabeth Jonker. 80

Effect of Emailed Messages on Return Use of a Nutrition Education Website and Subsequent Changes in Dietary Behavior ([e27](#))

W Woodall, David Buller, Laura Saba, Donald Zimmerman, Emily Waters, Joan Hines, Gary Cutter, Randall Starling. 90

Review

Internet-Based Physical Activity Interventions: A Systematic Review of the Literature

Marleen H van den Berg¹, PhD; Johannes W Schoones², BA; Theodora PM Vliet Vlieland¹, MD, PhD

¹Department of Rheumatology, Leiden University Medical Center (LUMC), Leiden, The Netherlands

²Walaeus Library, LUMC, Leiden, The Netherlands

Corresponding Author:

Marleen H van den Berg, PhD

Leiden University Medical Center

Department of Rheumatology (C1-R)

PO Box 9600, 2300 RC

Leiden, The Netherlands

Phone: +31 71 526 3598

Fax: +31 71 526 6752

Email: mhvandenberg@lumc.nl

Abstract

Background: Nowadays people are extensively encouraged to become more physically active. The Internet has been brought forward as an effective tool to change physical activity behavior. However, little is known about the evidence regarding such Internet-based interventions.

Objective: The aim of the study was to systematically assess the methodological quality and the effectiveness of interventions designed to promote physical activity by means of the Internet as evaluated by randomized controlled trials.

Methods: A literature search was conducted up to July 2006 using the databases PubMed, Web of Science, EMBASE, PsycINFO, and Cochrane Library. Only randomized controlled trials describing the effectiveness of an Internet-based intervention, with the promotion of physical activity among adults being one of its major goals, were included. Data extracted included source and year of publication, country of origin, targeted health behaviors, participants' characteristics, characteristics of the intervention, and effectiveness data. In addition, the methodological quality was assessed.

Results: The literature search resulted in 10 eligible studies of which five met at least nine out of 13 general methodological criteria. The majority of the interventions were tailored to the characteristics of the participants and used interactive self-monitoring and feedback tools. Six studies used one or more theoretical models to compose the contents of the interventions. One study used an objective measure to assess the amount of physical activity (activity monitor), and six studies used multiple subjective measures of physical activity. Furthermore, half of the studies employed measures of physical fitness other than physical activity. In three studies, an Internet-based physical activity intervention was compared with a waiting list group. Of these three studies, two reported a significantly greater improvement in physical activity levels in the Internet-based intervention than in the control group. Seven studies compared two types of Internet-based physical activity interventions in which the main difference was either the intensity of contact between the participants and supervisors (4 studies) or the type of treatment procedures applied (3 studies). In one of these studies, a significant effect in favor of an intervention with more supervisor contact was seen.

Conclusions: There is indicative evidence that Internet-based physical activity interventions are more effective than a waiting list strategy. The added value of specific components of Internet-based physical activity interventions such as increased supervisor contact, tailored information, or theoretical fidelity remains to be established. Methodological quality as well as the type of physical activity outcome measure varied, stressing the need for standardization of these measures.

(*J Med Internet Res* 2007;9(3):e26) doi:[10.2196/jmir.9.3.e26](https://doi.org/10.2196/jmir.9.3.e26)

KEYWORDS

Physical activity; exercise; Internet; behaviour change; systematic review; consumer health informatics

Introduction

Regular physical activity is associated with lower morbidity and mortality rates from cardiovascular disease [1-4], diabetes mellitus [5], cancer [6], and osteoporosis [7]. Despite these proven health benefits, the majority of the adult population in Western nations does not meet the public health recommendations for physical activity [8-12]. Therefore, there is a need for the delivery of effective interventions aimed at positively influencing physical activity behavior.

Traditionally, most physical activity interventions use face-to-face modes of delivery (eg, individual consultations or group meetings). Their mainly short-term effectiveness has been extensively documented in a number of systematic reviews [13-18]. In addition, these reviews demonstrated that many of the physical activity studies suffer from several methodological weaknesses. The main methodological shortcomings identified by these reviews included use of physical activity measures without validity/reliability data, exclusive reliance on self-report measures, inadequate control of confounding factors, small sample sizes, lack of data on follow-up, and low follow-up rates.

With the number of people having access to and using the Internet rapidly increasing [19], the Internet is more and more used as a mode of delivery for physical activity programs. The strength of Internet-based physical activity interventions lies in the fact that with this mode of delivery large numbers of individuals can be reached at lower costs than with face-to-face interventions [20]. Moreover, by using the Internet, participants can access large amounts of information, and they can choose the time when they would like to interact and receive information [21]. Previous reviews on the effectiveness of Web-based physical activity interventions have indicated that the Internet can indeed serve as a promising mode of delivering physical activity interventions [20-24]. However, most of these reviews need to be updated as they comprised studies that were conducted between 2000 and 2003. This is all the more important as previous reviews included mainly observational and anecdotal studies, whereas a number randomized controlled trials have been published over recent years. Moreover, specific methodological characteristics of studies on physical activity interventions, such as the measurement of physical activity, have not yet been addressed in reviews that were exclusively aimed at Internet-based interventions.

The aim of this review is therefore to systematically assess both the methodological quality and the effectiveness of interventions designed to promote physical activity by means of the Internet as evaluated by randomized controlled trials.

Methods

Definitions

Physical activity and exercise represent different concepts: physical activity is defined as any bodily movement resulting in energy expenditure; exercise is a subset of physical activity that is planned, structured, repetitive, and aimed at improving or maintaining physical fitness [25]. Since exercise falls under

the broader concept of physical activity, in this paper we will use the term physical activity.

In addition, since email communication is based on Internet technology, both the use of websites and email will be designated as an Internet-based intervention.

Search Strategy

In cooperation with a trained librarian (JS), a search strategy was composed. The following databases were searched: PubMed (1949 to July 2006), Web of Science (1945 to July 2006), EMBASE (OVID-version, 1980 to July 2006), PsycINFO (1887 to July 2006), and Cochrane Library (1990 to July 2006). The search strategy consisted of the AND combination of three main concepts: Internet, physical activity, and intervention. For these three concepts, all relevant keyword variations were used, not only keyword variations in the controlled vocabularies of the various databases, but the free text word variations of these concepts as well. In general, the search consisted of the combination of the following terms: (1) internet or worldwideweb or world wide web or information technology or cyber* or web or website* or interactive or email or e-mail or e mail or emails or e-mails or e mails or emailing or e-mailing or e mailing or electronic mail; (2) physical education and training or exercise therapy or physical fitness or exercise or motor activity or physical training or physical education or fitness or exercise* or physical activity or physical activities or physical inactivity; and (3) intervention or interventions or intervention* or treatment outcome or intervention studies or epidemiologic study characteristics or study characteristics or epidemiologic methods or program or programs or programme or programmes or programmed or program evaluation.

This search strategy was optimized for all consulted databases, taking into account the differences of the various controlled vocabularies as well as the differences of database-specific technical variations (eg, different truncation symbols). Details of the database searches can be obtained from the author.

Selection of Articles

To be included, articles had to describe an intervention in which one of the primary goals was the promotion of physical activity among adults (18 years or older). Furthermore, the intervention had to be delivered predominantly by means of the Internet in one of the following ways: (1) exchange of information via the World Wide Web between a health care setting and an individual (eg, between a clinic and a participant's home or workplace), (2) use of email for communication between a therapist or health care professional and a patient (or patient group). Internet-based physical activity interventions that promoted physical activity in order to achieve a secondary goal, such as weight reduction, were also included.

Only randomized controlled trials with pretest and posttest outcome data for both the control and intervention groups were considered for inclusion in this review. No restrictions were defined regarding the type and contents of the control group: this could be assignment to a waiting list, a non-Internet-based intervention, or a different type of Internet-based intervention. At least one of the outcomes had to be described in terms of change in physical activity level (eg, change in amount or

quantity of physical activity). Furthermore, because of limited resources for translation, this review was restricted to publications in English, Dutch, and German.

The reference lists of the selected articles were checked for additional eligible articles, using the same inclusion criteria. Review articles could not be included in the review; however, the reference lists of relevant review articles were also checked for additional eligible articles. The articles were independently selected and assessed by two reviewers (MvdB and TVV).

Assessment of Methodological Quality

With respect to the guidelines for evaluating methodological quality of intervention studies, the literature does not provide a gold standard. We used a list of criteria recommended by Van Tulder et al [26], which has proven to be appropriate in other reviews evaluating physical activity or exercise interventions [27,28]. This list was based on the guidelines for systematic reviews as set by the editorial board of the Cochrane Collaboration Back Review Group, which address the main steps in conducting a systematic review: literature search, inclusion criteria, methodological quality, data extraction, and data analysis. The list of Van Tulder et al contains 19 methodological criteria. The criteria “care provider blinded,” “patient blinded,” “co-interventions avoided,” and “description of adverse effects” were not regarded as being suitable or relevant by the reviewers because of the character of the interventions and were removed from the list. The criteria “relevant outcome measures” and “short-term follow-up outcome” were already used as inclusion criteria for articles in this review; therefore, these criteria were not used for assessing methodological quality. Finally, the criterion “acceptable compliance” was reformulated as “description of compliance,” and “description of and acceptable dropout rate” was reformulated as “description of dropout rate plus comparison of dropouts with completers.” The final number of criteria used to assess methodological quality was 13 (see the Multimedia Appendix). All criteria were scored as “yes,” “no,” or “unclear.” Equal weight was applied to all criteria, resulting in a methodological summary score ranging from 0 to 13. The literature provides no guidelines for choosing cutoff points in order to rate the methodological quality [29]. In this review, we rated the studies as having good methodological quality if two thirds or more of the criteria were met (ie, a summary score of 9 or higher).

In addition, we evaluated the studies included in this review with respect to quality criteria that apply to physical activity interventions and Internet-based interventions in particular. These criteria were derived from previous literature on physical activity assessment in general [30] and on evaluation methods of Internet-based behavioral interventions [31,32] and comprised the following:

- Intervention-related: (1) tailoring of program to participants’ characteristics, (2) use of interactive self-monitoring and feedback, (3) theoretical fidelity (degree to which interventions follow their planned procedures or theoretical models)
- Process-related: (4) information on use of intervention tools or facilities

- Outcome-related: (5) use of a combination of physical activity measurements (rather than one measure), (6) use of objective methods of data collection, such as activity monitors, heart rate monitors, pedometers, direct observation, or doubly labelled water, (7) use of additional fitness-related outcomes

All quality criteria were scored as “yes,” “no,” or “unclear.”

Data Extraction

This review is a qualitative systematic review as the data extracted from the selected studies were summarized but not statistically combined. Aggregating findings across studies rather than pooling them was a more useful method of describing synthesis, as the outcome measures varied widely. The results of the selected studies were broken down, thoroughly analyzed, and then combined into a whole via a listing of themes. This has proven to be a suitable method for systematic reviews [33].

The following information was systematically extracted by the two reviewers: source and year of publication, country of origin, targeted health behaviors (physical activity, weight loss, nutrition behavior, or other), characteristics of the study population (number and type of participants, age, gender), characteristics of the intervention (duration, theoretical foundation, description of contents), and pretest and posttest physical activity outcomes of both intervention groups. With respect to the changes in physical activity level, only the posttest results measured directly after finishing the physical activity intervention were included.

In order to be able to make more valid comparisons, the selected studies are divided into three categories: section A contains studies in which Internet-based physical activity interventions were compared with a waiting list or an attention-control group; section B contains studies in which two types of Internet-based physical activity interventions were compared that mainly differed with respect to the amount or frequency of contact between the participants and supervisors; in section C, two types of Internet-based physical activity interventions were compared; however, in these studies, the two interventions varied with respect to the applied treatment procedures.

Reviewers were blinded to authorship, journal title, and other study-related information. Furthermore, screening for eligible articles as well as data extraction from the selected articles were done independently. Any discrepancies between the two reviewers were settled by consensus.

Results

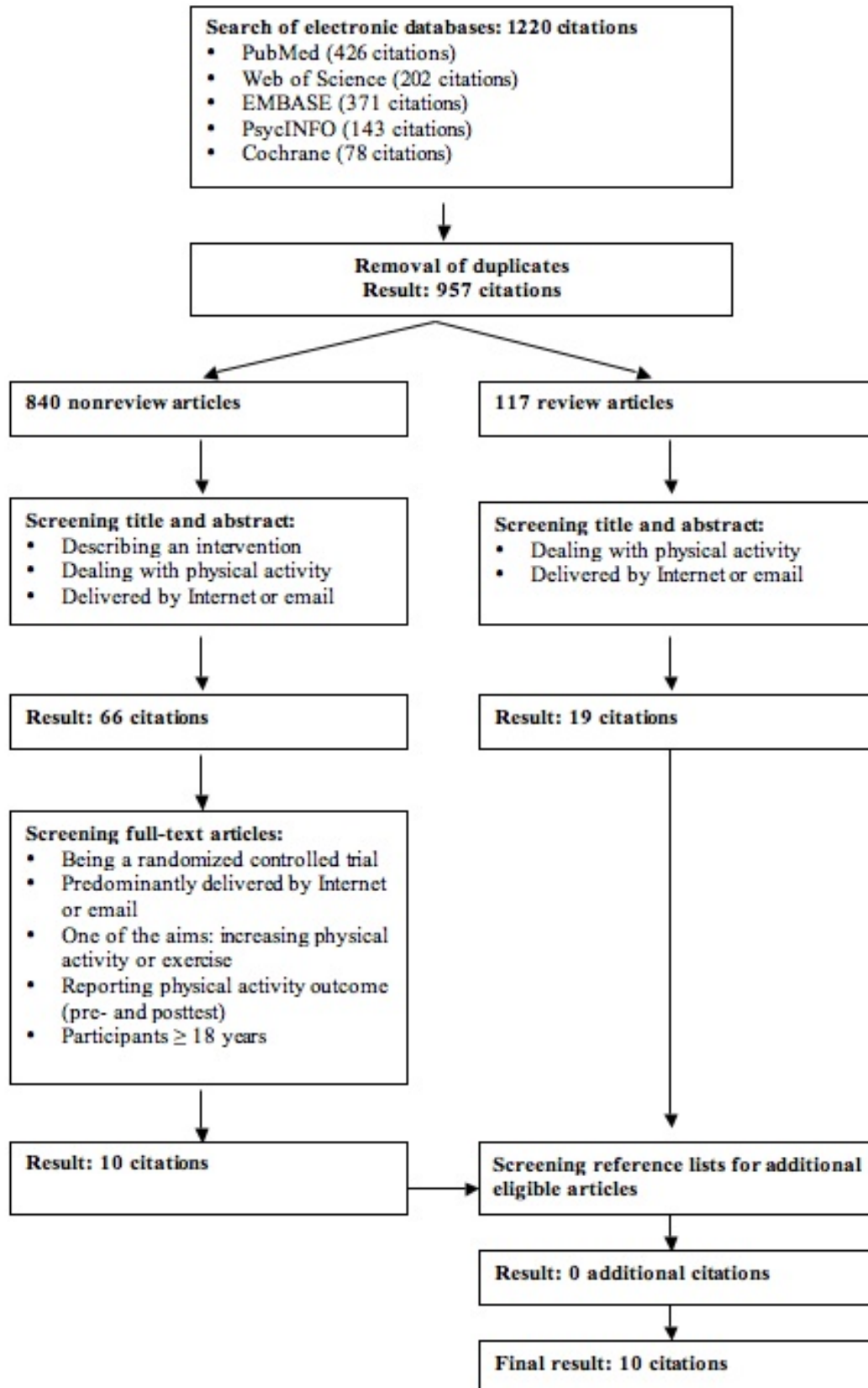
Selection of Articles

Figure 1 illustrates the search and selection process. The initial database search yielded 1220 citations. After eliminating duplicates, this was reduced to 957 citations, of which 117 were review articles. Screening titles and abstracts of the 840 nonreview articles resulted in 66 citations potentially meeting eligibility criteria. After completely reviewing the corresponding full-text articles, the total number of articles was reduced to 10. Reasons for exclusion of the other 56 citations were not reporting pretest and posttest physical activity outcomes (n =

25), intervention not predominantly delivered by the Internet (n = 16), not being a randomized controlled trial (n = 13), and participants being younger than 18 years (n = 2). Screening the titles and abstracts of the 117 review articles resulted in 19

relevant reviews. Screening both the reference list of these reviews, as well as the reference lists of the 10 selected articles, did not bring up any additional articles. As a result, 10 articles were included.

Figure 1. Article search and selection process



Assessment of Methodological Quality

Results of the methodological assessment are described in Table 1. Five studies met nine or more criteria [32,34-37], implying a good methodological quality. One study described the method of random assignment and stated that this assignment was performed by an independent person [37]. Information about the blinding of the outcome assessor was given in two studies [34,37]. None of the studies performed a full intention-to-treat analysis according to the definition of intention-to-treat given

by Hollis and Campbell [38], stating that “a full application of intention-to-treat is possible only when complete outcome data are available for all randomised subjects.” All studies reported a dropout rate, with six of the 10 studies comparing the characteristics of these dropouts with the subjects that completed all outcome measurements [32,35,36,39-41]. In two studies [41,42], the study sample included only those participants who completed both the baseline as well as the follow-up measurements, excluding dropouts from the analysis.

Table 1. Methodological quality of the studies

	Kosma et al [41]	Plotnikoff et al [42]	Napolitano et al [39]	Marshall et al [34]	Van den Berg et al [37]	Hageman et al [43]	Rovniak et al [32]	Tate et al [36]	McKay et al [40]	Tate et al [35]
Specification of eligibility criteria	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Description of randomization method	no	no	no	yes	yes	no	yes	yes	no	no
Random assignment performed by independent person	unclear	unclear	unclear	unclear	yes	unclear	unclear	unclear	unclear	unclear
Groups similar at baseline	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Sufficient description of interventions	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Description of compliance with interventions	no	no	yes	yes	yes	yes	yes	yes	yes	yes
Blinding of outcome assessor	unclear	unclear	unclear	yes	yes	unclear	no	unclear	unclear	unclear
Description of dropout rate plus comparison of dropouts and completes	yes	no	yes	no	no	no	yes	yes	yes	yes
Outcome assessment ≥ 6 months after randomization	no	no	no	no	yes	no	yes	yes	no	yes
Timing of assessments comparable	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Description of sample size calculation	no	no	no	yes	yes	yes	no	no	no	yes
Intention-to-treat analysis	no	no	no	no	no	no	no	no	no	no
Presentation of point estimates and variability measures	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Total number of criteria fulfilled	6	5	7	9	11	7	9	9	7	9

Concerning the quality criteria that apply to physical activity interventions and Internet-based interventions in particular, the results show that in the majority of the studies the interventions were tailored to the characteristics of the participants and used interactive self-monitoring and feedback tools (Table 2). Six studies used one or more theoretical models to compose the information delivered to the intervention group [39-42] or to both the intervention and control groups [32,34]. These models were the Transtheoretical Model [34,39,41,42], the Protection

Motivation Theory [42], the Theory of Planned Behavior [42], the Social Cognitive Theory [32,39,42], and a social-ecological model [40].

With respect to outcome measurement, one study [37] used an objective measure to assess the amount of physical activity (activity monitor), and six studies used multiple subjective measures of physical activity [34,37,39,40,42,43]. Half of the studies employed measures of physical fitness other than physical activity [32,35-37,43].

Table 2. Characteristics of intervention, process, and outcome measures of the studies

	Kosma et al [41]	Plotnikoff et al [42]	Napolitano et al [39]	Marshall et al [34]	Van den Berg et al [37]	Hageman et al [43]	Rovniak et al [32]	Tate et al [36]	McKay et al [40]	Tate et al [35]
Intervention										
Program tailored to participants' characteristics	no	no	yes	yes	yes	yes	yes	yes	yes	yes
Use of interactive self-monitoring and feedback	no	no	yes	yes	yes	no	yes	yes	yes	yes
Intervention developed according to theoretical guidelines	yes	yes	yes	yes	no	no	yes	no	yes	no
Process										
Use of intervention tools/facilities	no	no	no	yes	yes	yes	yes	yes	yes	yes
Outcome										
Use of combination of physical activity assessment measures	no	yes	yes	no	yes	yes	no	no	no	no
Use of objective physical activity assessment methods	no	no	no	no	yes	no	no	no	no	no
Use of additional physical fitness-related outcomes	no	no	no	no	yes	yes	yes	yes	no	yes

Data Extraction

Characteristics of Selected Studies

Study characteristics are described in [Table 3](#). Seven of the 10 selected studies were performed in the United States, one in Canada, one in Australia, and one in The Netherlands. All

studies were published between 2001 and 2006. Three studies addressed interventions targeted at both physical activity and nutrition behavior; the other seven studies focused on interventions aimed at physical activity behavior only. The duration of the interventions varied from 1 to 12 months, with three studies describing interventions of 6 months or longer [35-37].

Table 3. Characteristics of studies and participants*

Study	Targeted Health Behavior	Duration of Intervention (months)	Sample Description		Type of Participants	Gender (% male)	Age (mean \pm SD; years)
			No. of Participants Randomized	No. of Participants With Complete Data			
Kosma et al [41], 2005, USA	PA	1	151 (I: 101, C: 50)	75 (I: 46, C: 29)	Inactive adults with physical disabilities, with Internet access	21	38.7 \pm 8.9
Plotnikoff et al [42], 2005, Canada	PA and nutrition behavior	3	2598 (I: ?, C: ?)	2121 (I: 1566, C: 555)	Employees of large workplaces with Internet and email access	26	44.9 \pm 6.3
Napolitano et al [39], 2003, USA	PA	3	65 (I: 30, C: 35)	52 (I: 21, C: 31)	Hospital employees participating in \leq 120 min of moderate PA/week or \leq 60 min of vigorous PA/week, with Internet and email access	14	42.8 \pm 10.0
Marshall et al [34], 2003, Australia	PA	2	655 (I: 327, C: 328)	512 (I: 250, C: 262)	University employees with email access	49	43 \pm 11
Van den Berg et al [37], 2006, The Netherlands	PA	12	160 (I: 82, C: 78)	152 (I: 77, C: 75)	Patients with rheumatoid arthritis not participating in 30 min of moderate PA on \geq 5 days/week, with Internet and email access	24	49.6 \pm 10.3
Hageman et al [43], 2005, USA	PA	2	31 (I: 15, C: 16)	30 (I: ?, C: ?)	Healthy women not participating in 30 min of moderate PA on \geq 5 days/week, with Internet access	0	56.1 \pm 4.9
Rovniak et al [32], 2005, USA	PA	3	61 (I: 30, C: 31)	50 (I: 25, C: 25)	Sedentary adult women participating in $<$ 90 min of PA/week, with email access	0	40.2 \pm 9.1
Tate et al [36], 2003, USA	Weight loss (PA and nutrition)	12	92 (I: 46, C: 46)	77 (I: 38, C: 39)	Overweight (BMI 27-40 kg/m ²) adults at risk of type 2 diabetes, with Internet and email access	10	48.5 \pm 9.4
McKay et al [40], 2001, USA	PA	2	78 (I: 38, C: 40)	68 (I: 35, C: 33)	Type 2 diabetic patients not participating in 30 min of moderate PA on \geq 5 days/week, with Internet and email access	47	52.3 \pm ?
Tate et al [35], 2001, USA	Weight loss (PA and nutrition)	6	91 (I: 46, C: 45)	71 (I: 36, C: 35)	Overweight (BMI 25-36 kg/m ²) adult hospital employees with Internet and email access	11	40.9 \pm 10.6

*PA indicates physical activity; I, intervention group; C, control group; ?, unknown; BMI, body mass index.

Characteristics of Study Populations

Table 3 shows that the total population size varied from 31 to 2598 participants. The study populations all consisted of healthy (overweight) adults, except for the studies of Kosma et al [41], McKay et al [40], and Van den Berg et al [37], which included physically disabled patients, diabetic patients, and patients with rheumatoid arthritis, respectively. Six of the 10 studies were specifically targeted at adults who were sedentary at baseline [32,37,39-41,43]; the other four studies did not employ any inclusion criteria regarding baseline physical activity level [34-36,42]. In two studies [34,40], the proportion of male and female participants was almost equal; in the other studies, the large majority of participants were female. Mean age varied from 39 to 56 years.

Characteristics of the Interventions

Table 4 describes the characteristics of the Internet-based physical activity programs and control conditions.

Section A of Table 4 describes the three studies in which an Internet-based physical activity intervention was compared a waiting list group [39,41,42]. In two of these studies [39,41], the participants in the Internet-based intervention had access to a website and received emails; in the other study, the intervention group received only emails [42].

Section B of Table 4 describes the four studies that compared two types of Internet-based physical activity intervention with the main difference being the amount or frequency of contact between the participants and supervisors [35-37,40]. These studies investigated whether more intensive supervision would

lead to a greater increase in physical activity level. In three studies, the difference in the amount of supervisor contact was, in fact, a difference in the degree of tailoring or personalization [35,37,40], in which participants from the intervention group had access to a website and received emails, whereas the control group had website access only. In the other study [36], website access and email communication was offered to participants in both the intervention and control groups.

Section C of Table 4 describes the three studies that compared two types of Internet-based physical activity intervention in which the main difference was the treatment procedures that were used, whereas the amount of contact between the participants and supervisors did not differ. One study [32] investigated whether precision in replicating theory-based recommendations influenced the effectiveness of an Internet-based physical activity intervention. In the second study [34], the means by which the physical activity interventions were delivered differed (print-based versus Web-based). The third study was designed to explore the net effect of tailored versus standard information.

Eight studies aimed to increase any type of physical activity, whereas two studies were specifically targeted at walking [32] or cycling on a bicycle ergometer [37].

Effectiveness of Intervention

The physical activity outcome measures of both the intervention and control groups are expressed as pretest and posttest results and are described in Table 4. Four studies included one physical

activity outcome parameter [32,35,36,41], five studies included two physical activity parameters [34,39,40,42,43], and one study reported more than two physical activity parameters [37]. Five of the 10 selected studies reported additional physical fitness-related outcome measures such as cardiorespiratory fitness, flexibility, and body weight [32,35-37,43]. In three of these five studies [32,35,36], the reported changes in physical activity level were considered a secondary outcome; primary outcomes in these studies were changes in body weight and waist circumference [35,36], cardiorespiratory fitness, and walking speed [32].

Regarding the four studies described in Section A of Table 4, in which Internet-based interventions were compared with a waiting list, two studies reported significant differences between the intervention and control groups [39,42]. With respect to the four studies described in Section B of Table 4, in which the intensity of contact in two types of Internet-based physical activity intervention varied, one study reported significant differences between the intervention and control groups with respect to change in physical activity level [37]. Two of the four studies [35,36] in Section B were not primarily aimed at increasing physical activity level, but rather to decrease body weight and waist circumference.

The changes in physical activity level were all nonsignificant in the three studies in which the applied treatment procedures of two Internet-based physical activity interventions varied (Table 4, Section C). This section comprised one study in which physical activity was not the primary outcome measure [32].

Table 4. Characteristics and results of the Internet-based physical activity interventions*

Study	Description of Intervention Group	Description of Control Group	PA Outcome Measures [†]		PA Post-test Results (mean ± SD)	Additional Fitness-Related Outcomes	Conclusion
			Type of PA Outcome Variable	PA Pre-test Results (mean ± SD)			
Section A: comparison of an Internet-based physical activity intervention with a waiting list or attention-control group							
Kosma et al [41], 2005	Weekly emails containing a Web link to motivational PA lesson plans; opportunity to participate in Web-based discussion board, for half of intervention group	Weekly emails containing messages not related to PA	Leisure time PA (MET hours/day)	I: 6.1 ± 7.4 C: 9.3 ± 7.7	I: 8.2 ± 6.8 C: 6.9 ± 7.8	–	No significant between-group differences for leisure time PA
Plotnikoff et al [42], 2005	Weekly emails containing PA information operationalizing social-cognitive items and beliefs predicting PA behavior and links to other websites about PA and healthy eating	No weekly emails (nothing)	Moderate and vigorous PA (MET min/week) Workplace activity status (1 = sedentary to 4 = very active)	PA: I: 664.1 ± 726.1 C: 668.6 ± 752.6 Workplace status: I: 1.3 ± 0.6 C: 1.3 ± 0.5	PA: I: 683.7 ± 702.3 C: 592.7 ± 652.8 Workplace status: I: 1.4 ± 0.6 C: 1.4 ± 0.6	–	Significant between-group differences for moderate and vigorous PA, not for workplace status
Napolitano et al [39], 2003	Access to stage-based PA website containing the following sections: activity quiz, safety tips, becoming active, PA and health, overcoming barriers, planning PA, and benefits of PA Weekly tip sheets sent by email containing PA-related information about getting started, monitoring progress, setting goals, rewarding, and support Opportunity to contact helpline by email or telephone in case of questions, concerns, or problems	Waiting list	Moderate intensity PA (min/week) Walking (min/week)	Moderate PA: I: 68.8 ± 58.1 C: 80.9 ± 77.8 Walking: I: 57.2 ± 56.9 C: 87.6 ± 177.4	Moderate PA: I: 112.0 ± 75.7 C: 82.0 ± 87.3 Walking: I: 99.8 ± 68.3 C: 68.4 ± 85.2	–	Significant between-group differences for moderate intensity PA and walking
Section B: comparison of two types of Internet-based physical activity interventions that differ with respect to amount of contact between the participants and supervisors							
Van den Berg et al [37], 2006	Access to website containing a personalized PA program consisting of weekly personalized physical activity schedules with weekly personalized feedback provided by physical therapist Access to online discussion forum to contact other participants Access to face-to-face group meetings every 3 months A bicycle ergometer was given on loan during intervention period	Access to website containing general PA information, which was updated once a month Opportunity to order free copy of PA-related CD-ROM	Moderate PA (% patients meeting moderate PA recommendations) Vigorous PA (% patients meeting vigorous PA recommendations)	Moderate proportions: I: 0 C: 0 Vigorous proportions: I: 6 C: 1	Moderate %: I: 26 C: 15 Vigorous %: I: 34 C: 10	Functional ability	Significant between-group differences for vigorous PA, not for moderate PA

Study	Description of Intervention Group	Description of Control Group	PA Outcome Measures [†]			Additional Fitness-Related Outcomes	Conclusion
			Type of PA Outcome Variable	PA Pre-test Results (mean ± SD)	PA Post-test Results (mean ± SD)		
Tate et al [36], 2003	<p>One introductory face-to-face group weight loss session (1 hour) in which instructions regarding weight loss and increasing PA levels were given by clinical therapist</p> <p>Access to educational website containing information about weight loss, including tips, links, and other Internet resources</p> <p>Instructions to report dietary and PA self-monitoring information weekly by means of website diary</p> <p>5 emails per week sent by therapist in the first month, weekly emails for remaining 11 months; emails contained personalized feedback, recommendations, reinforcements, answers to participants' questions, and general support</p>	<p>One introductory face-to-face group weight loss session (1 hour) in which instructions regarding weight loss and increasing PA levels were given by clinical therapist</p> <p>Access to educational website containing information about weight loss including tips, links and other Internet resources</p> <p>Encouragement to use online dietary and PA self-monitoring tools</p> <p>Weekly email reminders sent by therapist to submit self-monitoring data</p>	Exercise energy expenditure (kcal/week) [‡]	I: 886 ± 832 C: 803 ± 1015	I: 342 ± 945 [§] C: 63 ± 1211 [§]	Body weight and waist circumference	No significant between-group differences for exercise energy expenditure
McKay et al [40], 2001	<p>Access to website containing a personalized PA program based on baseline online assessment of PA level; PA program consisted of personalized goal setting, activity selection, scheduling PA, overcoming barriers</p> <p>Access to personal PA database containing additional PA-related information and PA logs with graphs of progress</p> <p>Provision of personalized counseling and support provided by a personal coach by means of online messages</p> <p>Access to peer-to-peer support groups</p>	<p>Access to website containing diabetes specific articles plus real-time blood glucose tracking with graphic feedback</p>	<p>Moderate-to-vigorous intensity exercise (min/day)</p> <p>Walking (min/day)</p>	<p>Exercise: I: 5.6 ± 6.2 C: 7.3 ± 6.2</p> <p>Walking: I: 6.4 ± 6.2 C: 8.4 ± 8.4</p>	<p>Exercise: I: 17.6 ± 15.3 C: 18.0 ± 17.3</p> <p>Walking: I: 12.5 ± 9.5 C: 16.8 ± 22.8</p>	–	No significant between-group differences for moderate-to-vigorous intensity exercise or walking

Study	Description of Intervention Group	Description of Control Group	PA Outcome Measures [†]			Additional Fitness-Related Outcomes	Conclusion
			Type of PA Outcome Variable	PA Pre-test Results (mean ± SD)	PA Post-test Results (mean ± SD)		
Tate et al [35], 2001	<p>One introductory face-to-face group weight loss session (1 hour) in which instructions regarding weight loss and increasing PA levels were given by clinical therapist</p> <p>Access to educational website containing information about weight loss, such as diet, exercise, self-monitoring, social support, stimulus control, and managing stress</p> <p>A brief 15 min face-to-face check-in with therapist every 3 months</p> <p>Instructions to report dietary and PA self-monitoring information weekly by means of website diary</p> <p>Weekly emails sent by therapist containing a behavioral weight loss lesson, personalized feedback, recommendations, reinforcements, answers to participants' questions, and general support</p> <p>Access to electronic bulletin board</p>	<p>One introductory face-to-face group weight loss session (1 hour) in which instructions regarding weight loss and increasing PA levels were given by clinical therapist</p> <p>Access to educational website containing information about weight loss, such as diet, exercise, self-monitoring, social support, stimulus control, and managing stress</p> <p>A brief 15 min face-to-face check-in with therapist every 3 months</p> <p>Encouragements to use online dietary and PA self-monitoring tools</p>	Exercise energy expenditure (kcal/week) [‡]	I: 1360 ± 1415 C: 1031 ± 981	I: 1289 ± 919 C: 1125 ± 1320	Body weight and waist circumference	No significant between-group differences for exercise energy expenditure

Section C: comparison of two types of Internet-based physical activity interventions that differ with respect to the applied treatment procedures

Hageman et al [43], 2005	<p>One initial face-to-face assessment of behavioral markers and biomarkers</p> <p>Three online newsletters containing individually tailored information about PA goals, benefits, and barriers to PA and self-efficacy delivered by Internet every month</p>	<p>One initial face-to-face assessment of behavioral markers and biomarkers</p> <p>Three online newsletters containing general information about PA goals, benefits, and barriers to PA and self-efficacy delivered by Internet every month</p>	<p>Moderate or vigorous PA (min/week)</p> <p>Energy expenditure (kcal/kg/day)</p>	<p>PA:</p> <p>I: 937.6 ± 616.5 C: 1228.1 ± 119.7</p> <p>Expenditure:</p> <p>I: 28.7 ± 5.0 C: 28.9 ± 5.7</p>	<p>PA:</p> <p>I: 672.5 ± 643.9 C: 906.0 ± 775.8</p> <p>Expenditure:</p> <p>I: 26.5 ± 5.0 C: 27.3 ± 4.6</p>	Cardiorespiratory fitness, flexibility, body composition	No significant between-group differences for moderate or vigorous PA or energy expenditure
--------------------------	---	---	---	---	--	--	--

Study	Description of Intervention Group	Description of Control Group	PA Outcome Measures [†]			Additional Fitness-Related Outcomes	Conclusion
			Type of PA Outcome Variable	PA Pre-test Results (mean ± SD)	PA Post-test Results (mean ± SD)		
Rovniak et al [32], 2005	One 30 min face-to-face session providing information about walking plus modeling of 3 walking skills Specific and tailored email-based walking prescription by supervisor Immediate and precise self-monitoring of walking information by participants by means of online walking logs Weekly specific feedback by supervisor about walking performance relative to past accomplishments and normative standards sent by email	One 30 min face-to-face session only providing information about walking General email-based walking prescription by supervisor General self-monitoring of walking information by participants by means of online walking logs Weekly general feedback sent by supervisor about walking performance	Walking (min/week) [‡]	I: 17.5 ± 20.9 C: 16.4 ± 24.8	I: 74.5 ± 49.9 C: 61.2 ± 38.8	Cardiorespiratory fitness, walking speed, body mass index	No significant between-group differences for walking time
Marshall et al [34], 2003	Access to a stage-targeted PA website containing stage-based quizzes with feedback, personalized sections on goal setting, activity planning, targeted heart rates, and a PA readiness questionnaire Personalized reinforcement emails sent every 2 weeks containing stage-targeted PA information and links to study website	Stage-targeted printed booklets sent by postal mail containing PA information based on Trans-theoretical Model of Behavior Change Additional printed reinforcement letters sent by postal mail every 2 weeks containing stage-targeted PA information	Total amount of PA (MET min/week) Total amount of sitting (MET min/week)	PA: [¶] I: 2425 ± 113 C: 2413 ± 115 Sitting time: [¶] I: 2263 ± 57 C: 2221 ± 56	PA: [¶] I: 2433 ± 121 C: 2518 ± 115 Sitting time: [¶] I: 2158 ± 48 C: 2150 ± 49	–	No significant between-group differences for PA and sitting time

*PA indicates physical activity; I, intervention group; C, control group; MET, metabolic equivalent

[†]PA outcome measures are outcomes that measure (changes in) the amount of physical activity.

[‡]Physical activity outcome variable in this study was considered a secondary outcome.

[§]Values of posttest data represent change scores (mean ± SD).

^{||}Posttest data not measured directly after the intervention (1 month after sending last newsletter).

[¶]Values of pre- and posttest data represent mean ± SE.

Discussion

The number of randomized controlled trials on the effectiveness of Internet-based physical activity interventions is limited. This review represents the best available evidence so far. Two investigators independently assessed all articles and abstracts, and consensus was reached concerning both the inclusion of the studies and the data extraction.

Three studies were identified that investigated whether an Internet-based physical activity intervention was more effective than a waiting list. Two of these studies reported a significantly greater increase in physical activity in the Internet-based

intervention than in the waiting list group. However, the effect sizes, which were reported in only one of these two studies, were small, indicating that the clinical relevance remains questionable.

In four studies, two types of Internet-based intervention were compared in which the most important difference between the intervention and control groups was the amount of contact with the supervisors. Of these studies, only one reported significant differences between the two interventions with respect to change in physical activity level. However, in this study, the amount of personalized supervision was not the only difference between the intervention and control groups. As opposed to the

participants from the control group, participants from the intervention group also received a bicycle ergometer and were offered peer-to-peer group contacts. Therefore, it could not be established if the increased amount of contact caused the increase in effectiveness. None of the three studies in which different types of treatment procedures of two Internet-based physical activity interventions were compared reported significant differences.

The methodological quality of the selected studies in this review varied. Only half of the 10 studies were rated as having a good methodological quality. Lack of information about blinding of the outcome assessor, no description of sample size calculation, and insufficient description of the randomization and concealment method were the most important reasons for low scores on methodological quality. This may have influenced the results of these studies since it has been shown that inadequate methodological approaches in controlled trials, particularly those representing poor allocation concealment, are associated with bias [44]. Furthermore, none of the studies applied an intention-to-treat analysis. However, a full application of the intent-to-treat model according to the definition given by Hollis and Campbell [38] may not be possible for most physical activity studies because, in most of these studies, there will be at least some subjects who drop out, refuse to complete final assessments, or change residence.

In addition, we evaluated the quality of the studies by assessing whether or not the interventions fulfilled criteria that apply to Internet-based physical activity interventions in particular, including intervention measures, process measures, and outcome measures. It was shown that in six studies the researchers used one or more theoretical models to compose the interventions. The Transtheoretical Model and the Social Cognitive Theory were the two most frequently used theories. This review could not demonstrate that theory-based physical activity interventions conducted through the Internet are more effective than non-theory-based interventions. Although there is some evidence that interventions in which these models are incorporated are effective in increasing physical activity level [45-47], other researchers still question this effectiveness [48]. Further research on the surplus value of these models in promoting complex health behavior such as physical activity is needed.

Furthermore, the results show that most of the studies used a single physical activity outcome measure, and objective measures such as activity monitors or pedometers were rarely used. In order to be able to better establish the effect of Internet-based physical activity interventions, future studies should incorporate multiple physical activity outcomes, preferably accompanied by one or two objective measures. Moreover, there is a need for more uniform physical activity outcome measures; in our review, studies reported their outcomes in time, energy expenditure, or categorical variables such as proportions of persons meeting physical activity recommendations.

On the basis of the above-mentioned results of this review, we conclude that there is indicative evidence that Internet-based physical activity interventions are more effective than a waiting list group. With respect to which components serve as the key

components (ie, amount of contact or type of treatment procedure), the evidence is scanty.

Several factors may have contributed to the limited evidence of effectiveness. First, the number of eligible studies was limited. The Internet is a relatively new tool for delivering physical activity interventions. Moreover, many of the interventions that did use the Internet for program delivery did not report their outcomes in terms of changes in physical activity level, but used indirect measures such as stages of motivational readiness, weight change, heart rate, or maximal oxygen uptake. Our review included three studies in which the changes in physical activity level were considered secondary outcomes; these interventions were not primarily aimed at changing physical activity behavior. These three studies all compared two different types of Internet-based intervention.

Second, this review comprised mainly short-term physical activity interventions. Only three studies incorporated interventions of 6 months or longer. The literature suggests that long-term changes in physical activity behavior can only be accomplished by studies with long-term follow-up [18]. However, no guidelines exist regarding the optimal duration of interventions. Therefore, more research should be done to evaluate the minimal duration of physical activity interventions in order to produce long-term physical activity behavior change.

Third, the baseline physical activity levels of the participants differed, making it difficult to report on the overall effectiveness of these interventions. Moreover, four studies in this review did not report any baseline physical activity levels. Since physically active persons in general are better able to comply with physical activity interventions and maintain a healthy lifestyle than sedentary persons [49-51], incomplete or inconsistent information about baseline physical activity levels may have influenced our results.

A final limitation is the fact that the contents of the control intervention differed widely. In some studies, participants from the control group received more general or standard versions of the Internet-based physical activity intervention; in other studies, these participants received a print-based version of the intervention or were assigned to a waiting list. The exact surplus value of adding personalized supervision to an Internet-based physical activity intervention could not be established because, in most studies, in addition to this supervision, other components were added as well. The two trials that compared the Internet-based physical activity intervention with a waiting list both reported significant differences between the intervention and control groups. This may indicate that, when trying to increase people's physical activity level, providing an Internet-based physical activity intervention is more effective than doing little or nothing. However, more studies are needed to establish this conclusion. With respect to determining the effectiveness of different components of an Internet-based physical activity intervention, more studies are needed that use appropriate research designs (ie, designs in which the only difference between the intervention and control groups is the addition of a specific component, such as providing personalized supervision).

In conclusion, the methodological quality as well as the type of physical activity outcome measure of Internet-based physical activity interventions varied. However, Internet-based physical activity interventions appear to be more effective when compared to a waiting list strategy. Whether or not adding specific components to Internet-based physical activity interventions will result in greater effectiveness compared to Internet-based interventions in which these components are missing or offered less intensely remains to be established. An

important advantage of Internet-based interventions is that they can reach large numbers of people at relatively low cost. However, more cost-effectiveness studies should be done in order to establish the exact surplus value of this delivery method when compared with more traditional methods such as face-to-face sessions. Moreover, future research should properly define the control groups and incorporate both long-term as well uniform and objective physical activity outcome measures.

Conflicts of Interest

None declared.

Multimedia Appendix

Criteria of Methodological Quality

1. Were the eligibility criteria specified?
2. Was the method of randomization described?
3. Was the random allocation concealed? (ie, Was the assignment generated by an independent person not responsible for determining the eligibility of the patients?)
4. Were the groups similar at baseline regarding important prognostic indicators?
5. Were both the index and the control interventions explicitly described?
6. Was the compliance or adherence with the interventions described?
7. Was the outcome assessor blinded to the interventions?
8. Was the dropout rate described and were the characteristics of the dropouts compared with the completers of the study?
9. Was a long-term follow-up measurement performed (outcomes measured ≥ 6 months after randomization)?
10. Was the timing of the outcome measurements in both groups comparable?
11. Was the sample size for each group described by means of a power calculation?
12. Did the analysis include an intention-to-treat analysis?
13. Were point estimates and measures of variability presented for the primary outcome measures?

References

1. Berlin JA, Colditz GA. A meta-analysis of physical activity in the prevention of coronary heart disease. *Am J Epidemiol* 1990 Oct 1;132(4):612-628. [Medline: [2144946](#)]
2. Kaplan GA, Strawbridge WJ, Cohen RD, Hungerford LR. Natural history of leisure-time physical activity and its correlates: associations with mortality from all causes and cardiovascular disease over 28 years. *Am J Epidemiol* 1996 Oct 15;144(8):793-797 [FREE Full text] [Medline: [8857828](#)]
3. Sesso HD, Paffenbarger RS, Lee IM. Physical activity and coronary heart disease in men: The Harvard Alumni Health Study. *Circulation* 2000 Aug 29;102(9):975-980 [FREE Full text] [Medline: [10961960](#)]
4. Lee IM, Rexrode KM, Cook NR, Manson JE, Buring JE. Physical activity and coronary heart disease in women: is "no pain, no gain" passé? *JAMA* 2001 Mar 21;285(11):1447-1454 [FREE Full text] [Medline: [11255420](#)] [doi: [10.1001/jama.285.11.1447](#)]
5. Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *N Engl J Med* 1991 Jul 18;325(3):147-152. [Medline: [2052059](#)]
6. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. *Ann Intern Med* 1995 Mar 1;122(5):327-334 [FREE Full text] [Medline: [7847643](#)]
7. Bravo G, Gauthier P, Roy PM, Payette H, Gaulin P, Harvey M, et al. Impact of a 12-month exercise program on the physical and psychological health of osteopenic women. *J Am Geriatr Soc* 1996 Jul;44(7):756-762. [Medline: [8675921](#)]
8. Jones DA, Ainsworth BE, Croft JB, Macera CA, Lloyd EE, Yusuf HR. Moderate leisure-time physical activity: who is meeting the public health recommendations? A national cross-sectional study. *Arch Fam Med* 1998;7(3):285-289 [FREE Full text] [Medline: [9596466](#)] [doi: [10.1001/archfam.7.3.285](#)]
9. ; Centers for Disease Control and Prevention (CDC). Physical activity trends--United States, 1990-1998. *MMWR Morb Mortal Wkly Rep* 2001 Mar 9;50(9):166-169 [FREE Full text] [Medline: [11393487](#)]
10. Hootman JM, Macera CA, Ham SA, Helmick CG, Sniezek JE. Physical activity levels among the general US adult population and in adults with and without arthritis. *Arthritis Rheum* 2003 Feb 15;49(1):129-135 [FREE Full text] [Medline: [12579604](#)] [doi: [10.1002/art.10911](#)]

11. Fontaine KR, Heo M. Changes in the prevalence of US adults with arthritis who meet physical activity recommendations, 2001-2003. *J Clin Rheumatol* 2005 Feb;11(1):13-16. [Medline: [16357691](#)] [doi: [10.1097/01.rhu.0000152143.25357.fe](#)]
12. Ooijendijk W, Hildebrandt V. Physical activity in the Netherlands 2000: first results from the monitoring study Physical Activity and Health [in Dutch]. In: Ooijendijk W, Hildebrandt V, Stiggelbout M, editors. *Trendrapport Bewegen en Gezondheid 2000/2001*. Heerhugowaard, The Netherlands: PlantijnCasparie; 2002:7-23.
13. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The effectiveness of interventions to increase physical activity. A systematic review. *Am J Prev Med* 2002 May;22(4 Suppl):73-107. [Medline: [11985936](#)] [doi: [10.1016/S0749-3797\(02\)00434-8](#)]
14. Van Der Bij AK, Laurant MGH, Wensing M. Effectiveness of physical activity interventions for older adults: a review. *Am J Prev Med* 2002 Feb;22(2):120-133. [Medline: [11818183](#)] [doi: [10.1016/S0749-3797\(01\)00413-5](#)]
15. Smith BJ. Promotion of physical activity in primary health care: update of the evidence on interventions. *J Sci Med Sport* 2004 Apr;7(1 Suppl):67-73. [Medline: [15214604](#)] [doi: [10.1016/S1440-2440\(04\)80280-9](#)]
16. Conn VS, Minor MA, Burks KJ, Rantz MJ, Pomeroy SH. Integrative review of physical activity intervention research with aging adults. *J Am Geriatr Soc* 2003 Aug;51(8):1159-1168. [Medline: [12890083](#)] [doi: [10.1046/j.1532-5415.2003.51365.x](#)]
17. Rissanen A, Fogelholm M. Physical activity in the prevention and treatment of other morbid conditions and impairments associated with obesity: current evidence and research issues. *Med Sci Sports Exerc* 1999 Nov;31(11 Suppl):S635-S645. [Medline: [10593540](#)]
18. Simons-morton DG, Calfas KJ, Oldenburg B, Burton NW. Effects of interventions in health care settings on physical activity or cardiorespiratory fitness. *Am J Prev Med* 1998 Nov;15(4):413-430. [Medline: [9838981](#)] [doi: [10.1016/S0749-3797\(98\)00078-6](#)]
19. World Internet Users and Population Statistics 2006. Internet World Stats. URL: <http://www.internetworldstats.com/stats.htm> [accessed 2007 Aug 16] [WebCite Cache ID 5R8b46u5g]
20. Marcus BH, Nigg CR, Riebe D, Forsyth LH. Interactive communication strategies: implications for population-based physical-activity promotion. *Am J Prev Med* 2000 Aug;19(2):121-126. [Medline: [10913903](#)] [doi: [10.1016/S0749-3797\(00\)00186-0](#)]
21. Napolitano MA, Marcus BH. Targeting and tailoring physical activity information using print and information technologies. *Exerc Sport Sci Rev* 2002 Jul;30(3):122-128. [Medline: [12150571](#)] [doi: [10.1097/00003677-200207000-00006](#)]
22. Marcus BH, Owen N, Forsyth LH, Cavill NA, Fridinger F. Physical activity interventions using mass media, print media, and information technology. *Am J Prev Med* 1998 Nov;15(4):362-378. [Medline: [9838978](#)] [doi: [10.1016/S0749-3797\(98\)00079-8](#)]
23. Marshall AL, Owen N, Bauman AE. Mediated approaches for influencing physical activity: update of the evidence on mass media, print, telephone and website delivery of interventions. *J Sci Med Sport* 2004 Apr;7(1 Suppl):74-80. [Medline: [15214605](#)] [doi: [10.1016/S1440-2440\(04\)80281-0](#)]
24. Winett RA, Tate DF, Anderson ES, Wojcik JR, Winett SG. Long-term weight gain prevention: a theoretically based Internet approach. *Prev Med* 2005 Aug;41(2):629-641. [Medline: [15917062](#)] [doi: [10.1016/j.ypmed.2004.12.005](#)]
25. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100(2):126-131 [FREE Full text] [Medline: [3920711](#)]
26. Van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997 Oct 15;22(20):2323-2330. [Medline: [9355211](#)] [doi: [10.1097/00007632-199710150-00001](#)]
27. Steultjens EMJ, Dekker J, Bouter LM, Van Schaardenburg D, Van Kuyk MAH, Van Den Ende CHM. Occupational therapy for rheumatoid arthritis: a systematic review. *Arthritis Rheum* 2002 Dec 15;47(6):672-685 [FREE Full text] [Medline: [12522844](#)] [doi: [10.1002/art.10801](#)]
28. Van Den Ende CH, Vliet Vlieland TP, Munneke M, Hazes JM. Dynamic exercise therapy in rheumatoid arthritis: a systematic review. *Br J Rheumatol* 1998 Jun;37(6):677-687 [FREE Full text] [Medline: [9667624](#)] [doi: [10.1093/rheumatology/37.6.677](#)]
29. Van Tulder M, Furlan A, Bombardier C, Bouter L; Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003 Jun 15;28(12):1290-1299. [Medline: [12811274](#)] [doi: [10.1097/00007632-200306150-00014](#)]
30. Welk GJ. *Physical Activity Assessments for Health-Related Research*. Champaign, IL: Human Kinetics Publishers; 2002.
31. 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] [Medline: [15631964](#)] [doi: [10.2196/jmir.6.4.e40](#)]
32. Rovniak LS, Hovell MF, Wojcik JR, Winett RA, Martinez-Donate AP. Enhancing theoretical fidelity: an e-mail-based walking program demonstration. *Am J Health Promot* 2005;20(2):85-95. [Medline: [16295700](#)]
33. Estabrooks PA, Field P, Morse J. Aggregating qualitative findings: an approach to theory development. *Qual Health Res* 1994;4(4):503-511. [doi: [10.1177/104973239400400410](#)]
34. Marshall AL, Leslie ER, Bauman AE, Marcus BH, Owen N. Print versus website physical activity programs: a randomized trial. *Am J Prev Med* 2003 Aug;25(2):88-94. [Medline: [12880874](#)] [doi: [10.1016/S0749-3797\(03\)00111-9](#)]

35. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. *JAMA* 2001 Mar 7;285(9):1172-1177 [FREE Full text] [Medline: [11231746](#)] [doi: [10.1001/jama.285.9.1172](#)]
36. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *JAMA* 2003 Apr 9;289(14):1833-1836 [FREE Full text] [Medline: [12684363](#)] [doi: [10.1001/jama.289.14.1833](#)]
37. Van Den Berg MH, Ronday HK, Peeters AJ, Le Cessie S, Van Der Giesen FJ, Breedveld FC, et al. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. *Arthritis Rheum* 2006 Dec 15;55(6):935-945. [Medline: [17139640](#)] [doi: [10.1002/art.22339](#)]
38. 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 [FREE Full text] [Medline: [10480822](#)]
39. Napolitano MA, Fotheringham M, Tate D, Sciamanna C, Leslie E, Owen N, et al. Evaluation of an internet-based physical activity intervention: a preliminary investigation. *Ann Behav Med* 2003;25(2):92-99. [Medline: [12704010](#)] [doi: [10.1207/S15324796ABM2502_04](#)]
40. Mckay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. *Diabetes Care* 2001 Aug;24(8):1328-1334 [FREE Full text] [Medline: [11473065](#)] [doi: [10.2337/diacare.24.8.1328](#)]
41. Kosma M, Cardinal BJ, Mccubbin JA. A pilot study of a web-based physical activity motivational program for adults with physical disabilities. *Disabil Rehabil* 2005 Dec 15;27(23):1435-1442. [Medline: [16418058](#)] [doi: [10.1080/09638280500242713](#)]
42. Plotnikoff RC, Mccargar LJ, Wilson PM, Loucaides CA. Efficacy of an E-mail intervention for the promotion of physical activity and nutrition behavior in the workplace context. *Am J Health Promot* 2005;19(6):422-429. [Medline: [16022206](#)]
43. Hageman PA, Walker SN, Pullen CH. Tailored versus standard internet-delivered interventions to promote physical activity in older women. *J Geriatr Phys Ther* 2005;28(1):28-33. [Medline: [16236225](#)]
44. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995 Feb 1;273(5):408-412. [Medline: [7823387](#)] [doi: [10.1001/jama.273.5.408](#)]
45. Mihalko SL, Wickley KL, Sharpe BL. Promoting physical activity in independent living communities. *Med Sci Sports Exerc* 2006 Jan;38(1):112-115. [Medline: [16394962](#)] [doi: [10.1249/01.mss.0000183230.08341.6b](#)]
46. Kirk AF, Mutrie N, Macintyre PD, Fisher MB. Promoting and maintaining physical activity in people with type 2 diabetes. *Am J Prev Med* 2004 Nov;27(4):289-296. [Medline: [15488358](#)] [doi: [10.1016/j.amepre.2004.07.009](#)]
47. Woods C, Mutrie N, Scott M. Physical activity intervention: a transtheoretical model-based intervention designed to help sedentary young adults become active. *Health Educ Res* 2002 Aug;17(4):451-460 [FREE Full text] [Medline: [12197590](#)] [doi: [10.1093/her/17.4.451](#)]
48. Adams J, White M. Why don't stage-based activity promotion interventions work? *Health Educ Res* 2005 Apr;20(2):237-243 [FREE Full text] [Medline: [15253998](#)] [doi: [10.1093/her/cyg105](#)]
49. King AC, Kiernan M, Oman RF, Kraemer HC, Hull M, Ahn D. Can we identify who will adhere to long-term physical activity? Signal detection methodology as a potential aid to clinical decision making. *Health Psychol* 1997 Jul;16(4):380-389. [Medline: [9237091](#)] [doi: [10.1037/0278-6133.16.4.380](#)]
50. Worcester MUC, Murphy BM, Mee VK, Roberts SB, Goble AJ. Cardiac rehabilitation programmes: predictors of non-attendance and drop-out. *Eur J Cardiovasc Prev Rehabil* 2004 Aug;11(4):328-335. [Medline: [15292767](#)] [doi: [10.1097/01.hjr.0000137083.20844.54](#)]
51. Bock BC, Marcus BH, Pinto BM, Forsyth LH. Maintenance of physical activity following an individualized motivationally tailored intervention. *Ann Behav Med* 2001;23(2):79-87. [Medline: [11394558](#)] [doi: [10.1207/S15324796ABM2302_2](#)]

submitted 13.12.06; peer-reviewed by L Rovniak; comments to author 19.03.06; revised version received 11.06.07; accepted 23.07.07; published 30.09.07.

Please cite as:

van den Berg MH, Schoones JW, Vliet Vlieland TPM

Internet-Based Physical Activity Interventions: A Systematic Review of the Literature

J Med Internet Res 2007;9(3):e26

URL: <http://www.jmir.org/2007/3/e26/>

doi: [10.2196/jmir.9.3.e26](#)

PMID:

Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Tutorial

Reliability and Validity Issues Related to Interactive Tailored Patient Assessments: A Case Study

Cornelia M Ruland^{1,2,3}, RN, PhD; Suzanne Bakken³, RN, DNSc; Jo Røislien^{4,5}, PhD

¹Center for Shared Decision Making and Nursing Research, Rikshospitalet Medical Center, Oslo, Norway

²Faculty of Medicine, University of Oslo, Oslo, Norway

³Department of Biomedical Informatics, Columbia University, New York, NY, USA

⁴Biostatistics, Rikshospitalet Medical Center, Oslo, Norway

⁵Department of Biostatistics, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:

Cornelia M Ruland, RN, PhD

Center for Shared Decision Making and Nursing Research

Rikshospitalet Medical Center

Forskningsvn. 2b

0027 Oslo

Norway

Phone: +47 23075460

Fax: +47 23075450

Email: cornelia.ruland@rr-research.no

Abstract

Recently there has been a proliferation of interactive tailored patient assessment (ITPA) tools. However, evidence of the reliability and validity of these instruments is often missing, which makes their value in research studies questionable. Because several of the common methods to evaluate instrument reliability and validity are not applicable to interactive tailored patient assessments, informatics researchers may benefit from some guidance on which methods of reliability and validity assessment they can appropriately use. This paper describes the main differences between interactive tailored patient assessments and assessment instruments based on psychometric, or classical test, theory; it summarizes the measurement techniques normally used to ascertain the validity and reliability of assessment instruments based on psychometric theory; it discusses which methods are appropriate for interactive tailored patient assessments and which are not; and finally, it illustrates the application of some of the feasible techniques with a case study that describes how the reliability and validity of the tailored symptom assessment instrument called Choice were evaluated.

(*J Med Internet Res* 2007;9(3):e22) doi:[10.2196/jmir.9.3.e22](https://doi.org/10.2196/jmir.9.3.e22)

KEYWORDS

Validity; reliability; interactive tailored patient assessments

Introduction

Recent years have seen a proliferation of interactive health communication tools, together with a growing trend toward empowering patients to take a more active role in their own health care. A prerequisite to effectively helping patients in need of care is to elicit their symptoms and health problems from their perspective. Interactive tailored patient assessments (ITPAs) have become increasingly important as a means of eliciting patients' illness experiences and tailoring patient care or self-care recommendations to each patient's individual needs. The ease of deployment of Web-based surveys has made the use of interactive tailored questionnaires more common, and

software that allows researchers to rapidly develop custom-tailored questionnaires has started to emerge.

Interactive tailored patient assessments have a number of advantages compared with standardized assessments, in which respondents are required to complete all questions. In interactive tailored patient assessments, the questions can be tailored to each patient individually based on his or her initial responses. Superfluous questions are eliminated, and the questions that remain are more relevant to the patient. For example, the Dialogix system developed at Columbia University implements structured interviews on a series of Web pages. It supports complex branching and conditional tailoring so that questions and summary reports can be tailored to the subject's responses

[1]. The system has been used for surveys on children in the community, the diagnosis of sleep disorder, and depression. Another example is the Composite International Diagnostic Interview (CIDI) used for assessment of mental health disorders, in which positive responses to symptom questions are followed up by other questions, while negative responses often lead to subsequent questions being skipped [2].

One might argue that there do exist traditional assessments that behave somewhat like interactive tailored patient assessments; for example, “if you answered ‘yes’ to question 4, skip questions 5 through 8,” and so on. However, in assessments of this type, *most* questions are answered by all respondents, with additional information gathered for selected subgroups. In interactive tailored patient assessments, however, anything goes; responding, or not, to *any* item can be totally up to the respondent, effectively resulting in each patient completing a “different” assessment. For example, patients can branch into sections that focus on their specific symptoms and problems without being bothered by other questions that are not relevant to them. Because patients complete only a subset of the total number of items available, the response burden is decreased. Consequently, interactive tailored patient assessments allow for an expansion in the breadth and depth of the assessment that helps patients find a closer match between symptom or problem descriptions and their actual illness experiences.

The credibility of interactive tailored patient assessments depends on their ability to adequately capture patients’ experienced symptoms and problems. Validity and reliability are, therefore, crucial issues. Despite an increasing number of studies that use interactive tailored patient assessments as research tools, even in randomized controlled trials, information about reliability and validity is often missing. Consequently, those wishing to implement a specific interactive tailored patient assessment in practice have little assurance about the instrument’s reliability and validity. Also, without such evidence, it is difficult to disseminate study results outside the informatics community and into the clinical literature where a minimum standard for reporting reliability and validity is required for publication. A minimal standard for research instruments should at least include test results of one type of reliability for the group being tested, one type of content validity, and at least one type of criterion-related or construct validity [3].

Psychometric theory offers a number of techniques to examine the reliability and validity of research instruments. However, many of these techniques only apply to instances in which individuals respond to the same set of items, in contrast to interactive tailored patient assessments, in which each informant responds to a different subset of individually selected items. Thus, informatics researchers who are interested in developing an interactive tailored patient assessment are left with the question of which methods they can appropriately use to establish its reliability and validity.

The purpose of this paper is to provide some guidance on evaluating reliability and validity of interactive tailored patient assessments. In it, we (1) describe the main differences between interactive tailored patient assessments and assessment

instruments based on classical test theory, using a tailored symptom assessment instrument called Choice as an example, (2) summarize the psychometric techniques normally used to ascertain the validity and reliability of instruments for self-reported assessments, (3) discuss which methods are appropriate for interactive tailored patient assessments and which are not, and finally, (4) illustrate the application of some of the feasible techniques with a case study that describes measurement of the reliability and validity of the Choice instrument. This may serve as a model for other researchers for evaluating reliability and validity of interactive tailored patient assessments.

Example of an Interactive Tailored Patient Assessment: The Choice Assessment

Choice is the name of a suite of tailored symptom assessment tools designed to help patients report their experienced symptoms and health problems so that their care providers can tailor patient care to each patients’ individual symptoms, problems, and needs. The Choice application used here as an example targets patients with chronic and serious long-term illnesses such as cancer. However, interactive tailored patient assessments are also applicable to other patient populations.

The application is contained and administered via a tablet computer with a touch-sensitive screen or is administered via an Internet application. It supports complex branching, so only relevant questions are asked, and conditional tailoring, so questions are tailored to a subject’s previous responses. For example, in the Choice cancer module, patients first identify among 19 problem categories those that apply to themselves. This triggers a subset of related symptoms from which patients again only select those that apply. For example, if patients initially select the “Problems with eating and drinking” category, they are presented with a more detailed list that helps them specify their eating and drinking problems (eg, taste changes, lack of appetite). The patients then rate the degree of bother and their priorities for care for the selected symptoms. When they are done, the system creates an assessment summary that displays patients’ selected symptoms ranked by their priorities for care. This summary can be used by patients and clinicians for subsequent shared care planning. The Choice instrument has consistently been demonstrated to significantly increase congruence between patients’ reported symptoms and patient care in both rehabilitation and cancer patients [4-6].

Main Differences Between Interactive Tailored Patient Assessments and Traditional Instruments

Traditional Instruments

Interactive tailored patient assessments such as the Choice instrument are different in several respects from other standardized measurement approaches that rely on patient self-report. The primary goal of traditional instruments is to support research, that is, to describe, contrast, or compare populations and to arrive at more generalizable conclusions

based on specific observations [7]. Instruments can be scales or subscales that are composed of theoretically homogeneous items and that measure an internally consistent construct (eg, depression). Scales meet the criteria of classical test theory, and reliability and validity assessments that are based on measures of internal consistency are appropriate. Another type of instrument is an index, which consists of items that are not necessarily correlated and that together compose the index (eg, a measure of quality of life). Rather than being indicators of the underlying theoretical construct, as in scales, items of an index themselves define the construct. Indexes do not meet assumptions of classical test theory, and internal consistency is, therefore, not a good estimate of reliability or validity [8]. Interactive tailored patient assessments are similar to indexes and, therefore, the same statistical limitations apply.

In the application of either scales or indexes, all respondents complete a given set or subset of items [8,9]. This naturally limits the total number of items that can be, or preferably, should be, contained in an instrument. An indicator of a “good” instrument is parsimony—the instrument’s ability to explain the greatest amount of variance of the concept being measured with the fewest number of items [7]. Given that there is evidence of the reliability and validity of the instrument, higher and lower scores represent higher and lower presence, respectively, of the concept being measured.

Interactive Tailored Patient Assessments

Interactive tailored patient assessments are primarily designed for clinical application. Thus, the main focus of interest is to elicit characteristics that are unique to a particular person. The purpose is to provide the person with individually tailored care, information, or behavioral change strategies [10]. This is different from the “one size fits all” approach of traditional measurement instruments, in which the focus of interest is the characteristics of populations rather than the individual.

Another difference from traditional assessment instruments is that an interactive tailored patient assessment may be purposely designed to capture each patient’s personal experience. For example, in the Choice instrument, the goal is to help patients find descriptions of their symptoms and health problems that reflect their personal experiences as closely as possible. Thus, patients may choose between relatively similar symptoms that are expressed with synonymous terms, selecting those that they feel are closest to their experience. Such comprehensiveness of symptom descriptions would be difficult in traditional measurement instruments with a parsimonious set of items and would be considered redundant.

There may also be differences in how questions in the instrument are organized and structured. For example, scales combine items into internally consistent scales, or subscales, which tap the same underlying concept. An example is the Center for Epidemiological Studies Depression Scale (CES-D), described later in the case study, which consists of four subscales for which indicators of depression include “problems concentrating” and “sleeping problems” [11]. However, laypersons may not necessarily understand the associations between these two symptoms and depression. To ensure that patients can branch into their symptoms and health problems without difficulty, items in an interactive tailored patient assessment may be grouped according to laypersons’ knowledge structure rather than according to a theoretical concept such as depression. For example, in the Choice instrument, items are organized based on insights gained from systematic investigations of how laypersons organize and label problems and symptoms into meaningful groups [12]. While such a structure supports patient comprehension and recognition, it does not necessarily fit the structure of an internally consistent scale.

Table 1 summarizes differences between traditional measurement instruments and interactive tailored patient assessments.

Table 1. Differences Between Traditional Measurement Instruments and Interactive Tailored Patient Assessments

	Traditional Measurement Instruments	ITPA Example: Choice
Focus of Interest	Understanding characteristics of populations; generalizability	Understanding characteristics of individuals
Primary Purpose	Research	Clinical practice; to tailor patient care / advice to each individual
Scale	1. Each subscale measures one latent concept at a time. Different concepts are contained in internally consistent subscales. 2. Items of an index serve as causal indicators that define the concept being measured.	May capture patients’ symptom and problem experiences on different dimensions Concepts are not necessarily structured into internally consistent subscales, but are organized to fit the patients’ “lay” knowledge structures.
Set of Questions	Every respondent completes more or less the same set of questions.	Every respondent completes a different set of questions, based on initial item selection.
Goal	Parsimony: to explain the greatest amount of variance in the concept measured with the fewest numbers of items.	Comprehensiveness: to help patients find a close match between the item description and their actual experience.

Techniques to Measure Reliability and Validity and Their Applicability to Interactive Tailored Patient Assessments

Measurement is the process of linking abstract concepts to empirical indicators. This can happen in two ways. The first is by focusing on the crucial relationship between the observable response and the underlying unobservable theoretical concept. This is the case with concepts such as “intelligence,” which we cannot observe directly, but implications of it, such as peoples’ vocabulary, mathematical ability, and knowledge about the world, stem from this quality. Instruments constructed to capture such concepts have come to be called scales [8,9]. The other possibility is that the unobservable theoretical concept under study is the response to observable explanatory factors. This is the case with, for example, socioeconomic status, which is a

function of, say, income and level of education, not the other way around. Instruments constructed to capture such concepts are called indexes, as described earlier [8,9]. The choice of the specific items is much more important in the construction of indexes than of scales.

Reliability and validity are the two basic properties of empirical measurements. Reliability concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials. Validity is the degree to which an instrument measures what it purports to measure. Reliability is a necessary but not a sufficient condition for validity [13,14]. While reliability and validity are equally important for interactive tailored patient assessments as for other standardized assessments, not all common techniques for measuring reliability and validity are appropriate for interactive tailored patient assessments (Table 2).

Table 2. Psychometric concepts, definitions, and methods

Psychometric Concept	Definition	Methods	Appropriateness for ITPAs
Reliability			
Internal consistency	Average intercorrelation among items	Cronbach alpha, split-half	Inappropriate due to highly variable number of assessment items among respondents
Test-retest	Association between measurements on the same respondents at multiple points in time using the same version of the measurement instrument; coefficient of stability	Correlation between two measurements	Inappropriate if concept being measured changes over time; otherwise appropriate. Even small changes over time might fundamentally change the patient’s response to the interactive tailored patient assessment.
Alternate forms	Association between measurements on the same respondents at multiple points in time using two forms of the “same” measurement instrument; coefficient of equivalence	Correlation between two measurements	Inappropriate if concept being measured changes over time; otherwise appropriate. Due to the nature of the interactive tailored patient assessment, with possibly detailed items, coming up with an alternate form might be difficult.
Validity			
Content	Extent to which a specific measure depicts a domain of content	Literature review, expert review	Appropriate
Criterion-related	Extent of correlation between the test and the criterion	Concurrent validity (test and criterion at same point in time); predictive (test and criterion at a future point in time)	Appropriate. Be aware that it might be difficult to find a sensible criterion when many issues are addressed simultaneously, as often is the case.
Construct	Extent to which a particular measure performs in accordance with theoretically derived hypotheses concerning the concepts (or constructs) being measured	Factor analysis, convergent validation, discriminant validation, known group differences, multitrait-multi-method matrix	Factor analysis is often inappropriate due to variable number of assessment items among respondents, or the large sample size that otherwise would be required. Other methods are usually appropriate.

Measures of Reliability

Common approaches to examine reliability include test-retest, alternate forms, split-half, and tests of internal consistency [13,15].

In the test-retest method, the same test is given to the same people after a period of time [13]. The correlations between the scores in the two administrations of the same test are calculated,

and the correlation between two parallel measures equals the reliability coefficient. A prerequisite for test-retest reliability is that the second administration be conducted within a small enough time frame so that the concept being measured (eg, pain) does not change. This is, however, often a problem. Test-retest reliability is appropriate for traditional assessments as well as for interactive tailored patient assessments that measure stable traits, but it is inappropriate for assessments of volatile concepts

that change rapidly over time (eg, how bothersome a symptom is).

The alternate form method requires two testing situations with the same people, but an alternate form of the same test is administered [13]. The two forms are intended to measure the same concept. The correlation between the alternative forms provides the estimate of reliability. Similar to the test-retest method, the alternate form of the instrument must be given within a small enough time frame so that the concept being measured has not changed. Under these conditions, the alternate form approach can be appropriate for interactive tailored patient assessments.

In the split-half technique, items of the scale are split in two. To obtain a measure of reliability, the scores of the halves are correlated. This follows the same logic as in the test-retest technique, where the correlation between two parallel measures equals the reliability coefficient. The issue of how to split the items in half, however, is not clear cut.

By far the most popular approach is the internal consistency reliability coefficient Cronbach alpha [16]. Among the reasons for its popularity is the fact that it, like the split-half technique, requires only a single test administration. It does, however, expand on that methodology of the split-half technique, and the calculation of alpha is based on the inter-item correlations among all the items of the scale. The higher the alpha, the higher the reliability [13].

A problem with all the above measures is that they indirectly depend on all respondents completing more or less the same consistent set of items, making the measures difficult to apply to interactive tailored patient assessments. A scale's reliability is mainly addressed by looking at correlations—mathematical expressions of association. The calculations are done by pairing data and comparing whether variable values behave in a similar manner; if the value of one variable goes up, and the value of another tends to do so as well, the two variables will be more correlated than if this was not so. Problems arise, however, in the presence of missing data (ie, there is no value for a given variable to compare with another). Usually, the issue of missing values in a data set constitutes no major problem when calculating correlations. For example, for 100 patients measured on weight and shoe size, with two persons missing out on the weighting because they were in the gym, this still leaves 98 people for the calculation of the correlation between weight and shoe size for that group of patients. Generally, the amount of missing data in reviewing scales is negligible. There will most likely be some patients that have not answered one item or another, but the amount of pairs left for correlation calculations is rarely affected to such an extent that these calculations suffer severely.

In interactive tailored patient assessments, however, the amount of missing data could be devastatingly high, effectively making well-known techniques useless. Take the Choice instrument. It has a total of 141 symptoms that the patients can choose from. In the testing of the system, the average number of symptoms the patients reported was 10 [17]. That is, for every patient, the average amount of “missing data” after an assessment was more than 90%. Note that these non-answers are actual missing data

in the definition of the term: if a patient has not chosen to say something about symptom A, it is not the same as having reported “no bother with symptom A,” which would give a zero value (or similar measure of “nothing”) to use in calculations. But here we do not have any information about how the patient felt about symptom A at the time. Maybe the patient actually had something to say about symptom A but prioritized other items which were more important or simply forgot to respond to that item.

This lack of a fixed system of items to perform calculations on in order to verify the reliability of an interactive tailored patient assessment constitutes a major statistical challenge. All correlation calculations are deemed to be suffering from this fact, and all correlations will be calculated less precisely since the unanswered questions will contribute a “missing,” erasing that piece of information totally, rather than a zero or similar value, as in more traditional assessments. For example, a patient answering items 1 through 5 in one administration of an interactive tailored patient assessment and items 2 through 10 in another administration of the same interactive tailored patient assessment, would, in a test-retest, only have four items in common for the two administrations, even though five items were answered the first time and nine the second time, for a total of 10 different items.

The calculation of Cronbach alpha [16] depends on the number of items and the mean inter-item correlation. For interactive tailored patient assessments, however, one needs an adjustment for the fact that each patient only responds to a small subdomain of N , which will differ from patient to patient. Further, the inter-item correlation is then based on an extremely sparsely filled scale. Finally, the shared size of the interactive tailored patient assessment instrument is a possible problem in itself; with 100 items, an average inter-item correlation of only 0.04 is enough to ensure an alpha of .80.

Factor analysis is closely linked to reliability measures, but makes less stringent assumptions than alpha-type methods. Such methods are, however, also deemed to be unreliable in the setting described above. Factor analysis does nothing more than redefine and simplify the correlation matrix, a matrix that may be calculated on the basis of a huge amount of missing data and very sparse real information. The number of assessments needed in order to have a trustworthy correlation matrix would then have to be extremely high. There are several guidelines for sample size. Among others, Tinsley and Tinsley [18] suggest a ratio of 5-10 subjects per item, up to about 300 subjects. Thereafter, the ratio can be somewhat relaxed. Comrey [19], on the other hand, stated that a sample size of 200 is adequate in most cases of ordinary factor analysis that involve no more than 40 items. However, this calculation breaks down for a 141-item assessment in which each individual selects approximately 10 items; the exact sample size needed in these instances thus becomes very difficult to calculate. Cronbach alpha and other similar measures, as well as factor analysis methodologies, are indirectly based on the fact that all patients fill out the same fixed set of items or close thereto. To our knowledge, nobody has refined these statistical measures to cope with the problems described above. Validating interactive tailored patient assessments thus relies on carefully reviewing

the options at hand to see whether they will be applicable for a given instrument. For the Choice instrument, a hybrid of test-retest and alternate forms was used for reliability assessment. It is described in more detail below in the case study.

Measures of Validity

The main methods to assess the validity of a test for a group of people under certain circumstances are content validity, criterion-related validity, and construct validity. Fundamentally, content validity depends on the extent to which an empirical measurement reflects a specific domain of content and whether the items reflect the meaning associated with each dimension or subdimension [13] of that measure. Content validity is crucial for all measurements, including interactive tailored patient assessments, but unfortunately there is no rigorous way to assess it [13].

Criterion-related validity refers to the correlation of a measure with a criterion variable that is external to the measuring instrument itself [15]. The higher the correlation, the more valid is the measure for the particular criterion. The measurements may be collected at the same point in time (concurrent validity), or the measurement under study may be used to predict a future measurement (predictive validity). For example, the degree to which a test for college admission can predict later academic achievement reflects criterion-related validity of the test. The availability of a criterion measurement (ie, a gold standard) is a prerequisite to examining criterion-related validity of any assessment, tailored or untailored. Because such a gold standard is often missing, measuring criterion-related validity is difficult.

In contrast to content validity and criterion-related validity, construct validity has a more generalized applicability and lends itself easier to empirical investigation. Constructs concern domains of variables [15]. Construct validity is concerned with the extent to which a particular measure relates to other measurements consistent with theoretically derived hypotheses concerning the construct being measured [13]. There are three major aspects of construct validation: (1) specifying the domain of observables related to the construct, (2) determining the extent to which observables measure the same thing, and (3) performing subsequent experiments to determine the extent to which supposed measures of the construct are consistent with “best guesses” about the construct [7,15].

A number of techniques for examining construct validity are applicable to interactive tailored patient assessments. For example, convergent and discriminant approaches, including known group differences, are based on hypothesized relationships between the measurement of concern and another variable. Convergent validity is demonstrated when two independent methods that measure the same variable or attribute are highly correlated. Divergent validity is demonstrated when measures of different attributes do not highly correlate.

In their seminal paper on construct validation, Campbell and Fiske [20] proposed the multitrait-multimethod matrix as an approach to examining convergent and discriminant validity. The multitrait-multimethod matrix includes two traits (one of primary interest) and two methods that are applied to both traits.

The basic premise is that the measurements of a trait will converge across methods and diverge between traits. For example, measurements related to dyspnea severity should converge across paper-based and computer-based assessment methods, but the measurement of dyspnea severity should be less highly correlated with the measurement of nausea severity using the same method.

Other techniques to establish construct validity that examine the internal structure of a measurement instrument, such as factor analysis, are, however, often inappropriate for interactive tailored patient assessments because of their dependence on a reliable correlation matrix. The sheer size of the population needed to verify the instrument, coping with both the possible three-digit number of items and the possible close-to-100% missing data, could approach numbers way out of practical reach. Table 2 summarizes psychometric concepts, measurement methods, and their appropriateness for interactive tailored patient assessments.

A Case Study: Examining Reliability and Validity of the Choice Instrument

Reliability Assessment

When testing the reliability of Choice, it was evident that we needed a way of being able to pair observations on the different items without encountering an overwhelming amount of missing data. Because questions in the Choice instrument are tailored to each respondent based on initial response, reliability measures that are built on internal consistency could not be appropriately used for the evaluation of reliability.

A first thought was to perform a test-retest, as it would be natural to assume that an individual would correlate higher with himself or herself (ie, having the same bothersome symptoms and same priorities for treatment if the time frame between the tests was sufficiently short), reducing the amount of missing data in the correlation pairing. A complete test-retest using the Choice instrument felt inappropriate, however, because of the risk that patients' symptom reports could change to such an extent that the discrepancy between items chosen in the test and the retest would make the correlation calculations unreliable. This concern was strengthened by the fact that several of the items address issues that change fairly quickly with time.

The alternate form approach seemed a logical second option, but as the Choice instrument contains 141 symptoms with several nuances in the wording to capture the specific disease pattern of the particular patient, as described earlier, an alternate form could run the risk of being different in such a way that patients would choose other symptoms merely due to the wording of the items. It seemed difficult to come up with an acceptable, completely alternative form of the instrument. There did, however, exist a somewhat alternative format of the Choice instrument that would at the same time minimize the amount of missing data: the full list of the 141 symptoms. We used this to assess the reliability of the Choice instrument.

To collect the reliability data, we conducted a separate study independent from our clinical trial. Because reliability is

sample-specific, patients in this new study were recruited from the same population and setting and had to meet the same inclusion criteria as patients in the clinical trial. After Institutional Review Board approval was obtained, 100 patients undergoing cancer treatment were recruited. First, patients were asked to complete the tailored Choice assessment similar to patients in the clinical trial. Immediately after and in the same data collection session, they were asked to complete a questionnaire, the alternate form that included the full set of 141 symptom descriptions contained in Choice. The correlation between Choice and questionnaire data was 0.74 for all symptoms, and 0.85 for moderately or very bothersome symptoms [17]. According to Nunnally and Bernstein, correlations greater than 0.70 provide evidence of the satisfactory reliability of a measurement instrument [15].

It may at first be surprising that the correlation coefficients between the two formats were not higher. The main reason was that in the Choice instrument it is possible to choose different terms to express almost the same symptom. For example, a patient who chose "lack of energy" in the interactive tailored patient assessment version, chose instead "fatigued" in the paper-based form. While the patient may not have been aware of this distinction, this weakened the correlations between the two forms, making them somewhat lower than one might expect.

Validity Assessment

Content Validity

As above mentioned, content validity depends greatly on the adequacy with which a specific domain of content is sampled [15]. While this is difficult to measure directly, thorough and appropriate procedures used during the development of a new instrument are a prerequisite of content validity. It is impossible to specify exactly how many items need to be developed for a particular domain of content. However, it is always preferable to initially create too many items rather than too few as inadequate items can always be eliminated [13]. This is particularly true for interactive tailored patient assessments, in which patients complete only those subsets of items relevant to them, and the total number of items thus matters less. Here we describe the process for developing and ensuring content validity of the Choice module for cancer patients.

The goal when constructing the tailored Choice instrument was to assist patients in communicating their illness experience along physical, psychosocial, and functional dimensions as close as possible to their actual experiences. It was, therefore, important to include a comprehensive set of items that reflected all dimensions of patients' illness experiences in sufficient level of detail and that were expressed in lay language to support patient recognition and communication.

To identify items to be included, we conducted a thorough review of the scientific literature to identify problems, specific symptoms, and functional limitations encountered by cancer patients. This search and review included the health care bibliographic databases as well as the World Wide Web and resulted in a preliminary list of symptoms and functional problems for potential inclusion. Expert groups of specialists in cancer care (physicians, nurses, social workers) then critically

reviewed this list for relevance, comprehensibility, completeness, and level of detail and supplemented it with expert opinion [6]. Particular attention was paid to expressing symptoms and problems in simple, understandable, nonmedical lay language. Next, the revised symptom list was presented to 15 cancer inpatients and outpatients (9 women, 6 men; age 40-74 years) who were asked to complete and evaluate a paper-based version of the symptom assessment for clarity of meaning, appropriateness, wording, completeness, redundancy, and format, and to add comments. This resulted in further suggestions for revisions, which were discussed in the cancer expert groups. The subsequently refined symptom list was then implemented in the tailored computer application and pilot tested with 56 outpatients with varying cancer diagnoses [6]. Based on this pilot study, a few item descriptions were revised to better describe symptoms from the perspective of the patients. The final version was used for the reliability testing described above and in the clinical trial that provided data for the validity testing described below.

Construct Validity

To evaluate construct validity of the Choice instrument, we used known group differences techniques as well as assessments of convergent and discriminant validity. We performed three evaluations of known group differences based on data collected in a clinical trial of 148 patients who received active cancer treatment for leukemia and lymphoma.

The first test was based on the hypothesis that patients undergoing a stem cell transplant would report more symptoms with the Choice instrument than patients treated with chemotherapy only. This hypothesis is consistent with empirical evidence on treatment side effects and was supported by the data. Patients undergoing a stem cell transplant reported significantly more symptoms than patients in the chemotherapy group (14.6 vs 9.2, $P < .001$).

In the second test, we examined gender differences in self-reported symptoms. Because the literature has provided some evidence that women report more symptoms than men [21], we expected that this difference would also be found with the Choice instrument. This was again supported. In our clinical trial, women reported significantly more symptoms than men (13.7 vs 10.0, $P < .001$).

Finally, we examined whether the most reported symptoms during patients' illness trajectories were consistent with expected symptom patterns during different phases of treatment and rehabilitation. This was again supported. The most frequently selected symptoms 1 to 2 months into treatment were side effects related to chemotherapy and stem cell transplant, including nausea, vomiting, and mouth sores. During the third and fourth months of treatment, long-term side effects such as neurological problems, memory problems, and weight loss started to occur more frequently. During rehabilitation, the number of physical symptoms decreased and the focus of self-reported symptoms shifted to issues regarding resuming a normal life and worries about the future. Thus, all three known group difference tests performed as expected and provided support for the validity of the Choice instrument.

To measure convergent and discriminant validity, we compared the performance of the Choice instrument in our clinical trial data set with two other measures taken at the same time point: the CES-D [11] and the SF-36, a multidimensional measure of health-related quality of life [22]. Ideally, measures of a similar trait should correlate higher with each other than they do with measures of different traits. To estimate convergent validity, we computed the correlations between the psychosocial subscales of the Choice instrument and both the CES-D depression subscale and the SF-36 mental health index subscale. A correlation of 0.57 was found with the CES-D depression subscale and -0.64 with the SF-36 mental health index. Similar evidence of convergent validity was found for physical symptoms. The physical symptom subscales of the Choice instrument strongly correlated with the SF-36 bodily pain scale ($r = -0.61$), the SF-36 physical health component subscale ($r = -0.54$), and the SF-36 physical functioning subscale ($r = -0.44$).

To assess discriminant validity, we performed correlations between Choice subscales and CES-D and SF-36 subscales that measured different attributes, hypothesizing that they would not correlate to a very high degree. This was supported by our data. The physical symptom subscales of the Choice instrument correlated only weakly with the CES-D depression subscale ($r = 0.25$) and the SF-36 mental health index ($r = -0.28$). Similarly, psychosocial symptoms in the Choice instrument correlated weakly with the SF-36 physical functioning subscale ($r = -0.18$) and the physical health component subscale ($r = -0.13$).

Conclusion

In this paper, we strongly advocate evaluating and reporting reliability and validity of interactive tailored patient assessments,

which is crucial for the credibility of interactive tailored patient assessments as research instruments. However, several of the common measurement techniques available to assess these psychometric properties are not applicable to interactive tailored patient assessments. The advantage of computerized tailored assessments is that patients can skip unimportant items and hone in on problems that matter to them and that reflect their actual experience. However, this advantage makes reliability and validity assessments of interactive tailored patient assessments a challenge for informatics researchers. To assist in this task, we have discussed which techniques might be feasible for establishing reliability and validity of interactive tailored patient assessments and demonstrated their application in a case study of the Choice instrument.

Although assessment of reliability of an interactive tailored patient assessment may require collection of a separate data set in addition to the clinical trial data, this is well worth the effort. A basic core of evidence of reliability and validity is needed for any instrument. Reliability is a prerequisite for validity, and an unreliable instrument cannot be valid. Unreliable and invalid instruments are not worth further investigation [3]. Reporting of interactive tailored patient assessment reliability and validity should become a requirement for publishable informatics research, so researchers can trust the data. Evidence of reliability and validity has long been a requirement for publication of research instruments in the clinical literature, and is, therefore, a prerequisite for the dissemination of informatics tools outside the informatics community. The adoption of a similar requirement in scientific informatics journals would greatly enhance the state of science in the field of tailored assessments and health interventions.

Acknowledgments

This work was supported by the Norwegian Research Council, grant #154739/320. The authors would like to thank members of the Choice research team for their assistance in data collection, data entry, and scoring of the instruments: Dr. Glenys Hamilton, Jørn Kristiansen, and Heidi Sandbæk.

Conflicts of Interest

None declared.

References

1. White TM, Hauan MJ. The capture and use of detailed process information in the Dialogix system for structured web-based interactions. Proc AMIA Symp 2001;761-765. [Medline: [11825288](#)]
2. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. J Psychiatr Res 1994;28(1):57-84. [Medline: [8064641](#)] [doi: [10.1016/0022-3956\(94\)90036-1](#)]
3. Norbeck JS. What constitutes a publishable report of instrument development? Nurs Res 1985;34(6):380-382. [Medline: [3852252](#)] [doi: [10.1097/00006199-198511000-00022](#)]
4. Ruland CM. Decision support for patient preference-based care planning: effects on nursing care and patient outcomes. J Am Med Inform Assoc 1999;6(4):304-312 [FREE Full text] [Medline: [10428003](#)]
5. Ruland CM. Handheld technology to improve patient care: evaluating a support system for preference-based care planning at the bedside. J Am Med Inform Assoc 2002;9(2):192-201 [FREE Full text] [Medline: [11861634](#)] [doi: [10.1197/jamia.M0891](#)]

6. Ruland CM, White T, Stevens M, Fanciullo G, Khilani SM. Effects of a computerized system to support shared decision making in symptom management of cancer patients: preliminary results. *J Am Med Inform Assoc* 2003;10(6):573-579 [FREE Full text] [Medline: [12925545](#)] [doi: [10.1197/jamia.M1365](#)]
7. DeVellis RF. *Scale Development: Theory and Applications*. 2nd edition. Thousand Oaks, CA: Sage Publications; 2003.
8. Streiner DL. Being inconsistent about consistency: when coefficient alpha does and doesn't matter. *J Pers Assess* 2003 Jun;80(3):217-222. [Medline: [12763696](#)] [doi: [10.1207/S15327752JPA8003_01](#)]
9. Bollen K, Lennox R. Conventional wisdom on measurement: a structural equation perspective. *Psychol Bull* 1991;110(2):305-314. [doi: [10.1037/0033-2909.110.2.305](#)]
10. Kreuter MW, Bull FC, Clark EM, Oswald DL. Understanding how people process health information: a comparison of tailored and nontailored weight-loss materials. *Health Psychol* 1999 Sep;18(5):487-494. [Medline: [10519465](#)] [doi: [10.1037/0278-6133.18.5.487](#)]
11. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1(3):385-401. [doi: [10.1177/014662167700100306](#)]
12. Andersen R, Ruland C, Slaughter L, Andersen T, Jacobsen W. Clustering techniques for organizing cancer-related concepts into meaningful groups for patients. *AMIA Annu Symp Proc* 2005:882. [Medline: [16779169](#)]
13. Carmines EG, Zeller RA. *Reliability and Validity Assessment*. Newbury Park, CA: Sage Publications; 1979.
14. Pedhazur EJ, Schmelkin LP. *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991.
15. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd edition. New York: McGraw-Hill; 1994.
16. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16(3):297-334. [doi: [10.1007/BF02310555](#)]
17. Ruland C, Røslie J, Bakken S, Kristiansen J. Comparing tailored computerized symptom assessments to interviews and questionnaires. *AMIA Annu Symp Proc* 2006:1081 [FREE Full text] [Medline: [17238700](#)]
18. Tinsley DJ, Tinsley HEA. Uses of factor analysis in counselling psychological research. *J Couns Psychol* 1987;34(4):414-424. [doi: [10.1037/0022-0167.34.4.414](#)]
19. Comrey AL. Factor-analytic methods of scale development in personality and clinical psychology. *J Consult Clin Psychol* 1988 Oct;56(5):754-761. [Medline: [3057010](#)] [doi: [10.1037/0022-006X.56.5.754](#)]
20. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959 Mar;56(2):81-105. [Medline: [13634291](#)] [doi: [10.1037/h0046016](#)]
21. Miaskowski C. Gender differences in pain, fatigue, and depression in patients with cancer. *J Natl Cancer Inst Monogr* 2004(32):139-143. [Medline: [15263057](#)] [doi: [10.1093/jncimonographs/lgh024](#)]
22. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998 Nov;51(11):903-912. [Medline: [9817107](#)] [doi: [10.1016/S0895-4356\(98\)00081-X](#)]

Abbreviations

CES-D: Center for Epidemiological Studies Depression Scale

ITPA: interactive tailored patient assessment

SF-36: Medical Outcomes Study 36-Item Short Form Health Survey

submitted 20.12.06; peer-reviewed by D Lewis, D Streiner; comments to author 05.01.06; revised version received 16.04.07; accepted 20.05.07; published 01.08.07.

Please cite as:

Ruland CM, Bakken S, Røslie J

Reliability and Validity Issues Related to Interactive Tailored Patient Assessments: A Case Study

J Med Internet Res 2007;9(3):e22

URL: <http://www.jmir.org/2007/3/e22/>

doi: [10.2196/jmir.9.3.e22](#)

PMID:

© Cornelia M Ruland, Suzanne Bakken, Jo Røslie. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 01.08.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

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)

Ellen Burke Beckjord^{1,2}, PhD, MPH; Lila J Finney Rutten², PHD, MPH; Linda Squiers³, PhD; Neeraj K Arora⁴, PhD; Lindsey Volckmann², BA; Richard P Moser⁵, PhD; Bradford W Hesse², PhD

¹Cancer Prevention Fellowship Program, Division of Cancer Prevention, Office of Preventive Oncology, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA

²Division of Cancer Control and Population Sciences (DCCPS), Behavioral Research Program (BRP), Health Communication and Informatics Research Branch (HCIRB), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA

³Cancer Information Service, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA

⁴Division of Cancer Control and Population Sciences (DCCPS), Applied Research Program, Outcomes Research Branch, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA

⁵Division of Cancer Control and Population Sciences (DCCPS), Behavioral Research Program (BRP), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA

Corresponding Author:

Ellen Burke Beckjord, PhD, MPH
National Cancer Institute
6130 Executive Boulevard (EPN), 4051A, MSC 7365
Bethesda, MD 20892
USA
Phone: +1 301 435 2842
Fax: +1 301 480 2198
Email: beckjore@mail.nih.gov

Abstract

Background: Despite substantial evidence that the public wants access to Internet-based communication with health care providers, online patient-provider communication remains relatively uncommon, and few studies have examined sociodemographic and health-related factors associated with the use of online communication with health care providers at a population level.

Objective: The aim of the study was to use nationally representative data to report on the prevalence of and changes in use of online patient-provider communication in 2003 and 2005 and to describe sociodemographic and health-related factors associated with its use.

Methods: Data for this study are from two iterations of the Health Information National Trends Survey (HINTS 2003, HINTS 2005). In both years, respondents were asked whether they had ever used email or the Internet to communicate with a doctor or a doctor's office. Adult Internet users in 2003 (n = 3982) and 2005 (n = 3244) were included in the present study. Multivariate logistic regression analysis was conducted to identify predictors for electronic communication with health care providers.

Results: In 2003, 7% of Internet users had communicated online with a health care provider; this prevalence significantly increased to 10% in 2005. In multivariate analyses, Internet users with more years of education, who lived in a metro area, who reported poorer health status or who had a personal history of cancer were more likely to have used online patient-provider communication.

Conclusions: Despite wide diffusion of the Internet, online patient-provider communication remains uncommon but is slowly increasing. Policy-level changes are needed to maximize the availability and effectiveness of online patient-provider communication for health care consumers and health care providers. Internet access remains a significant barrier to online patient-provider communication.

(*J Med Internet Res* 2007;9(3):e20) doi:[10.2196/jmir.9.3.e20](https://doi.org/10.2196/jmir.9.3.e20)

KEYWORDS

Internet; patient-provider communication; electronic mail; information services, trends and utilization; medical informatics, trends; health education; health services; demography; data collection; health care surveys; neoplasms; regression analysis

Introduction

For more than a decade, studies have consistently shown that some members of the public want access to Internet-based communication with health care providers, with preference estimates for online patient-provider communication ranging from 40% to 83% [1-9]. The Institute of Medicine has characterized patient-centered care as health care that addresses patient needs and preferences, and it has identified information technology as crucial to advancing health care quality [10]. There is evidence that online communication with health care providers promotes health-related quality of life (eg, [11]) and, further, that health care consumers would benefit from increased partnerships between health information technology and health care providers: health care providers are a more trusted source of health information than the Internet, but the Internet is the source most often used by the public to retrieve health information [12]. Thus, as the presence of eHealth applications such as online patient-provider communication grows within the health care landscape, it is important to examine the prevalence of Internet-based patient-provider communication and to identify sociodemographic and health-related factors associated with its use to ensure that the potential benefits of online communication with health care providers are available to all health care consumers.

Despite the penetration of Internet access (an estimated 73% of American adults are online, 91% of whom use email [13]) and the growing public endorsement of the Internet as a useful tool in health-related decision making [14], online patient-provider communication remains uncommon [15-18]. While acknowledging potential benefits of online patient-provider communication (eg, for scheduling appointments [19]), health care providers have concerns regarding confidentiality, reimbursement, and workload related to online communication with patients [20]. In general, preference for online communication is higher among patients than among health care providers [21].

In 2003, the National Cancer Institute launched the first biennial Health Information National Trends Survey (HINTS) [22,23]. HINTS was designed to capture nationally representative data on the impact of the cancer information environment and to specifically assess the public's use of information technology for health. Though previous estimates of the online public's use of the Internet to communicate with health care providers have been reported from nationally representative data [15,17,24], the purpose and design of HINTS allow for a more in-depth examination of sociodemographic and health-related factors in relation to online communication with health care providers than has been possible in previous investigations. Further, using cross-sectional iterations of HINTS from 2003 and 2005, we can examine whether the prevalence of online communication with health care providers among American Internet users has changed over time.

In 2001, an estimated 6% of Internet users had emailed health care providers [15], up to 7% in 2003 [17]. In this paper, we provide estimates of Internet-based patient-provider communication derived from national probability samples of online American adults in 2003 and 2005. In addition, we report on changes in use of the Internet to communicate with health care providers over this 2-year period and identify sociodemographic and health-related factors associated with online patient-provider communication. Overall, we expect that the prevalence of online patient-provider communication will be low, but that use of the Internet to communicate with health care providers will significantly increase between 2003 and 2005. Further, based on results of previous studies [2,25], we expect that Internet users with more years of education, who are of higher socioeconomic status, and who are more engaged with the health care system will be more likely to report having communicated online with an health care provider.

Methods**Data Source**

Data for this study are from HINTS 2003 and HINTS 2005, two iterations of the nationally representative survey designed to assess the impact of the cancer information environment and the public's knowledge of, attitudes toward, and behaviors related to cancer and cancer prevention [26]. Comprehensive reports on the conceptual framework of HINTS and sample designs are published elsewhere [22,23].

Data for HINTS 2003 were collected from October 2002 through April 2003. Data for HINTS 2005 were collected from February through August 2005. The cross-sectional surveys were administered by trained interviewers to representative samples of American households drawn from all telephone exchanges in the United States. Exchanges with high numbers of African Americans and Hispanics were oversampled in 2003. One adult (age 18 or older) was selected from each household to participate in the full survey during a household screening. In 2003, response rates were 55% at the household screening level (ie, the initial contact with the household used for sampling purposes) and 63% at the sampled person interview level (ie, completion of the interview by the sampled household member); in 2005, the respective response rates were 34% and 61%. Every sampled adult who completed a questionnaire in HINTS 2003 and HINTS 2005 was assigned a final sampling weight and a set of 50 replicate sampling weights. These sampling weights were used for the purpose of computing nationally representative estimates, to adjust for nonresponse and to reduce the sampling variance of estimators through utilization of information with less sampling and nonsampling error than the corresponding HINTS estimates (eg, estimates obtained through the Current Population Survey, which has much larger sample sizes than HINTS).

Complete interviews were conducted with 6369 adults for HINTS 2003 and with 5394 adults for HINTS 2005. In both

surveys, only Internet users ($n = 3982$ in 2003; $n = 3244$ in 2005) were asked whether they had used email or the Internet to communicate with a doctor or a doctor's office in the past 12 months (yes/no). Thus, Internet users in HINTS 2003 and 2005 served as the study populations for the current investigation.

Study Variables

As in prior investigations of Internet users' use of online patient-provider communication (eg, [15]), sociodemographic variables included in the present study were age, gender, education, annual income, race/ethnicity, and metropolitan statistical area (metro or non-metro county). Previous studies have also suggested that health-related variables, such as poorer self-reported health status [2] and having health insurance [25], are associated with use of online patient-provider communication. Health-related variables included in the present study were self-reported health status, possession of health insurance, and personal cancer history.

Data Analysis

Analyses were conducted using SAS-callable SUDAAN, version 9.0 [27] to account for the complex survey design of HINTS and to obtain appropriate standard errors and 95% confidence intervals (CIs) for point estimates. Responses of "refused" or "don't know" were counted as missing. Unweighted and weighted descriptive statistics are presented, and weighted data were used in all inferential statistical analyses. Bivariate analyses (chi-square) estimated changes in the prevalence of online communication with health care providers between 2003 and 2005 and associations between sociodemographic or health-related factors with online communication in 2003 and 2005. Three multivariate logistic regression models were used

to estimate the odds of having used online communication with health care providers. The first used the combined data set of Internet users from 2003 and 2005 to model changes in use of online patient-provider communication over this 2-year period. We then examined sociodemographic and health-related factors separately in 2003 and 2005 to determine whether study variables associated with use of online patient-provider communication were consistent over time. The regression models used a forced-entry variable selection wherein all study variables were entered in one step.

Results

Sample Characteristics and Bivariate Analyses

Table 1 displays the sample characteristics of Internet users in HINTS 2003 and HINTS 2005 (weighted and unweighted). In 2003, 7% of Internet users reported communicating online with an health care provider in the past 12 months, consistent with previous prevalence estimates of Internet-based patient-provider communication in 2003 [17]. In 2005, 10% of Internet users reported communicating online with an health care provider. In bivariate analyses, this increase in use of Internet-based patient-provider communication from 2003 to 2005 was statistically significant ($\chi^2_1 = 9.44$; $P = .003$).

Bivariate associations between study variables and emailing health care providers are displayed in Table 2. In 2003, respondents who had communicated online with an health care provider had significantly more years of education and were more likely to reside in a metro county. In 2005, they were more likely to be female, had significantly more years of education, and were more likely to have a personal history of cancer.

Table 1. Sample characteristics of Internet users in HINTS 2003 and HINTS 2005*

	% of HINTS 2003 Internet Users (n = 3982)		% of HINTS 2005 Internet Users (n = 3244)	
	Unweighted	Weighted	Unweighted	Weighted
Communicated online with an health care provider in the past 12 months	8	7	10	10
Age (years)				
18-34	32	38	24	38
35-49	37	36	33	33
50-64	23	20	30	21
65-74	6	4	9	5
75 or older	2	2	4	2
Gender				
Male	41	50	36	48
Female	59	50	64	52
Education				
Less than high school	4	7	4	5
High school graduate	22	25	19	24
Some college	31	33	32	38
College graduate	44	35	45	33
Annual Income (US \$)				
< 25000	17	16	12	13
25000 to < 35000	12	11	8	7
35000 to < 50000	18	18	13	12
50000 to < 75000	22	22	22	21
75000 or more	31	33	31	33
Race/Ethnicity				
White	73	75	82	77
Hispanic/Latino	8	7	6	8
African American	10	8	7	9
Asian American/Other [†]	5	6	6	7
Health Insurance				
No	9	11	9	12
Yes	90	89	91	88
Health Status				
Excellent/very good/good	84	84	85	84
Fair/poor	15	16	15	16
History of Cancer				
Yes	10	8	12	9
No	90	92	88	91
Metropolitan Statistical Area				
County in metro area	84	84	81	82
County in non-metro area	16	16	19	18

*Within-category cell values that add up to less than 100% reflect missing data due to responses of "refused" or "don't know."

[†]Other includes Pacific Islander, Native Hawaiian, American Indian, Alaska Native, and multiple races mentioned.

Table 2. Bivariate associations between communicating online with health care providers and study variables in HINTS 2003 and HINTS 2005

	HINTS 2003		HINTS 2005	
	% Communicated online with an health care provider*	<i>P</i> †	% Communicated online with an health care provider*	<i>P</i> †
Age (years)		.19		.35
18-34	6.4		10.2	
35-49	6.8		9.3	
50-64	9.3		10.1	
64-75	4.6		6.4	
75 or older	6.7		7.0	
Gender		.19		.02
Male	7.6		7.9	
Female	6.4		11.2	
Education		< .001		.049
Less than high school	3.2		8.3	
High school graduate	3.5		6.6	
Some college	7.3		10.1	
College graduate	10.3		11.7	
Annual Income (US \$)		.07		.17
< 25000	7.9		7.6	
25000 to < 35000	6.5		9.9	
35000 to < 50000	5.5		7.6	
50000 to < 75000	6.5		8.4	
75000 or more	9.2		12.8	
Race/Ethnicity		.88		.34
White	7.3		9.5	
Hispanic/Latino	6.4		5.9	
African American	6.1		11.3	
Asian American/Other‡	7.2		13.5	
Health Insurance		.54		.27
No	7.3		10.1	
Yes	6.3		7.8	
Health Status		.21		.79
Excellent/very good/good	6.9		9.9	
Fair/poor	8.3		9.3	
History of Cancer		.13		.03
Yes	9.4		14.5	
No	6.8		9.1	
Metropolitan Statistical Area		< .001		.12
County in metro area	7.5		10.1	
County in non-metro area	4.3		7.4	

*Weighted percents of Internet users who communicated online with an health care provider in the past 12 months within each study variable category.

†From chi-square tests (with degrees of freedom equaling number of categories minus 1).

‡Other includes Pacific Islander, Native Hawaiian, American Indian, Alaska Native, and multiple races mentioned.

Multivariate Analyses

Our first multivariate logistic regression estimated changes in the odds of having communicated online with an health care provider between 2003 and 2005. This analysis was done with a combined data set of Internet users from HINTS 2003 and 2005 ($n = 7134$). The year of HINTS administration (2003 or 2005) was included to examine whether the increase in prevalence of online patient-provider communication remained significant after adjustment for the study variables in [Table 1](#) (data not shown). Consistent with bivariate results, the increase in prevalence of online patient-provider communication among adult Internet users between 2003 and 2005 was significant; there was a 33% increase in the odds of having communicated online with an health care provider among respondents in HINTS 2005 compared to respondents in HINTS 2003 (OR = 1.33; 95% CI = 1.04-1.70; $P = .03$).

[Table 3](#) displays the results of the multivariate analyses by HINTS year. Consistent with bivariate results, in 2003, education and metropolitan statistical area were associated with use of online patient-provider communication. Specifically,

Internet users who were college graduates had over three times the odds of communicating online with an health care provider compared to those with less than a high school education (OR = 3.73; 95% CI = 1.10-12.59; $P = .03$). Those who lived in a non-metro area were less likely to have used online patient-provider communication compared to Internet users who resided in metro area counties (OR = 0.62; 95% CI = 0.41-0.95; $P = .03$). Finally, Internet users who reported “fair” or “poor” health status had higher odds of communicating online with an health care provider (OR = 1.46; 95% CI = 1.00-2.04; $P = .05$).

For Internet users in 2005, women were more likely to have communicated online with an health care provider compared to men (OR = 1.47; 95% CI = 1.00-2.15; $P = .05$), and cancer survivors were more likely to have used online patient-provider communication compared to those without a history of cancer (OR = 1.99; 95% CI = 1.27-3.12; $P = .002$). These results are consistent with bivariate analyses; however, education was not associated with online patient-provider communication in the multivariate model.

Table 3. Multivariate logistic regressions of having used online patient-provider communication in HINTS 2003 or HINTS 2005

	Odds of Communicating Online With an health care provider in the Past 12 Months			
	HINTS 2003 (n = 3527)		HINTS 2005 (n = 2649)	
	OR (95% CI)	P*	OR (95% CI)	P*
Age (years)		.33		.35
18-34	1.00		1.00	
35-49	0.82 (0.55-1.23)	.33	0.77 (0.44-1.33)	.34
50-64	1.14 (0.76-1.70)	.53	0.76 (0.41-1.42)	.38
65-74	0.57 (0.26-1.23)	.15	0.45 (0.18-1.13)	.09
75 or older	1.07 (0.30-3.77)	.91	0.47 (0.15-1.51)	.20
Gender				
Male	1.00		1.00	
Female	0.75 (0.56-1.02)	.07	1.47 (1.00-2.15)	.05
Education		< .001		.26
Less than high school	1.00		1.00	
High school graduate	1.20 (0.34-4.31)	.77	0.56 (0.16-1.95)	.35
Some college	2.44 (0.71-8.42)	.16	0.93 (0.26-3.35)	.91
College graduate	3.73 (1.10-12.59)	.03	0.99 (0.28-3.48)	.99
Annual Income (US \$)		.45		.34
< 25000	1.00		1.00	
25000 to < 35000	0.75 (0.36-1.55)	.42	1.35 (0.53-3.44)	.52
35000 to < 50000	0.59 (0.31-1.14)	.11	0.95 (0.49-1.88)	.89
50000 to < 75000	0.65 (0.37-1.16)	.14	1.09 (0.62-1.90)	.76
75000 or more	0.78 (0.44-1.36)	.37	1.56 (0.86-2.81)	.14
Race/Ethnicity		.98		.62
White	1.00		1.00	
Hispanic/Latino	0.92 (0.49-1.75)	.80	0.53 (0.16-1.72)	.28
African American	0.92 (0.53-1.59)	.75	1.26 (0.63-2.55)	.50
Asian American/Other [†]	1.00 (0.46-2.16)	.99	1.05 (0.43-2.58)	.91
Health Insurance				
Yes	1.00		1.00	
No	1.00 (0.52-1.90)	.99	0.99 (0.51-1.90)	.96
Health Status				
Excellent/very good/good	1.00		1.00	
Fair/poor	1.43 (1.00-2.04)	.05	0.88 (0.56-1.39)	.58
History of Cancer				
No	1.00		1.00	
Yes	1.34 (0.85-2.13)	.21	1.99 (1.27-3.12)	.002
Metropolitan Statistical Area				
County in metro area	1.00		1.00	
County in non-metro area	0.62 (0.41-0.95)	.03	0.76 (0.49-1.18)	.21

*P values reported for category headings for study variables with more than 2 categories refer to main effects.

[†]Other includes Pacific Islander, Native Hawaiian, American Indian, Alaska Native, and multiple races mentioned.

Discussion

Despite over a decade of research and the availability of guidelines for use of Internet-based communication by health care providers [28], the number of health care consumers using online patient-provider communication is still far below estimates of the number who would prefer to do so. Though data from HINTS suggest that use is slowly increasing, diffusion of online patient-provider communication is occurring at a pace far slower than diffusion of Internet use in general [29].

Thus, the question remains: Why is the overall prevalence of online communication with health care providers so low? While health care consumers and health care providers express concerns about communicating online, ratings of satisfaction and predictions about impact on health care quality regarding Internet-based communication have been generally favorable among both health care consumers and health care providers (eg, [5,30,31]). Therefore, use of online patient-provider communication will likely not significantly increase through efforts to change the primarily positive attitudes of health care consumers or health care providers, but rather, through changes in policies related to health care delivery [32] and through development of systems that prioritize usability [33]. Recent increased availability and adoption of online personal health records and electronic health records will likely affect the prevalence of online patient-provider communication [34,35], as will policies at the state and federal levels designed to promote diffusion of health information technology (eg, [36]). Continued implementation of policies that provide an architecture of support for online patient-provider communication and that address issues related to consumer and health care provider preferences, system interoperability, data security, and health care costs will be critical for maximizing the availability, adoption, and effectiveness of Internet-based communication between health care consumers and health care providers [32,33,35,37,38].

Associations between Internet users' sociodemographic and health-related characteristics and use of online patient-provider communication reveal insights regarding who may be taking the lead with online health care provider communication and who may be left behind. In 2003, Internet users with high levels of education were more likely to have communicated online with a health care provider, consistent with previous studies [5]. That education was nonsignificant in 2005 may suggest that health care consumers' level of education is less of a barrier to communicating online with health care providers as the prevalence of online patient-provider communication increases. Similarly, though Internet users residing in non-metro counties were less likely to have used online patient-provider communication in 2003, metropolitan statistical area was not associated with use in 2005. Deeper penetration of high-speed Internet access into more rural areas [39,40] may have decreased, over time, the degree to which location prevented online communication with health care providers. In both years, indicators of poorer health status (poor/fair self-reported health status, personal cancer history) were associated with online health care provider communication, suggesting that Internet

users with more medical problems or who are more engaged with the health care system due to a significant medical history may be more "hooked in" to Internet-based health communication resources or may have more a frequent need to use them. Finally, in 2005, women were more likely to use online patient provider communication compared to men. This result is consistent with findings that online women are more likely to search specifically for health information compared to men [41] and that higher percentages of women use the Internet for interpersonal communication related to health (eg, use of online support groups or health-based chat rooms [42]).

We did not observe associations between online communication with health care providers and characteristics such as race/ethnicity or annual income that have been documented in other studies (eg, [2,25]) as evidence of a "digital divide" [43,44]. Nonetheless, research and policy should continue to address groups potentially affected by the digital divide to ensure that advances in health information technology benefit all health care consumers [45]. Finally, our results were not consistent with previous studies that observed younger Internet users to be more likely to engage in online communication with health care providers (eg, [2,6]), suggesting a potential growth in comfort with online communication among Internet users of all ages.

Limitations

Though HINTS data are nationally representative, the generalizability of our results may be limited by survey response rates and the drop in response rates between 2003 and 2005. However, HINTS response rates are comparable to those of other national random digit dial surveys [46], and the agreement between our findings regarding prevalence of online patient-provider communication with other reports [17] supports the reliability of the HINTS estimates. Further, estimates of Internet penetration vary widely in the published literature; HINTS penetration estimates may be more conservative than data reported through market analysis firms due to the degree of sampling precision mandated for federal surveys that provide publicly available data. Due to item wording, we can only discuss our results at a generalized level of "online patient-provider communication" or "Internet-based communication with health care providers" and cannot characterize this behavior in more specific ways (eg, use of personal email, use of a Web portal) that could potentially affect our findings and resulting conclusions [38]. Finally, though HINTS provides a valuable population-level perspective on the prevalence of Internet-based health care provider communication and information on the characteristics of those who use it, all data are based on self-report, and HINTS does not allow for more in-depth examinations of barriers to communicating online with health care providers or the perceived benefits for those who do. To best meet the needs of patients and health care providers, research should continue to assess health care consumers' and health care providers' perspectives on barriers and benefits related to use of Internet-based communication as health information technology increasingly becomes part of standard medical care.

Conclusions

Online patient-provider communication is increasing slowly but remains uncommon. Though lower levels of education and non-metro county residence may have been barriers to using Internet-based communication with health care providers in 2003, by 2005, these barriers were not evident in HINTS. However, use of online patient-provider communication is higher among Internet users who are experiencing health problems or who have significant medical histories; health care

consumers without specific medical issues may need increased prompting to use Internet-based communication with health care providers as they engage in preventive health care. Changes in health care policy will be necessary to increase diffusion and adoption of online patient-provider communication, and a significant barrier continues to be Internet access. Disparities in Internet access must be addressed to ensure that increasing use of online patient-provider communication does not widen the digital divide or amplify disparities in health care quality for the underserved and underrepresented [25,45].

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 N01-CO-12400.

Conflicts of Interest

None declared.

References

1. Couchman GR, Forjuoh SN, Rascoe TG. E-mail communications in family practice: what do patients expect? *J Fam Pract* 2001 May;50(5):414-418. [Medline: [11350705](#)]
2. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns. *Am J Manag Care* 2004 Sep;10(9):601-608 [FREE Full text] [Medline: [15515992](#)]
3. Leong SL, Gingrich D, Lewis PR, Mauger DT, George JH. Enhancing doctor-patient communication using email: a pilot study. *J Am Board Fam Pract* 2005;18(3):180-188 [FREE Full text] [Medline: [15879565](#)]
4. Lin CT, Wittevrongel L, Moore L, Beaty BL, Ross SE. An Internet-based patient-provider communication system: randomized controlled trial. *J Med Internet Res* 2005;7(4):e47 [FREE Full text] [Medline: [16236699](#)] [doi: [10.2196/jmir.7.4.e47](#)]
5. Moyer CA, Stern DT, Dobias KS, Cox DT, Katz SJ. Bridging the electronic divide: patient and provider perspectives on e-mail communication in primary care. *Am J Manag Care* 2002 May;8(5):427-433 [FREE Full text] [Medline: [12019595](#)]
6. Moyer CA, Stern DT, Katz SJ, Fendrick AM. "We got mail": electronic communication between physicians and patients. *Am J Manag Care* 1999 Dec;5(12):1513-1522 [FREE Full text] [Medline: [11066618](#)]
7. Neill RA, Mainous AG, Clark JR, Hagen MD. The utility of electronic mail as a medium for patient-physician communication. *Arch Fam Med* 1994 Mar;3(3):268-271. [Medline: [8180718](#)] [doi: [10.1001/archfami.3.3.268](#)]
8. Sittig DF, King S, Hazlehurst BL. A survey of patient-provider e-mail communication: what do patients think? *Int J Med Inform* 2001 Apr;61(1):71-80. [Medline: [11248604](#)] [doi: [10.1016/S1386-5056\(00\)00134-9](#)]
9. Virji A, Yarnall KSH, Krause KM, Pollak KI, Scannell MA, Gradison M, et al. Use of email in a family practice setting: opportunities and challenges in patient- and physician-initiated communication. *BMC Med* 2006;4(1):18 [FREE Full text] [Medline: [16911780](#)] [doi: [10.1186/1741-7015-4-18](#)]
10. ; Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy; 2001.
11. 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](#)] [doi: [10.1016/S0749-3797\(98\)00108-1](#)]
12. 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* ;165(22):2618-2624. [Medline: [16344419](#)] [doi: [10.1001/archinte.165.22.2618](#)]
13. ; Pew Internet & American Life Project. *Internet Activities*. Washington, DC: Pew Internet & American Life Project URL: http://www.pewinternet.org/trends/Internet_Activities_7.19.06.htm [accessed 2007 Jun 22] [WebCite Cache ID 5KObiNqcR]
14. Horrigan J, Rainie L. *The Internet's Growing Role in Life's Major Moments*. Washington, DC: Pew Internet & American Life Project; 2006. URL: http://www.pewinternet.org/pdfs/PIP_Major%20Moments_2006.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ9f9z44]
15. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. *JAMA* 2003 May 14;289(18):2400-2406 [FREE Full text] [Medline: [12746364](#)] [doi: [10.1001/jama.289.18.2400](#)]
16. Brooks RG, Menachemi N. Physicians' use of email with patients: factors influencing electronic communication and adherence to best practices. *J Med Internet Res* 2006;8(1):e2 [FREE Full text] [Medline: [16585026](#)] [doi: [10.2196/jmir.8.1.e2](#)]

17. Fox S, Fallows D. Internet Health Resources. Washington, DC: Pew Internet & American Life Project; 2003. URL: http://www.pewinternet.org/PPF/r/95/report_display.asp [accessed 2007 Jun 25]
18. Grant RW, Campbell EG, Gruen RL, Ferris TG, Blumenthal D. Prevalence of basic information technology use by U.S. physicians. *J Gen Intern Med* 2006 Nov;21(11):1150-1155. [Medline: [16879417](#)] [doi: [10.1111/j.1525-1497.2006.00571.x](#)]
19. Gaster B, Knight CL, Dewitt DE, Sheffield JVL, Assefi NP, Buchwald D. Physicians' use of and attitudes toward electronic mail for patient communication. *J Gen Intern Med* 2003 May;18(5):385-389 [FREE Full text] [Medline: [12795738](#)] [doi: [10.1046/j.1525-1497.2003.20627.x](#)]
20. Neville RG, Marsden W, Mccowan C, Pagliari C, Mullen H, Fannin A. A survey of GP attitudes to and experiences of email consultations. *Inform Prim Care* 2004;12(4):201-206. [Medline: [15808021](#)]
21. Katz SJ, Nissan N, Moyer CA. Crossing the digital divide: evaluating online communication between patients and their providers. *Am J Manag Care* 2004 Sep;10(9):593-598 [FREE Full text] [Medline: [15515991](#)]
22. Hesse BW, Moser RP, Rutten LJJ, Kreps GL. The health information national trends survey: research from the baseline. *J Health Commun* 2006;11 Suppl 1(1):vii-xvi. [Medline: [16641070](#)] [doi: [10.1080/10810730600692553](#)]
23. Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The Health Information National Trends Survey (HINTS): development, design, and dissemination. *J Health Commun* 2004;9(5):443-60; discussion 81-4. [Medline: [15513791](#)] [doi: [10.1080/10810730490504233](#)]
24. Sciamanna CN, Rogers ML, Shenassa ED, Houston TK. Patient access to U.S. physicians who conduct internet or e-mail consults. *J Gen Intern Med* 2007 Mar;22(3):378-381 [FREE Full text] [Medline: [17356972](#)] [doi: [10.1007/s11606-006-0076-1](#)]
25. Hsu J, Huang J, Kinsman J, Fireman B, Miller R, Selby J, et al. Use of e-Health services between 1999 and 2002: a growing digital divide. *J Am Med Inform Assoc* 2005;12(2):164-171 [FREE Full text] [Medline: [15561786](#)] [doi: [10.1197/jamia.M1672](#)]
26. Home page. Health Information National Trends Survey. URL: <http://cancercontrol.cancer.gov/hints/> [accessed 2006 Nov 9] [WebCite Cache ID 5KH8LZ9CN]
27. ; Research Triangle Institute: SUDAAN Users Manual. 9th edition. Research Triangle Park, NC: Research Triangle Institute; 2004.
28. Kane B, Sands DZ. Guidelines for the clinical use of electronic mail with patients. The AMIA Internet Working Group, Task Force on Guidelines for the Use of Clinic-Patient Electronic Mail. *J Am Med Inform Assoc* 1998;5(1):104-111 [FREE Full text] [Medline: [9452989](#)]
29. ; Pew Internet & American Life Project. Percentage of U.S. Adults Online. Washington, DC: Pew Internet & American Life Project URL: http://www.pewinternet.org/trends/Internet_Adoption_4.26.07.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ8FaE7m]
30. Liederman EM, Morefield CS. Web messaging: a new tool for patient-physician communication. *J Am Med Inform Assoc* 2003;10(3):260-270 [FREE Full text] [Medline: [12626378](#)] [doi: [10.1197/jamia.M1259](#)]
31. Patt MR, Houston TK, Jenckes MW, Sands DZ, Ford DE. Doctors who are using e-mail with their patients: a qualitative exploration. *J Med Internet Res* 2003 May 15;5(2):e9 [FREE Full text] [Medline: [12857665](#)] [doi: [10.2196/jmir.5.2.e9](#)]
32. Weiss N. E-mail consultation: clinical, financial, legal, and ethical implications. *Surg Neurol* 2004 May;61(5):455-9; discussion 459. [Medline: [15120220](#)] [doi: [10.1016/j.surneu.2003.09.029](#)]
33. Hesse BW, Shneiderman B. eHealth research from the user's perspective. *Am J Prev Med* 2007 May;32(5 Suppl):S97-103. [Medline: [17466825](#)] [doi: [10.1016/j.amepre.2007.01.019](#)]
34. Reinke T. Rise of the PHR. *Manag Care* 2007 Jan;16(1):40-42. [Medline: [17285811](#)]
35. Hesse BW. Harnessing the power of an intelligent health environment in cancer control. *Stud Health Technol Inform* 2005;118:159-176. [Medline: [16301777](#)]
36. Pérez-Peña R. Big Medical Step, Starting With the City's Poor. *New York Times*. 2007 Apr 16. 8217. URL: <http://www.nytimes.com/2007/04/16/nyregion/16info.html?ex=1177819200&en=b84ae3a56371e4e5&ei=5070> [accessed 2007 Apr 20] [WebCite Cache ID 5OQ8VqbFN]
37. Valdes I, Kibbe DC, Tolleson G, Kunik ME, Petersen LA. Barriers to proliferation of electronic medical records. *Inform Prim Care* 2004;12(1):3-9. [Medline: [15140347](#)]
38. Katz SJ, Moyer CA. The emerging role of online communication between patients and their providers. *J Gen Intern Med* 2004 Sep;19(9):978-983 [FREE Full text] [Medline: [15333064](#)] [doi: [10.1111/j.1525-1497.2004.30432.x](#)]
39. Horrigan JB. Home Broadband Adoption 2006. Washington, DC: Pew Internet & American Life Project; 2006. May 28. URL: http://www.pewinternet.org/pdfs/PIP_Broadband_trends2006.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ8kI9TI]
40. Horrigan J, Murray K. Rural Broadband Internet Use [data memo]. Washington, DC: Pew Internet & American Life Project; 2006. URL: http://www.pewinternet.org/pdfs/PIP_Rural_Broadband.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ9Dh2bz]
41. Rutten LJJ, 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(1):147-156. [Medline: [16641080](#)] [doi: [10.1080/10810730600637574](#)]

42. Fallows D. How Women and Men Use the Internet. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/pdfs/PIP_Women_and_Men_online.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ8tt5Kb]
43. Fox S. Digital Divisions. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/pdfs/PIP_Digital_Divisions_Oct_5_2005.pdf [accessed 2007 Jun 22] [WebCite Cache ID 5OQ8zpLuH]
44. Viswanath K. Science and society: the communications revolution and cancer control. *Nat Rev Cancer* 2005 Oct;5(10):828-835. [Medline: [16195753](#)] [doi: [10.1038/nrc1718](#)]
45. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. *Am J Prev Med* 2007 May;32(5 Suppl):S131-S133. [Medline: [17466818](#)] [doi: [10.1016/j.amepre.2007.02.012](#)]
46. Nelson DE, Powell-Griner E, Town M, Kovar MG. A comparison of national estimates from the National Health Interview Survey and the Behavioral Risk Factor Surveillance System. *Am J Public Health* 2003 Aug;93(8):1335-1341 [FREE Full text] [Medline: [12893624](#)]

Abbreviations

HINTS: Health Information National Trends Survey

submitted 27.11.06; peer-reviewed by D Nelson, P Bansil; comments to author 21.12.06; revised version received 19.05.07; accepted 22.05.07; published 12.07.07.

Please cite as:

Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, Hesse BW

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

URL: <http://www.jmir.org/2007/3/e20/>

doi: [10.2196/jmir.9.3.e20](#)

PMID: [17627929](#)

© Ellen Burke Beckjord, Lila J Finney Rutten, Linda Squiers, Neeraj K Arora, Lindsey Volckmann, Richard P Moser, Bradford W Hesse. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 12.07.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

Response Rate and Completeness of Questionnaires: A Randomized Study of Internet Versus Paper-and-Pencil Versions

Sissel Marie Kongsved¹, MS; Maja Basnov¹, MS; Kurt Holm-Christensen², MD; Niels Henrik Hjollund^{1,3}, MD, PhD

¹Department of Clinical Social Medicine, Centre of Public Health, Central Denmark Region, Aarhus, Denmark

²Department of Radiology, Randers Regional Hospital, Randers, Denmark

³Department of Clinical Social Medicine, Institute of Public Health, Aarhus University, Aarhus, Denmark

Corresponding Author:

Niels Henrik Hjollund, MD, PhD

Department of Clinical Social Medicine

Institute of Public Health

Aarhus University

Olof Palmes Allé 15

8200 Aarhus N

Denmark

Phone: +45 87 28 47 40

Fax: +45 87 28 47 02

Email: nhh@folkesundhed.au.dk

Abstract

Background: Research in quality of life traditionally relies on paper-and-pencil questionnaires. Easy access to the Internet has inspired a number of studies that use the Internet to collect questionnaire data. However, Internet-based data collection may differ from traditional methods with respect to response rate and data quality as well as the validity and reliability of the involved scales.

Objective: We used a randomized design to compare a paper-and-pencil questionnaire with an Internet version of the same questionnaire with respect to differences in response rate and completeness of data.

Methods: Women referred for mammography at a Danish public hospital from September 2004 to April 2005, aged less than 67 years and without a history of breast cancer, were eligible for the study. The women received the invitation to participate along with the usual letter from the Department of Radiology. A total of 533 women were invited to participate. They were randomized to receive either a paper questionnaire, with a prepaid return envelope, or a guideline on how to fill in the Internet-based version online. The questionnaire consisted of 17 pages with a total of 119 items, including the Short Form-36, Multidimensional Fatigue Inventory-20, Hospital Anxiety and Depression Scale, and questions regarding social status, education level, occupation, and access to the Internet. Nonrespondents received a postal reminder giving them the option of filling out the other version of the questionnaire.

Results: The response rate before the reminder was 17.9% for the Internet group compared to 73.2% for the paper-and-pencil group (risk difference 55.3%, $P < .001$). After the reminder, when the participant could choose between versions of the questionnaire, the total response rate for the Internet and paper-and-pencil group was 64.2% and 76.5%, respectively (risk difference 12.2%, $P = .002$). For the Internet version, 97.8% filled in a complete questionnaire without missing data, while 63.4% filled in a complete questionnaire for the paper-and-pencil version (risk difference 34.5%, $P < .001$).

Conclusions: The Internet version of the questionnaire was superior with respect to completeness of data, but the response rate in this population of unselected patients was low. The general population has yet to become more familiar with the Internet before an online survey can be the first choice of researchers, although it is worthwhile considering within selected populations of patients as it saves resources and provides more complete answers. An Internet version may be combined with the traditional version of a questionnaire, and in follow-up studies of patients it may be more feasible to offer Internet versions.

(*J Med Internet Res* 2007;9(3):e25) doi:[10.2196/jmir.9.3.e25](https://doi.org/10.2196/jmir.9.3.e25)

KEYWORDS

Questionnaire design; random allocation; Internet; postal service; evaluation; data collection methodology

Introduction

Research in quality of life traditionally relies on paper-and-pencil questionnaires. Internet surveys may have advantages compared to the traditional paper-and-pencil surveys with respect to turn-a-round time, expenses, and data management [1]. However, Internet-based data collection may differ from traditional methods with respect to response rate and data quality as well as validity and reliability of the involved scales. Only a few studies have systematically evaluated Internet-based survey methods. The main questions have addressed validity [2-7], response rate, response speed, and completeness of data [1,6-14].

Most studies report small differences in answers obtained in Internet and paper-and-pencil versions of questionnaires [2-7]. Pealer et al found no significant difference in response rates, the Internet version having a response rate of 62% compared to 58% for the paper-and-pencil version [12]. Ritter et al found a high response rate in both groups of a study population recruited on the Internet: 87% in the Internet group, and 83% in the paper-and-pencil group [9]. These studies either recruited their participants on the Internet or invited only participants with a known active email account, and, as a consequence, the results from these studies are not valid for a general population of patients. A Swedish study conducted in a general population sample obtained a response rate of 50% in the Internet group and 64% in the paper-and-pencil group. The method included two reminders, including a contact by telephone [6]. However, in a workplace health survey, a poor response rate was observed among the Internet group (19%) compared to the paper-and-pencil group (72%), but this study did not include a reminder procedure [1].

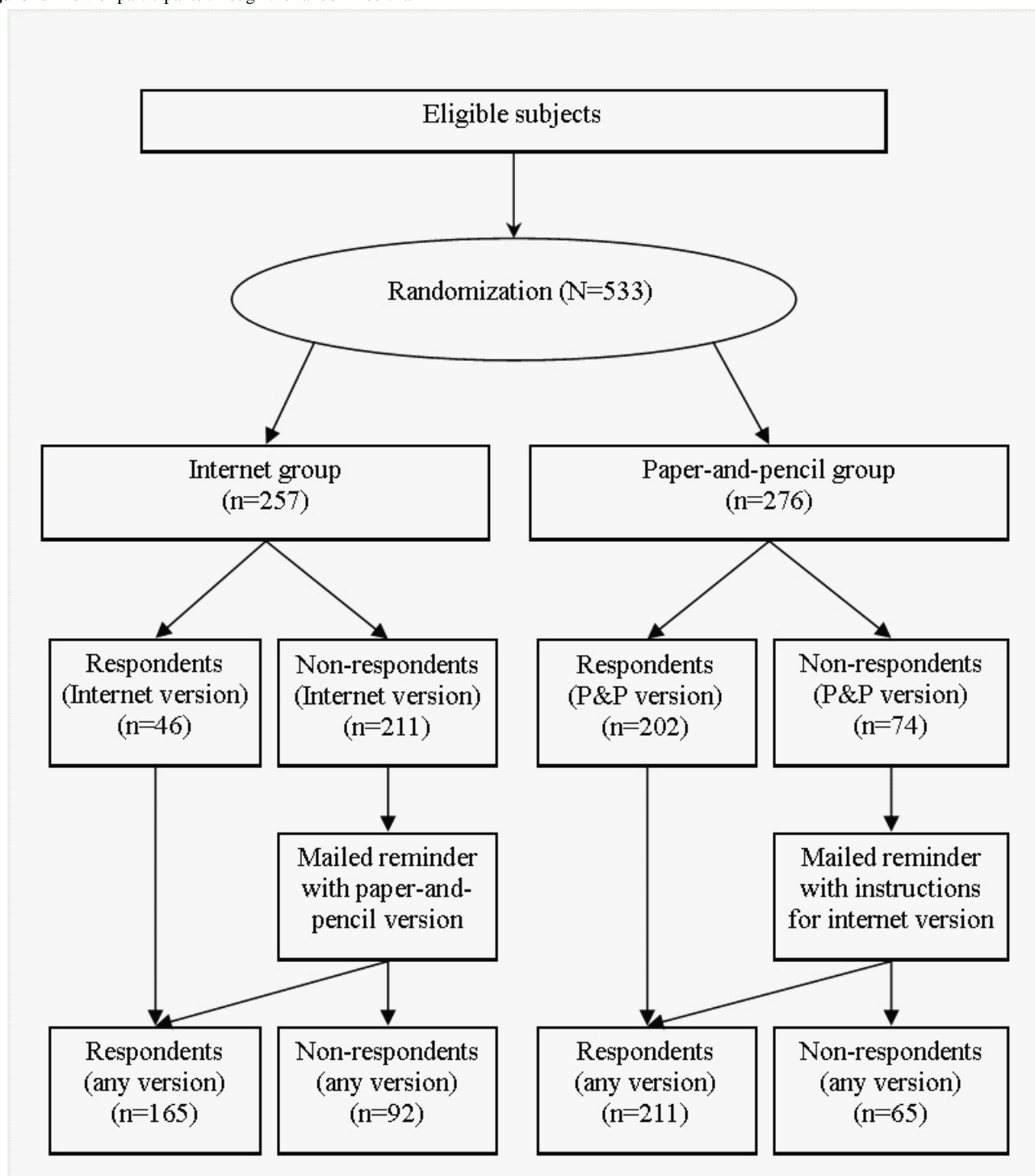
Overall, the results with respect to differences in response rate are inconsistent, which may reflect differences in methodology

and populations. We have not identified any randomized studies comparing Internet and paper-and-pencil questionnaires in patient populations unselected with respect to Internet access. Therefore, we aimed to evaluate an Internet survey method in comparison to paper-and-pencil with respect to response rate and completeness of data in a randomized controlled design among women referred for mammography.

Methods

Participants were women referred for mammography from September 2004 to April 2005 in the Department of Radiology at the public hospital, Randers Regional Hospital. The municipality of Randers has around 62000 inhabitants. Patients were referred by their family doctor. A consultant at the Department of Radiology assigned the referred patients to one of three categories: acute, subacute, or nonacute. Subsequently, a letter was sent to the woman informing her about the date, location, and other details of the mammography. The women were randomized to be invited to answer either an Internet version or a paper-and-pencil version of a questionnaire. We only invited women up to retirement age (67 years in Denmark) who did not have a history of breast cancer. Patients from all categories (acute, subacute, or nonacute) were invited until February 2005, whereafter only patients in the acute and subacute groups were invited to participate.

Nonrespondents in both groups were mailed a reminder after 10 days, given that the date of their mammography was not reached. The reminder informed the woman that she was free to answer the opposite version of the originally requested questionnaire if she so desired. Only questionnaires filled in before the date of the mammography were included in the analysis. The procedure is outlined in Figure 1. There were no incentives to promote the survey response.

Figure 1. Flow of participants through the randomized trial

The letter to women randomized to answer the paper-and-pencil version included a paper questionnaire and a prepaid return envelope, while the letter to women randomized to answer the Internet version included a guideline on how answer the Web-based version. Access to the Internet questionnaire required entry of a unique five-letter username. No password was needed since the first letter in the username was a redundancy code. The layout of the Internet version was as close to the paper version as possible (see [Figure 2](#) and [Figure 3](#)). In the Internet version, the participants were reminded of missing answers if they tried to leave a page incomplete. However, after pressing

an “OK” button, they were allowed to continue even if there were still missing answers [15]. The questionnaire consisted of 17 pages and 119 items and included Short Form-36 [16], Multidimensional Fatigue Inventory-20 [17], and The Hospital Anxiety and Depression Scale [18]. Questions regarding social status, education level, occupation, and access to the Internet were also asked.

All respondents were interviewed by telephone 1 month after they had their mammogram. They were invited to join a follow-up study and were asked to select the version of questionnaire they preferred.

Figure 2. Screenshot of the Internet version of the questionnaire

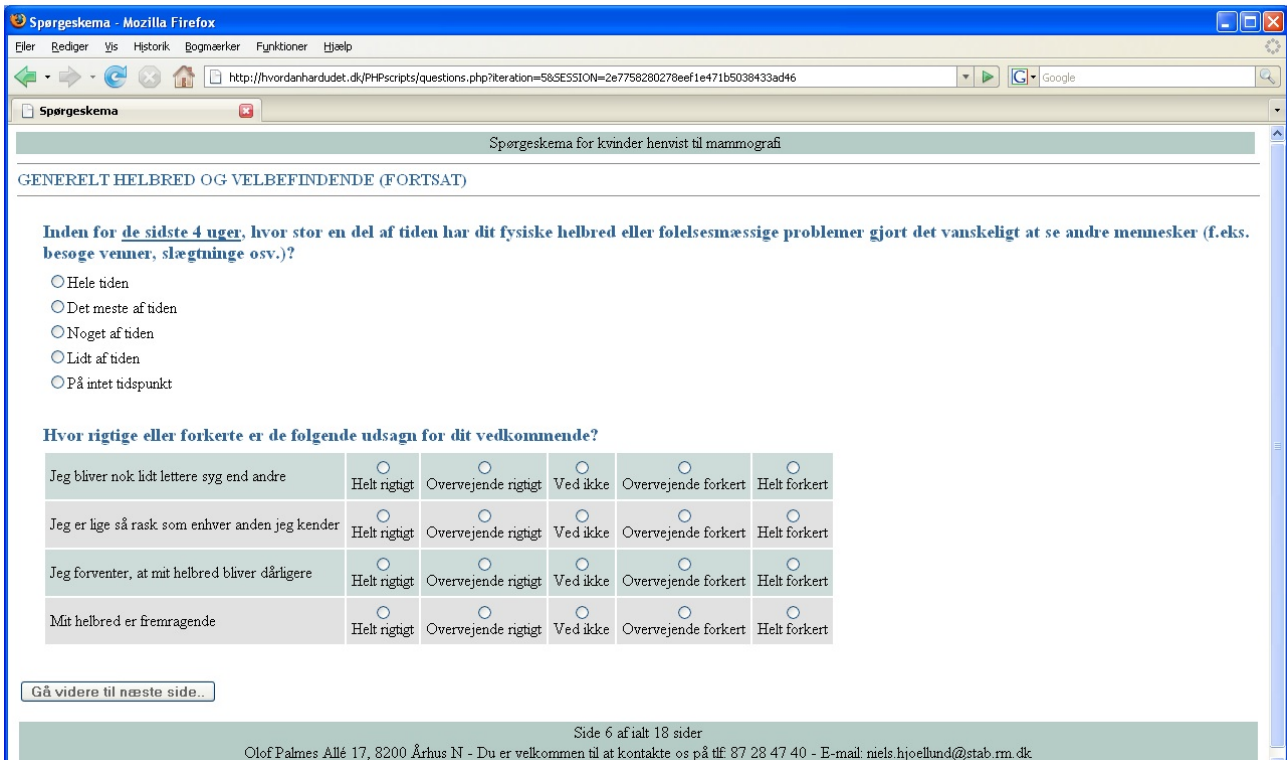
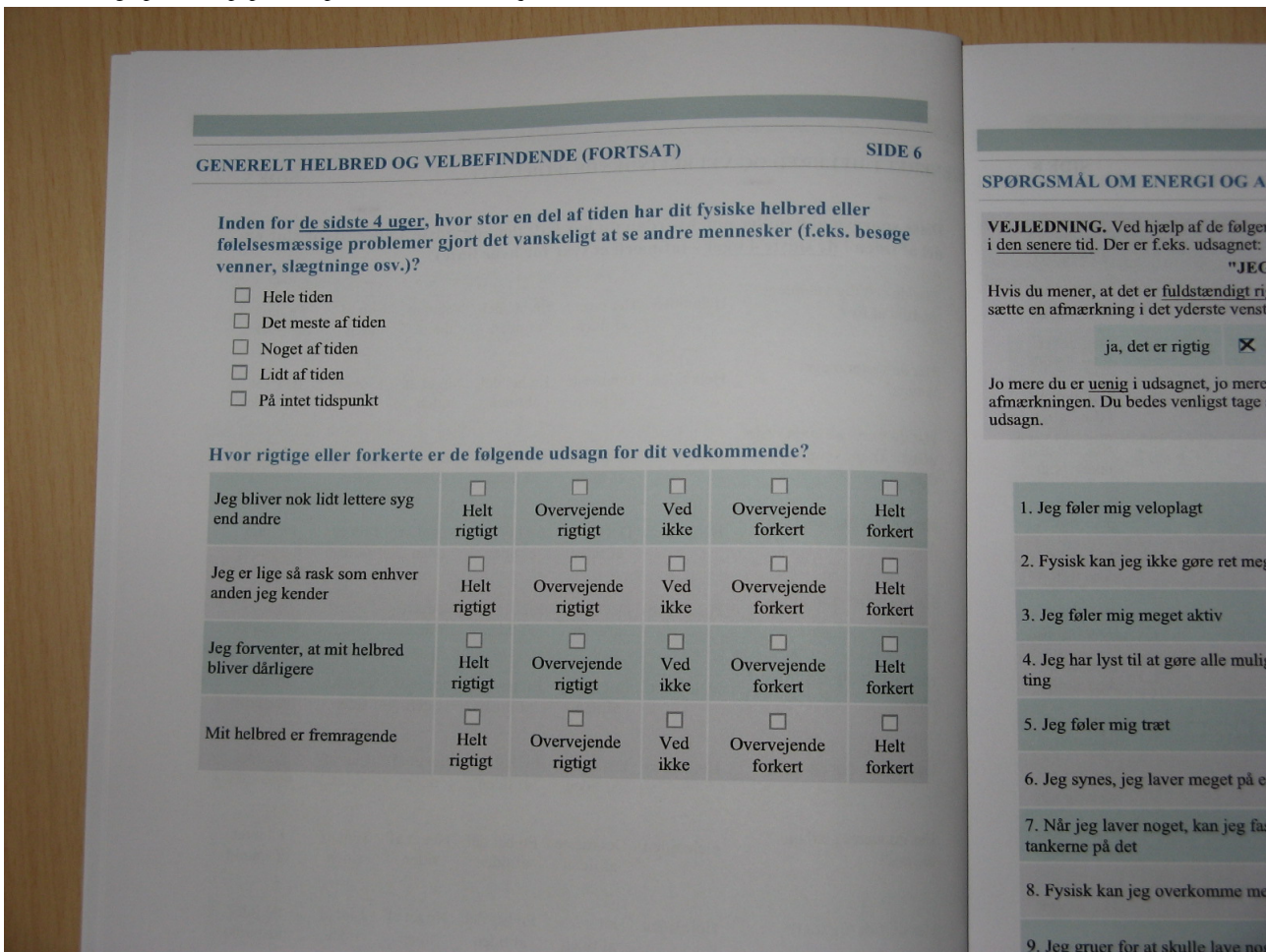


Figure 3. Photograph of the paper-and-pencil version of the questionnaire



The sample size was calculated to provide a statistical power of at least 90% to detect a true difference in response rate of 15%. The actual power was 93.8%. Women had an equal probability of assignment to the two groups. The randomization code was developed using a computer random number generator. We tested the significance of categorized variables by the chi-square test and compared continuous variables by risk differences with 95% confidence intervals. Homogeneity across strata was tested with the Mantel-Haenszel test.

Results

The characteristics of the invited women are shown in [Table 1](#). Approximately 80% of the women were between 30 and 59 years old. The distributions within the two randomized groups were similar with respect to age, place of residence, and category of referral.

Table 1. Characteristics of patients, by randomization group

	Internet Group, % (n = 257)	Paper-and-Pencil Group, % (n = 276)
Age (years)		
20-29	5.1	7.4
30-39	25.0	21.0
40-49	29.7	29.6
50-59	26.8	27.6
60-67	13.4	14.4
Place of Residence*		
Rural	60.1	53.3
Village/suburb	26.5	27.2
Urban	13.4	19.5
Category of Referral†		
Acute	50.0	47.9
Subacute	18.5	20.6
Nonacute	31.5	31.5

*Defined by postal code

†The acute group was called in for mammography within 3-14 days, the subacute group within 1 to 3 weeks, and the nonacute group, not before 5 months.

The response rate before the reminder was 17.9% in the Internet group compared to 73.2% in the paper-and-pencil group, corresponding to a 55.3% difference in response rate in favor of the paper-and-pencil version ([Table 2](#)). The same tendency was found in all strata with respect to age, place of residence, and category of referral (see [Table 2](#)).

After the reminder, the response rate improved distinctly in the group originally randomized to the Internet ([Table 3](#)). Among

the women assigned to the nonacute group, who had the longest respite before their mammogram, the response rate was even higher in the group randomized to the Internet version.

The completeness of answers in the two versions is summarized in [Table 4](#). The Internet version produced significantly more complete questionnaires than the paper-and-pencil version. For the paper-and-pencil version, there was a tendency toward more incomplete scales the longer the scales were.

Table 2. Response rate before reminder, by randomization group

	Internet Group, % (n = 257)	Paper-and-Pencil Group, % (n = 276)	Difference, % (95% CI)*
Total	17.9	73.2	55.3 (48.3-62.3)
Age (years)			
20-29	10.5	85.7	75.2 (52.2-98.1)
30-39	16.7	72.5	55.8 (41.3-70.3)
40-49	23.7	68.3	44.6 (30.7-58.5)
50-59	18.3	75.7	57.4 (44.1-70.7)
60-67	10.8	75.7	64.9 (47.8-81.9)
			$\chi^2_4 = 4.8, P = .30$
Place of Residence[†]			
Rural	17.5	74.7	57.2 (48.0-66.49)
Village/suburb	22.9	71.2	48.4 (34.1-62.7)
Urban	12.0	70.3	58.3 (41.0-75.5)
			$\chi^2_2 = 2.2, P = .33$
Category of Referral[‡]			
Acute	23.9	72.5	48.9 (38.3-59.5)
Subacute	13.2	70.6	57.4 (41.9-72.9)
Nonacute	12.4	75.9	63.5 (52.0-75.0)
			$\chi^2_2 = 5.1, P = .08$

*With Mantel-Haenszel test of homogeneity

[†]Defined by postal code

[‡]The acute group was called in for mammography within 3-14 days, the subacute group within 1 to 3 weeks, and the nonacute group, not before 5 months.

Table 3. Response rate after reminder, by randomization group

	Internet Group (n = 257)		Paper-and-Pencil Group (n = 276)		Difference in % (95% CI)
	%	% Reminded* (No.)	%	% Reminded (No.)	
Total	64.2	75 (159)	76.5	59 (44)	12.2 (4.5-20.0)
Category of Referral[†]					
Acute	49.6	58 (55)	74.6	39 (15)	25.0 (13.6-36.5)
Subacute	67.9	76 (35)	76.5	53 (8)	8.5 (-8.6-25.7)
Nonacute	84.0	97 (69)	79.3	100 (21)	-4.6 (-16.3-7.0)

*The percentage of primary nonrespondents who were reminded

[†]The acute group was called in within 3-14 days, the subacute group within 1 to 3 weeks, and the nonacute group, not before 5 months.

Table 4. Completeness of the three scales, by version

	Internet Version, % (n = 46)	Paper-and-Pencil Version, % (n = 202)	Difference, % (95% CI)
Total	97.8	63.4	34.5 (26.6-42.3)
Short Form-36 [16] (36 items)	100.0	71.3	28.7 (22.5-35.0)
Multidimensional Fatigue Inventory-20 [17] (20 items)	97.8	90.6	7.2 (1.4-13.1)
Hospital Anxiety and Depression Scale [18] (14 items)	100.0	92.6	7.4 (3.8-11.0)

The scores for the eight subscales of Short Form-36 are displayed in Table 5. There were no statistically significant differences between the two versions.

Table 5. Scores for subscales of Short Form-36, by randomization group

Subscale of Short Form-36 [16]	Internet Version, Mean (SD)	Paper-and-Pencil Version, Mean (SD)	Difference		
			%	<i>t</i>	<i>P</i>
Physical Function	91.4 (15.1)	90.1 (15.6)	1.3	0.5	.60
Role Physical	85.9 (30.6)	81.9 (30.3)	3.9	0.8	.43
Bodily Pain	81.2 (22.2)	76.3 (21.1)	4.9	1.4	.16
General Health	81.3 (14.7)	77.1 (18.7)	4.2	1.4	.16
Vitality	65.4 (22.7)	64.1 (22.2)	1.3	0.4	.71
Social Function	89.9 (16.4)	87.3 (19.0)	2.6	0.9	.39
Role Emotional	86.2 (25.9)	78.7 (32.8)	7.5	1.4	.14
Mental Health	75.2 (17.3)	71.8 (19.8)	3.4	1.1	.28

During the telephone interview with the respondents 1 month after they had their mammogram, they were invited to participate in the follow-up part of the study. They were asked to select the version of future questionnaires they preferred. The majority (55.4%) preferred the paper-and-pencil version, while 32.4% preferred the Internet version. The remaining 17.1% declined

further participation. Among the 46 respondents from the Internet group, 73.2% preferred to continue on the Internet compared to 17.1% who preferred to change to a paper-and-pencil version.

Access to Internet, estimated by answers from the paper-and-pencil group, is displayed in Table 6.

Table 6. Internet access among the paper-and-pencil group

	No.	At Home, %	Other*, %	None, %
Total	198	68.7	9.6	21.7
Age (years)				
20-29	12	58.3	25.0	16.7
30-39	48	70.8	8.4	20.8
40-49	56	82.1	3.6	14.3
50-59	55	63.6	18.2	18.2
60-65	27	51.9	0.0	48.1
$\chi^2_{8} = 25.6, P = .001$				
Place of Residence[†]				
Rural	120	62.5	12.5	25.0
Village	52	86.5	7.7	5.8
Urban	26	61.5	0	38.5
$\chi^2_{4} = 17.1, P = .002$				
Education Level (years)[‡]				
7-10	33	39.4	6.1	54.5
10-12	47	68.1	4.3	27.6
13-17	113	77.0	13.3	9.7
$\chi^2_{4} = 33.2, P < .001$				

* At work, local library, etc

† Defined by postal code

‡ According to International Standard Classification of Education

Discussion

We found an initial response rate of only 17.9% in the Internet group compared to 73.2% in the paper-and-pencil group. However, after a reminder, when the participants were free to choose between versions, the total response rate was similar in the two randomized groups. The quality of data regarding completeness was superior in the Internet version for all the involved scales. We did not identify any differences in Short Form-36 subscales. However, even in a randomized study, caution should be exercised when comparing the distribution of answers between the two groups since the distributions depend on differences in the two methods as well as selection bias, especially when the response rate in one of the groups is very low.

The population was unselected with respect to Internet access and experience. According to the 2005 Statistics Denmark survey, 77% of Danish women had access to the Internet [19]. Based on answers from the paper-and-pencil group, we estimate that 70% of the women in the present study had access to the Internet at home. Access was closely associated with level of education. The geographic area surrounding the public hospital includes rural locations as well as the fifth largest city in

Denmark. We consider our sample representative for female patients in Denmark.

The most prominent weakness of the Internet version was a low response rate, and we could not identify any single determinant factor. However, as expected, the response rate was highest in the age group with greatest access to the Internet. After a reminder letter, which stated that participants were free to fill out their preferred version of the questionnaire, the total response rates were nearly the same. However, women in the acute and subacute groups had less time to complete the questionnaire before their mammogram, which in some cases prevented the reminder.

Response rates to Internet questionnaires reported in the literature vary a lot between studies [1,6-14]. It is evident that studies conducted in populations with known access to the Internet are supposed to have higher response rate than studies of populations without known access, like the present study. However, differences in response rate may also be attributed to methodology and other characteristics of the population. A Swedish study compared the same paper-and-pencil questionnaire in two different versions with respect to ordering of questions and level of difficulty and found that the proportion of completers varied significantly [20]. It is plausible that populations of patients and general population samples may

react differently to an invitation to complete a Web questionnaire about health-related issues.

The fact that only 17.1% of respondents in the Internet group preferred to shift to the paper-and-pencil version when asked to join the follow-up study indicates that Internet versions may be more feasible in follow-up studies. One advantage of the Internet version is a high degree of completeness, and the design of Internet questionnaires allows the researcher to compensate for human error among participants who enter inconsistent answers or accidentally skip an item or even a page.

At present, Internet questionnaires can hardly stand alone as the method of data collection in studies of patients. Access to the Internet still depends on socioeconomic factors, and results obtained solely from Internet users may be biased. The general population must become more familiar with the Internet before an online survey can be the first choice of researchers, although it is worthwhile considering within selected populations of patients as it saves resources and provides more complete answers. An Internet version may be combined with a traditional version, and it may be more feasible to offer Internet versions in follow-up studies.

Acknowledgments

This work was supported by a grant from the Danish Cancer Society.

Conflicts of Interest

None declared.

References

1. Jones R, Pitt N. Health surveys in the workplace: comparison of postal, email and World Wide Web methods. *Occup Med (Lond)* 1999 Nov;49(8):556-558 [[FREE Full text](#)] [Medline: [10658310](#)] [doi: [10.1093/occmed/49.8.556](#)]
2. Fouladi RT, Mccarthy CJ, Moller NP. Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment* 2002 Jun;9(2):204-215. [Medline: [12066835](#)] [doi: [10.1177/10791102009002011](#)]
3. Bliven BD, Kaufman SE, Spertus JA. Electronic collection of health-related quality of life data: validity, time benefits, and patient preference. *Qual Life Res* 2001;10(1):15-22. [Medline: [11508472](#)] [doi: [10.1023/A:1016740312904](#)]
4. Davis RN. Web-based administration of a personality questionnaire: comparison with traditional methods. *Behav Res Methods Instrum Comput* 1999 Nov;31(4):572-577. [Medline: [10633976](#)]
5. Buchanan T, Smith JL. Using the Internet for psychological research: personality testing on the World Wide Web. *Br J Psychol* 1999 Feb;90 (Pt 1)(Pt 1):125-144. [Medline: [10085550](#)] [doi: [10.1348/000712699161189](#)]
6. Bälter KA, Bälter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiology* 2005 Jul;16(4):577-579. [Medline: [15951679](#)] [doi: [10.1097/01.ede.0000164553.16591.4b](#)]
7. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. *Eur J Epidemiol* 2006;21(2):103-111. [Medline: [16518678](#)] [doi: [10.1007/s10654-005-6030-4](#)]
8. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a randomized comparison (2). *J Med Internet Res* 2004 Sep 24;6(3):e30 [[FREE Full text](#)] [Medline: [15471756](#)] [doi: [10.2196/jmir.6.3.e30](#)]
9. Ritter P, Lorig K, Laurent D, Matthews K. Internet versus mailed questionnaires: a randomized comparison. *J Med Internet Res* 2004 Sep 15;6(3):e29 [[FREE Full text](#)] [Medline: [15471755](#)] [doi: [10.2196/jmir.6.3.e29](#)]
10. Boeckner LS, Pullen CH, Walker SN, Abbott GWA, Block T. Use and reliability of the World Wide Web version of the Block Health Habits and History Questionnaire with older rural women. *J Nutr Educ Behav* 2002 Mar;34(Suppl 1):S20-S24. [Medline: [12047825](#)] [doi: [10.1016/S1499-4046\(06\)60307-2](#)]
11. Harewood GC, Yacavone RF, Locke GR, Wiersema MJ. Prospective comparison of endoscopy patient satisfaction surveys: e-mail versus standard mail versus telephone. *Am J Gastroenterol* 2001 Dec;96(12):3312-3317. [Medline: [11774942](#)] [doi: [10.1111/j.1572-0241.2001.05331.x](#)]
12. Pealer LN, Weiler RM, Pigg RM, Miller D, Dorman SM. The feasibility of a web-based surveillance system to collect health risk behavior data from college students. *Health Educ Behav* 2001 Oct;28(5):547-559. [Medline: [11575685](#)] [doi: [10.1177/109019810102800503](#)]
13. Paolo AM, Bonaminio GA, Gibson C, Partridge T, Kallail K. Response rate comparisons of e-mail- and mail-distributed student evaluations. *Teach Learn Med* 2000;12(2):81-84. [Medline: [11228682](#)] [doi: [10.1207/S15328015TLM1202_4](#)]
14. Truell AD, Bartlett JE, Alexander MW. Response rate, speed, and completeness: a comparison of Internet-based and mail surveys. *Behav Res Methods Instrum Comput* 2002 Feb;34(1):46-49. [Medline: [12060989](#)]
15. Example of the Web-based questionnaire. *HvordanHarDuDet*. URL: <http://hvordanhardudet.dk/PHPscripts/testspm.htm> [accessed 2007 Jun 1] [[WebCite Cache ID 5O7tpDsPT](#)]
16. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-483. [Medline: [1593914](#)] [doi: [10.1097/00005650-199206000-00002](#)]

17. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995 Apr;39(3):315-325. [Medline: [7636775](#)] [doi: [10.1016/0022-3999\(94\)00125-O](#)]
18. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](#)] [doi: [10.1111/j.1600-0447.1983.tb09716.x](#)]
19. Access to the Internet 2005 [Befolkningens brug af Internet 2005]. Statistics Denmark. URL: http://www.dst.dk/upload/befolkningens_brug_af_internet_2005_002.pdf [accessed 2007 Jun 1] [WebCite Cache ID 5O3DQjy84]
20. Ekman A, Klint A, Dickman PW, Adami HO, Litton JE. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. *Eur J Epidemiol* 2007;22(5):293-300. [Medline: [17206467](#)] [doi: [10.1007/s10654-006-9091-0](#)]

submitted 08.11.06; peer-reviewed by J Olsen; comments to author 10.04.07; revised version received 23.04.07; accepted 18.05.07; published 31.09.07.

Please cite as:

Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH

Response Rate and Completeness of Questionnaires: A Randomized Study of Internet Versus Paper-and-Pencil Versions

J Med Internet Res 2007;9(3):e25

URL: <http://www.jmir.org/2007/3/e25/>

doi: [10.2196/jmir.9.3.e25](#)

PMID:

© Sissel Marie Kongsved, Maja Basnov, Kurt Holm-Christensen, Niels Henrik Hjollund. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 31.09.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

Impact of a Statewide Internet-Based Tobacco Cessation Intervention

Jessie E Saul¹, PhD; Barbara A Schillo¹, PhD; Sharrilyn Evered², PhD; Michael G Luxenberg³, PhD; Annette Kavanaugh³, MS; Nathan Cobb⁴, MD; Lawrence C An⁵, MD

¹ClearWay Minnesota, Minneapolis, MN, USA

²Center for Prevention, Blue Cross and Blue Shield of Minnesota, Saint Paul, MN, USA

³Professional Data Analysts, Inc, Minneapolis, MN, USA

⁴Pulmonary and Critical Care Unit, Massachusetts General Hospital, Boston, MA, USA

⁵University of Minnesota Medical School, Minneapolis, MN, USA

Corresponding Author:

Jessie E Saul, PhD
ClearWay Minnesota
8011 34th Ave S, Suite 400
Minneapolis, MN 55425
USA
Phone: +1 952 767 1415
Fax: +1 952 767 1422
Email: jsaul@clearwaymn.org

Abstract

Background: An increasing number of people have access to the Internet, and more people are seeking tobacco cessation resources online every year. Despite the proliferation of various online interventions and their evident acceptance and reach, little research has addressed their impact in the real world. Typically, low response rates to Internet-based follow-up surveys generate unrepresentative samples and large confidence intervals when reporting results.

Objectives: The aim of this study was to achieve a high response rate on follow-up evaluation in order to better determine the impact of an Internet-based tobacco cessation intervention provided to tobacco users in Minnesota, United States.

Methods: Participants included 607 men and women aged 18 and over residing in Minnesota who self-reported current tobacco use when registering for an Internet-based tobacco cessation program between February 2 and April 13, 2004. Participants were given access to an interactive website with features including social support, expert systems, proactive email, chat sessions, and online counselors. Mixed-mode follow-up (online survey with telephone survey for online nonrespondents) occurred 6 months after registration.

Results: Of the study participants, 77.6% (471/607) responded to the 6-month follow-up survey (39.4% online and 38.2% by telephone). Among respondents, 17.0% (80/471, 95% CI = 13.6%-20.4%) reported that they had not smoked in the past 7 days (observed rate). Assuming all nonrespondents were still smoking (missing=smoking rate), the quit rate was 13.2% (80/607, 95% CI = 10.5%-15.9%).

Conclusions: This mixed-mode follow-up survey of an online smoking cessation program achieved a high response rate and provides a more accurate estimate of long-term cessation rates than has been previously reported. Quit rates for the Internet-based tobacco cessation program were higher than those expected for unassisted quit attempts and are comparable to other evidence-based behavioral interventions. The similarities between quit rates demonstrates that an Internet-based cessation program may have as great an impact as, and can have wider reach than, other cessation programs such as those delivered by telephone. With over 100000 people having visited the website and over 23000 having registered, a 6-month self-reported quit rate of 13.2% suggests that the quitplan.com program helped over 3000 Minnesotans remain tobacco free for at least 6 months. Results of this study suggest that an Internet-based cessation program is a useful tool in states' efforts to provide comprehensive cessation tools for smokers.

(*J Med Internet Res* 2007;9(3):e28) doi:[10.2196/jmir.9.4.e28](https://doi.org/10.2196/jmir.9.4.e28)

KEYWORDS

Tobacco use cessation; Internet; behavior; evaluation studies

Introduction

An estimated 45.1 million Americans (20.9%) are current smokers [1]. More than 70% of US smokers want to quit, and 4 in 10 try to quit each year [1,2]. Unfortunately, most individuals who attempt to quit do so without receiving evidence-based treatments such as telephone quitlines, in-person counseling options, and pharmaceutical products [3]. Success rates for these unassisted quit attempts are low [4-7].

The Internet is a promising channel for improving delivery of tobacco treatment services. Approximately 60% of American adults reported having Internet access in the home in 2004 [8], and nearly 70% of US adults reported using the Internet at least occasionally in 2005 [9]. Searching for health information online is common [8], and it is estimated that as of 2004 over 8 million people had searched online for help to stop smoking [10].

The population impact of tobacco control programs is a product of both reach and effectiveness among participants [11]. Despite the proliferation of various online interventions and their evident acceptance and reach [12-16], little research has addressed their impact in the real world. Four recent randomized clinical trials have shown that individually tailored self-help materials delivered over the Internet result in modest increases in short-term abstinence [13,17-19], but information on longer term follow-up is limited. A number of demonstration and pilot projects of online cessation programs have been reported in the literature, with cessation rates ranging from 3% to 18% at time points ranging from 1 to 3 months [12,17,20,21]. However, results from these studies are difficult to interpret because of low response rates (10% to 56%) at follow-up [12,13,16-18,20,22].

While participant attrition is a usual, and even expected, aspect of online health-related applications [23], it poses a unique challenge for studies of online tobacco cessation interventions due to the strong association between response to follow-up and smoking status (more nonresponders are using tobacco than responders) [24-31]. Only one randomized clinical trial has produced response rates greater than 60% at 6 months: in this study, Muñoz et al [32] showed that tailored email messages increased the effectiveness of an online quit smoking guide. However, the study's use of monetary incentives to promote return to the site as well as the self-selected nature of the participants make the results less generalizable to the larger population of Internet users seeking help online to quit smoking.

ClearWay Minnesota, a nonprofit organization created as part of Minnesota's legal settlement with the tobacco industry, offers a range of statewide cessation services including Internet services through the quitplan.com website. Since the Internet

service was launched in July 2003, over 100000 individuals have visited the site, and over 23000 individuals have registered for the service, making it the most popular of ClearWay Minnesota's offerings [33]. At the time ClearWay Minnesota began providing Internet cessation services, there was limited information on the effectiveness of these programs. In response, ClearWay Minnesota contracted with an external evaluation firm to conduct an independent evaluation study of quitplan.com. The goal of this study was to determine a more precise estimate of the program's impact on its participants. This study was designed specifically to address gaps in the current literature by achieving a high response rate at a commonly used follow-up point (ie, 6 months after registration).

Methods

Quitplan.com Services

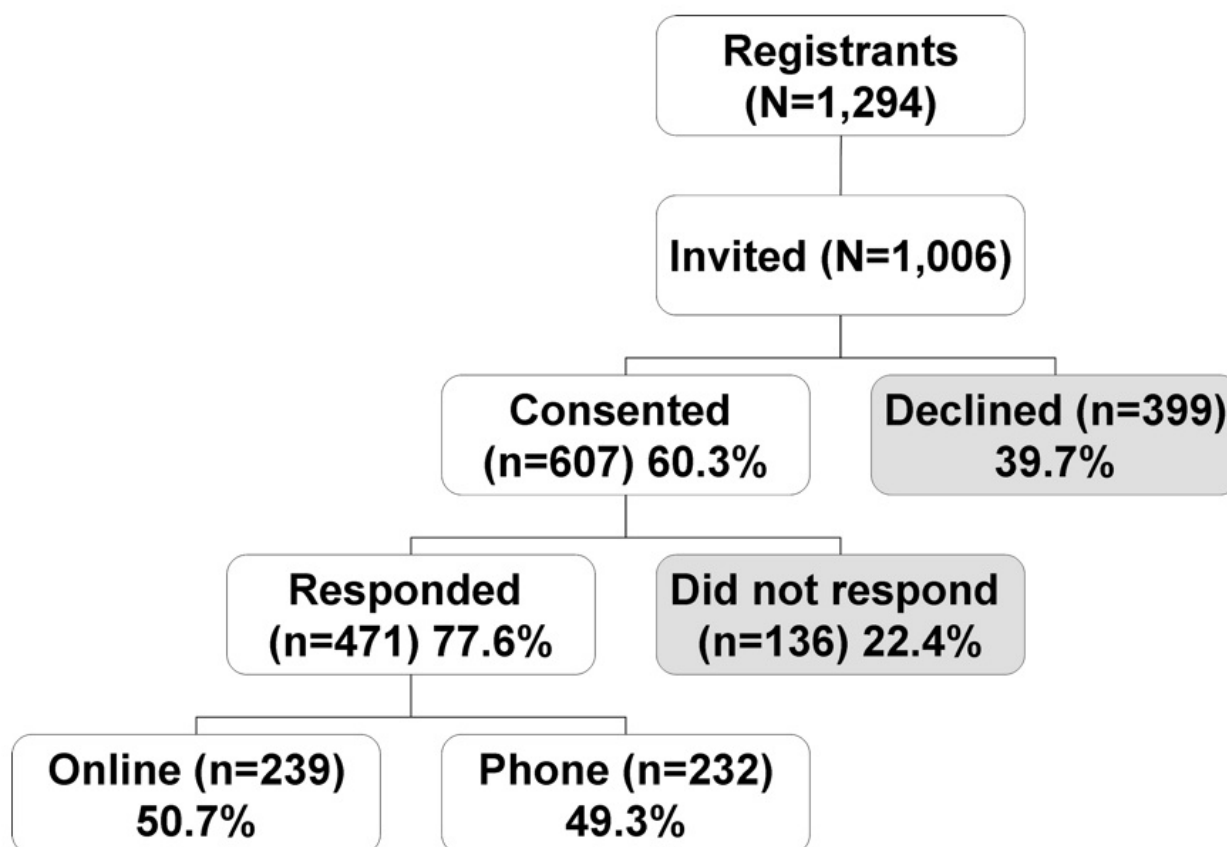
Content and programming for quitplan.com are provided by Healthways QuitNet Inc. The QuitNet service has been described elsewhere [16]. ClearWay Minnesota provides access to premium QuitNet services to all Minnesotans through a branded quitplan.com website. These services include online social support, expert systems, tailoring, and proactive email to enhance both cessation and relapse prevention. In addition, online counselors answer individual questions, and website staff moderate the forums and host chat sessions. Individuals in the quitplan.com program participate in the global social support community of all QuitNet powered websites.

Recruitment

All registrants included in the study (1) resided in Minnesota, (2) were at least 18 years old, (3) were accessing the site as a current tobacco user, and (4) did not report having already quit at the time of registration. All those reporting being in action or maintenance stages of tobacco cessation were excluded from the study. Of the 1294 registrants during the study period, 288 were not residents of Minnesota, were under 18 years old, were accessing the site on behalf of someone else, or self-identified as an evaluator or researcher. All 1006 eligible registrants who accessed the quitplan.com website between February 2 and April 13, 2004 were shown an additional screen during the registration process inviting them to participate in the study.

Figure 1 illustrates the flow of quitplan.com registrants through the consent and response protocol. A cohort of 607 individuals (60.3%) consented to participate in the study. The rate of consent for studies is often not well documented or reported in the literature. Of those who do report consent rates [13,20,32,34], the rates range between 20.9% and 76.8%. The consent rate for the current study is within the reported range.

Figure 1. Consent and response rates for study participants with registration dates between February 2 and April 13, 2004 (shaded boxes represent registrants for whom a full data set was not collected due to either lack of consent or response)



Evaluation

The study design consisted of a mixed-mode follow-up survey using email and, for those not responding by email, telephone. Participants were mailed a pre-notification letter 6 months after program registration and were then sent an email inviting them to complete an online evaluation survey. Reminder emails were sent to nonrespondents 3 and 7 days after this initial email, and attempts were made 12 days after the initial email to contact nonrespondents by phone to complete the follow-up evaluation. Up to 20 attempts to contact nonrespondents were made over multiple days at different times of day. All respondents received a US \$10 check for completing the follow-up survey.

Measures

Three data sources were used for this study: registration data, site usage data, and survey results. Demographic and clinical variables that were collected online at registration included age, gender, education, geographic region, employment status, health insurance status, marital status, readiness to quit (all types of tobacco), cigarettes per day, time to first morning cigarette, frequency of cigarette use, quit history (all types of tobacco), and quit attempts in the past year (all types of tobacco). In addition, this study examined the number of return visits to the site (log-ins) in the 6 months following registration, since the number of log-ins has been shown to be predictive of cessation outcomes in prior analyses [16]. Finally, the follow-up survey assessed quit status and other behaviors (7- and 30-day point

prevalence, use of medications or other quit aids since registration, number of quit attempts since registration, and duration of longest quit since registration) 6 months after study enrollment.

The primary outcome measure was self-reported 7-day point prevalence abstinence at 6 months post-registration. Quit status was assessed by self-report, which is consistent with the recommendation of the Society for Research on Nicotine and Tobacco Subcommittee on Biochemical Verification for low-demand studies [35]. Primary analysis of abstinence rates is by assuming all those lost to follow-up are still smoking (missing=smoking). Missing=smoking calculations are standard in the evaluation of cessation programs. We included all those who consented to participate in the study as the denominator for calculating the missing=smoking quit rate. Respondent-only (observed) quit rates are presented for comparison purposes.

Statistical Analysis

Analysis was performed using SPSS (Statistical Package for the Social Sciences) version 13.0 (SPSS Inc, Chicago, IL, USA). The analysis to calculate quit status and determine which variables were associated with quit status used both forced and forward step-wise logistic regressions. The impact of possible response bias on calculations of quit status was addressed by using a missing=smoking analysis that assumed that individuals not reached for follow-up were still smoking.

Registration and site usage data were examined to assess the degree that response bias influenced cessation outcomes. Baseline characteristics for respondents to the follow-up survey were compared to nonrespondents using chi-square tests for categorical variables and *t* tests or nonparametric tests for continuous variables, as appropriate.

Logistic regression was used to identify predictors of cessation. Several independent variables predictive of quit success in previous research include age [36], gender [37-39], education level [40,41], employment status [42,43], health insurance status [44,45], level of addiction [36,46], stage of readiness to quit [47], level of tobacco use [48,49], quit history [50,51], use of medications [4], and number of log-ins to Internet-based cessation programs [16]. Given the demonstrated correlations, a forced logistic regression analysis was conducted including these variables of interest. The forced model entered the variables in two blocks: all baseline variables (demographic and clinical) in the first block, and self-reported medication use (nicotine replacement therapy [NRT] or prescription medications such as Zyban) and the number of log-ins after registration (tricategorized) in the second block. The intent of this design was to test if log-ins and/or medication use was a significant predictor of quit status after adjusting for known baseline characteristics. Because the majority of registrants never logged in again after registration, the number of log-ins was divided into three categories (never logged in again after registration, 1-3 log-ins after registration, and 4 or more log-ins after registration) to provide groupings that both made sense cognitively and resulted in large enough numbers in each category to be able to conduct analysis. Forward step-wise logistic regression modeling was also performed and yielded essentially the same results. Therefore, the results of the a priori (ie, forced) model are presented here.

Institutional Review

This study was reviewed by the University of Minnesota's Institutional Review Board and determined to be exempt under federal guidelines 45 CFR 46.101 (b) for existing data.

Results

Response Rates and Response Bias

At 6 months, 471 of the 607 individuals in the study completed a follow-up survey, resulting in a final response rate of 77.6%

(39.4% online, 38.2% by telephone). Of those who did not respond online, over half (56.3%) were reached in 3 attempts, and nearly four fifths (79.0%) were reached in 6 attempts. Had the protocol included only 3 attempts, the total response rate would have been 70.7% (429/607). Sample size calculations indicated that a sample of 400 respondents was sufficient to determine 6-month quit rates with a 95% confidence level with a $\pm 5\%$ margin of error.

Participant Characteristics

Characteristics of respondents and nonrespondents are presented in Table 1. Some items were only assessed at follow-up (eg, marital status, medication use since registration). Of the 471 respondents, most were 25-44 years old (57.3%, 270/471), female (66.0%, 311/471), and lived in the 7-county Minneapolis/St. Paul metro area (63.6%, 288/453). A large majority had some college or postgraduate education (82.8%, 358/432) and were employed for wages (74.9%, 349/466). About half were married (50.4%, 237/470). Nearly all reported primary use of cigarettes (98.5%, 464/471), and most used tobacco daily (80.0%, 373/466). Roughly one quarter of respondents reported either light (27.6%, 130/471) or heavy (21.7%, 102/471) use of cigarettes, while about half reported moderate use (50.7%, 239/471).

Respondents were more likely to be older than nonrespondents (mean age 38.57 vs 35.75 years; $P = .008$), more likely to be insured (88.1% vs 80.9%; $P = .03$), and more likely to have ever quit for 30 days or more at registration (59.2% vs 47.1%; $P = .01$). There were no significant differences between respondents and nonrespondents in gender, geographic location, education level, employment status, type of tobacco used, daily versus occasional smoker, smoking intensity, time to first morning cigarette, stage of readiness to quit, 30-day quit in the past year, or ever quit for a year or more.

Nonresponders differed significantly from responders in terms of their utilization of the website. Nonresponders were more likely to have never logged in again after registration (68.4%, 93/136,) as compared to responders (47.1%, 222/471; $\chi^2_2 = 19.09$, $P < .001$).

Table 1. Quitplan.com 6-month follow-up survey participant characteristics

Variable	Respondents (n = 471)		Nonrespondents (n = 136)		P value * (respondents vs nonrespondents)	Participants (N = 607)	
	No.	%	No.	%		No.	%
Age group					.05		
18-24 years	55	12	27	20		82	14
25-44 years	270	57.3	78	57		348	57.3
45-64 years	139	29.5	30	22		169	27.8
65 or older	7	2	1	1		8	1
Gender					.12		
Female	311	66.0	80	59		391	64.4
Metro vs outstate					.81		
Outstate	165	36.4	47	38		212	36.7
7-county metro area	288	63.6	78	62		366	63.3
Education level (trichotomous)					.58		
High school or less	74	17	26	21		100	18.0
Some college	208	48.1	60	48		268	48.1
College graduate/postgraduate	150	34.7	39	31		189	33.9
Employed for wages					.95		
Yes (employed)	349	74.9	100	74.6		449	74.8
Health insurance status					.03		
Uninsured	54	12	25	19		79	14
Married (y/n)[†]					NA		
Yes (married)	237	50.4	NA			NA	
Primary form of tobacco used					.65 [‡]		
Cigarettes	464	98.5	135	99.3		599	98.7
Cigars	4	1	0	0		4	1
Pipe	1	0	0	0		1	0
Chewing tobacco or snuff	2	0	1	1		3	1
Cigarette use: daily or less than daily					.18		
Daily	373	80.0	115	85.2		488	81.2
Smoking intensity at registration					.09		
Light smoker (< 15 cigarettes/day)	130	27.6	38	28		168	27.7
Moderate smoker (15-24 cigarettes/day)	239	50.7	57	42		296	48.8
Heavy smoker (25+ cigarettes/day)	102	21.7	41	30		143	23.6
Time to first cigarette of the day (at registration)					.73		
Within 5 min	138	29.3	42	31		180	29.7
6-30 min	198	42.0	61	45		259	42.7
31-60 min	79	17	21	15		100	16.5
After 60 min	56	12	12	9		68	11
Stage of readiness to quit, 3 categories (at registration)					.75		
Precontemplation and contemplation	236	50.1	66	49		302	49.8
Preparation	235	49.9	70	52		305	50.2

Variable	Respondents (n = 471)		Nonrespondents (n = 136)		P value * (respondents vs nonrespondents)	Participants (N = 607)	
	No.	%	No.	%		No.	%
Attempted to quit in previous year (from registration)					.87		
Yes	298	63.3	85	63		383	63.1
Quit for 30 days or more in past 12 months?					.52		
Yes	43	9	10	7		53	9
Ever quit for 30 days or more?					.01		
Yes	279	59.2	64	47		343	56.5
Ever quit 1 year or more?					.38		
Yes	77	16	18	13		95	16
Number of log-ins in past 6 months (categorical)					< .001		
Never logged in	222	47.1	93	68		315	51.9
1-3 log-ins	149	31.6	28	21		177	29.2
4 or more log-ins	100	21.2	15	11		115	18.9
Cessation medication pattern (as reported at 6-month follow-up)[†]					NA		
None reported	243	51.6	NA			NA	
NRT only	138	29.3	NA			NA	
Zyban only	56	12	NA			NA	
NRT and Zyban	34	7	NA			NA	

NA, not available.

*P-values are from χ^2 statistics.

[†]Only asked at follow-up.

[‡]This calculation was done on a very small number of cases with a highly skewed distribution leading to small marginal expected values.

Cessation Outcomes

Both respondent-only and missing=smoking quit rates are presented here. Among respondents, 17.0% (80/471, 95% CI = 13.6%-20.4%) reported that they had not smoked in the past 7 days at the time of the 6-month follow-up. Using a missing=smoking analysis, the quit rate is 13.2% (80/607, 95% CI = 10.5%-15.9%).

There were no differences between telephone and online respondents in terms of any of the three cessation outcomes

(7-day point prevalence, 30-day point prevalence, or 30-day abstinence at some point during the past 6 months).

Results from the logistic regression model predicting 7-day abstinence at 6 months are shown in Table 2. The only variable with a significant odds ratio for 7-day abstinence at follow-up was "number of log-ins after registration." The odds of having quit were 2.90 (95% CI = 1.45-5.77) times higher for those logging in four or more times after registration at 6 months than for those who never logged in again after registration.

Table 2. Odds ratios for forced logistic regression model for 7-day abstinence at 6 months (N = 417)*

	<i>P</i> value	Odds Ratio (95.0% CI)
Number of log-ins		
1-3 log-ins in past 6 months vs none	.38	1.35 (0.69-2.67)
4+ log-ins in past 6 months vs none	.002	2.90 (1.45-5.77)
Age		
Age at registration	.29	0.99 (0.96-1.01)
Gender		
Gender (female vs male)	.91	1.04 (0.57-1.89)
Education		
Some college vs high school or less	.82	1.10 (0.49-2.47)
College graduate/postgraduate vs high school or less	.90	0.94 (0.39-2.27)
Employment		
Employed for pay	.84	1.08 (0.54-2.14)
Health insurance status		
Insured	.15	2.29 (0.75-7.02)
Time to first cigarette of the day		
6-30 min vs within 5 min	.63	1.19 (0.60-2.37)
31-60 min vs within 5 min	.27	0.58 (0.22-1.54)
After 60 min vs within 5 min	.48	1.43 (0.53-3.90)
Stage of change		
Preparation vs contemplation/precontemplation	.96	1.01 (0.57-1.80)
Smoking intensity		
Moderate smoker (15-24 cigs/day) vs light smoker (1-14 cigs/day)	.51	0.79 (0.39-1.58)
Heavy smoker (25+ cigs/day) vs light smoker (1-14 cigs/day)	.35	0.64 (0.25-1.64)
Quit history		
Ever quit for 30 days or more	.47	1.25 (0.69-2.26)
Use of medications		
Used meds (NRT and/or Zyban) in past 6 months	.10	1.63 (0.91-2.95)

* 11.5% excluded due to missing; Nagelkerke $R^2 = 0.102$.

Discussion

Principal Results

Several studies suggest that tobacco cessation programs can be delivered effectively via the Internet [13,18,19,32]. However, with the exception of the study by Muñoz et al [32], these studies were limited by large differences between observed and missing=smoking quit rates. For example, Cobb et al (2005) achieved a response rate of only 25.6% at 3 months, resulting in an observed quit rate of 30% and a missing=smoking rate of 7% [16]. The current study achieved a markedly higher response rate (78%), substantially closing the gap between observed (17.0%) and missing=smoking rates (13.2%).

By using a mixed-mode methodology for follow-up at 6 months, the present study resulted in a higher response rate, thus increasing our confidence in the precision of the estimated quit rate. Similar to a recent study by Couper et al (2007), we found

that many of those lost to online follow-up can be “brought back” through alternate modes of data collection [52]. It should be noted that our study, while not designed to test for mode effects, found no differences between telephone and online respondents in terms of any of the three cessation outcomes. Couper et al, however, found evidence that use of the telephone produced more socially desirable responses on weight loss outcomes when compared to mail as an alternate mode to online follow-up, pointing to the need to carefully consider mode effects in any future studies of online tobacco cessation interventions.

For the present study, we conclude that the best estimate of 7-day abstinence at 6 months after registration is between 13% (assuming missing=smoking) and 17% (among respondents only, ie, observed). Some have suggested that missing=smoking is an overly conservative approach for follow-up surveys because not all individuals who fail to answer the surveys continue to smoke [53,54]. This is particularly the case when

the goal is to evaluate the effectiveness of real-world programs and not to compare different groups in a trial setting. The results of missing=smoking analyses may be considered the lower estimate of program impact just as the observed quit results based on respondents only may be considered an upper estimate.

In the present study, the missing=smoking estimate is comparable to data from evaluations of other cessation programs funded by ClearWay Minnesota. For example, the missing=smoking quit rate for ClearWay Minnesota's QUITPLAN Helpline (prior to the introduction of free NRT to under- and uninsured callers) was 11.0% [55]. The similarities between quit rates suggest that an Internet-based cessation program may have a greater impact than behavioral-based telephone quitlines that do not provide NRT, given their greater reach (ie, easy access, availability, and participation) and noting that both types of programs produce comparable quit rates.

Several clinical trials are in progress to more fully evaluate the effectiveness of Internet-based cessation programs. Data from these trials will help to identify which elements of Internet-based tobacco cessation programs are critical for enhancing quit success. It may be that certain features or content may reduce effectiveness of the program, as has been shown in other studies [32].

In the present study, the number of log-ins was significantly correlated to quit status. It is interesting to note that an independently verifiable behavior occurring after registration, as opposed to baseline demographics or tobacco history, was predictive of quit status. Additional research is needed to determine the existence and direction of a causal relationship between log-ins and quitting and whether mechanisms for getting people to return to an Internet cessation website might increase the efficacy of the intervention.

Limitations

As an observational study, participants were not randomized into a control group, which limits the conclusions that can be drawn regarding effectiveness of the quitplan.com website. Muñoz et al have already shown that Internet-based cessation programs are effective [32]. However, this study addresses one limitation of the Muñoz study in that it measures quit rates for Internet users outside of the context of a randomized clinical trial and can thus be more easily generalized to users of Internet-based cessation programs in the real world.

Individuals who agreed to participate in a program evaluation did differ in terms of demographic or smoking-related characteristics and outcomes from those who did not agree to participate. As a result, participants and nonparticipants may be expected to differ in terms of their cessation outcomes. Given that only 60% of those who were invited to participate consented to do so, it may be that all site users have a different rate of abstinence than the subset of those who consented to participate. Future studies should consider strategies to increase initial consent rates to further improve generalizability.

Conclusions

This mixed-mode survey produced a high response rate, resulting in more accurate estimates of long-term cessation rates than previously reported. Quit rates for the Internet-based tobacco cessation program were better than those expected for unassisted quit attempts and are comparable to other evidence-based interventions. With over 100000 people having visited the site and over 23000 having registered since inception of the program in 2003, a 6-month self-reported quit rate of 13.2% suggests that the quitplan.com program has helped over 3000 Minnesotans remain tobacco free for at least 6 months. Results of this study suggest that an Internet-based cessation program is a useful tool in states' efforts to provide comprehensive cessation programs for smokers.

Acknowledgments

The authors would like to thank Dr. Michael Fiore and Paula Keller for their comments on the manuscript.

Funding for the Internet-based tobacco cessation program quitplan.com and its evaluation was provided by ClearWay Minnesota. ClearWay Minnesota is an independent nonprofit organization and is funded with 3% of the 1998 Minnesota tobacco settlement. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of ClearWay Minnesota.

Conflicts of Interest

Dr. Cobb served as Chief Medical Officer of QuitNet, Inc. and held a financial interest in the company during the period of the study. Professional Data Analysts, Inc. was contracted by ClearWay Minnesota to conduct the evaluation. All other authors have no personal financial interests or conflicts of interest to declare.

References

1. ; Centers for Disease Control and Prevention (CDC). Tobacco use among adults--United States, 2005. *MMWR Morb Mortal Wkly Rep* 2006 Oct 27;55(42):1145-1148 [FREE Full text] [Medline: [17065979](#)]
2. ; Centers for Disease Control and Prevention (CDC). Cigarette smoking among adults--United States, 2000. *MMWR Morb Mortal Wkly Rep* 2002 Jul 26;51(29):642-645 [FREE Full text] [Medline: [12186222](#)]

3. Fiore MC, Novotny TE, Pierce JP, Giovino GA, Hatzianandreu EJ, Newcomb PA, et al. Methods used to quit smoking in the United States. Do cessation programs help? *JAMA* 1990 May 23;263(20):2760-2765. [Medline: [2271019](#)] [doi: [10.1001/jama.263.20.2760](#)]
4. Fiore MC. US public health service clinical practice guideline: treating tobacco use and dependence. *Respir Care* 2000 Oct;45(10):1200-1262. [Medline: [11054899](#)]
5. Hughes JR, Gulliver SB, Fenwick JW, Valliere WA, Cruser K, Pepper S, et al. Smoking cessation among self-quitters. *Health Psychol* 1992;11(5):331-334. [Medline: [1425551](#)] [doi: [10.1037/0278-6133.11.5.331](#)]
6. Zhu S, Melcer T, Sun J, Rosbrook B, Pierce JP. Smoking cessation with and without assistance: a population-based analysis. *Am J Prev Med* 2000 May;18(4):305-311. [Medline: [10788733](#)] [doi: [10.1016/S0749-3797\(00\)00124-0](#)]
7. Pierce JP, Gilpin EA. Impact of over-the-counter sales on effectiveness of pharmaceutical aids for smoking cessation. *JAMA* 2002 Sep 11;288(10):1260-1264 [FREE Full text] [Medline: [12215133](#)] [doi: [10.1001/jama.288.10.1260](#)]
8. Fox S. Health information online: eight in ten internet users have looked for health information online, with increased interest in diet, fitness, drugs, health insurance, experimental treatments, and particular doctors and hospitals. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/pdfs/PIP_Healthtopics_May05.pdf [accessed 2007 Mar 7] [WebCite Cache ID 5NAh3q1Gg]
9. Fox S. Digital divisions: there are clear differences among those with broadband connections, dial-up connections, and no connections at all to the Internet. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/pdfs/PIP_Digital_Divisions_Oct_5_2005.pdf [accessed 2007 Mar 7] [WebCite Cache ID 5NAjXq6VI]
10. Fox S, Fallows D. Internet health resources: health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Washington, DC: Pew Internet & American Life Project; 2003. URL: http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf [accessed 2007 Sep 24] [WebCite Cache ID 5NBrkgAuo]
11. Abrams DB, Orleans C, Niaura R, Goldstein MG, Prochaska JM, Velicer WF. Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped-care and matching model. *Ann Behav Med* 1996;18(4):290-304.
12. Feil EG, Noell J, Lichtenstein E, Boles SM, McKay HG. Evaluation of an Internet-based smoking cessation program: lessons learned from a pilot study. *Nicotine Tob Res* 2003 Apr;5(2):189-194. [Medline: [12745491](#)] [doi: [10.1080/1462220031000073694](#)]
13. 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. [Medline: [15847626](#)] [doi: [10.1111/j.1360-0443.2005.01093.x](#)]
14. Zabinski MF, Wilfley DE, Pung MA, Winzelberg AJ, Eldredge K, Taylor CB. An interactive internet-based intervention for women at risk of eating disorders: a pilot study. *Int J Eat Disord* 2001 Sep;30(2):129-137. [Medline: [11449446](#)] [doi: [10.1002/eat.1065](#)]
15. Zabinski MF, Wilfley DE, Calfas KJ, Winzelberg AJ, Taylor CB. An interactive psychoeducational intervention for women at risk of developing an eating disorder. *J Consult Clin Psychol* 2004 Oct;72(5):914-919. [Medline: [15482051](#)] [doi: [10.1037/0022-006X.72.5.914](#)]
16. Cobb NK, Graham AL, Bock BC, Papanonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. *Nicotine Tob Res* 2005 Apr;7(2):207-216. [Medline: [16036277](#)] [doi: [10.1080/14622200500055319](#)]
17. Lenert L, Muñoz RF, Perez JE, Bansod A. Automated e-mail messaging as a tool for improving quit rates in an internet smoking cessation intervention. *J Am Med Inform Assoc* 2004 Jul;11(4):235-240 [FREE Full text] [Medline: [15064291](#)] [doi: [10.1197/jamia.M1464](#)]
18. Swartz LHG, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. *Tob Control* 2006 Feb;15(1):7-12 [FREE Full text] [Medline: [16436397](#)] [doi: [10.1136/tc.2003.006189](#)]
19. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. *J Med Internet Res* 2005 Mar 8;7(1):e2 [FREE Full text] [Medline: [15829474](#)] [doi: [10.2196/jmir.7.1.e2](#)]
20. Stoddard J, Delucchi K, Muñoz R, Collins N, Stable EP, Augustson E, et al. Smoking cessation research via the internet: a feasibility study. *J Health Commun* 2005 Jan;10(1):27-41. [Medline: [15764442](#)] [doi: [10.1080/10810730590904562](#)]
21. Woodruff SI, Edwards CC, Conway TL, Elliott SP. Pilot test of an Internet virtual world chat room for rural teen smokers. *J Adolesc Health* 2001 Oct;29(4):239-243. [Medline: [11587907](#)] [doi: [10.1016/S1054-139X\(01\)00262-2](#)]
22. Lenert L, Muñoz RF, Stoddard J, Delucchi K, Bansod A, Skoczen S, et al. Design and pilot evaluation of an internet smoking cessation program. *J Am Med Inform Assoc* 2003 Jan;10(1):16-20 [FREE Full text] [Medline: [12509354](#)] [doi: [10.1197/jamia.M1128](#)]
23. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [FREE Full text] [Medline: [15829473](#)] [doi: [10.2196/jmir.7.1.e11](#)]
24. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. *Eur J Epidemiol* 2006;21(2):103-111. [Medline: [16518678](#)] [doi: [10.1007/s10654-005-6030-4](#)]

25. Brøgger J, Bakke P, Eide GE, Gulsvik A. Contribution of follow-up of nonresponders to prevalence and risk estimates: a Norwegian respiratory health survey. *Am J Epidemiol* 2003 Mar 15;157(6):558-566 [FREE Full text] [Medline: [12631547](#)] [doi: [10.1093/aje/kwg003](#)]
26. Nakai S, Hashimoto S, Murakami Y, Hayashi M, Manabe K, Noda H. [Response rates and non-response bias in a health-related mailed survey]. *Nippon Koshu Eisei Zasshi* 1997 Mar;44(3):184-191. [Medline: [9175409](#)]
27. Hill A, Roberts J, Ewings P, Gunnell D. Non-response bias in a lifestyle survey. *J Public Health Med* 1997 Jun;19(2):203-207 [FREE Full text] [Medline: [9243437](#)]
28. O'Neill TW, Marsden D, Silman AJ. Differences in the characteristics of responders and non-responders in a prevalence survey of vertebral osteoporosis. European Vertebral Osteoporosis Study Group. *Osteoporos Int* 1995;5(5):327-334. [Medline: [8800782](#)] [doi: [10.1007/BF01622254](#)]
29. Hardie JA, Bakke PS, Mørkve O. Non-response bias in a postal questionnaire survey on respiratory health in the old and very old. *Scand J Public Health* 2003;31(6):411-417. [Medline: [14675932](#)] [doi: [10.1080/14034940210165163](#)]
30. Rönmark E, Lundqvist A, Lundbäck B, Nyström L. Non-responders to a postal questionnaire on respiratory symptoms and diseases. *Eur J Epidemiol* 1999 Mar;15(3):293-299. [Medline: [10395061](#)] [doi: [10.1023/A:1007582518922](#)]
31. Kotaniemi JT, Hassi J, Kataja M, Jönsson E, Laitinen LA, Sovijärvi AR, et al. Does non-responder bias have a significant effect on the results in a postal questionnaire study? *Eur J Epidemiol* 2001;17(9):809-817. [Medline: [12081098](#)] [doi: [10.1023/A:1015615130459](#)]
32. 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. [Medline: [16497602](#)] [doi: [10.1080/14622200500431940](#)]
33. Betzner A, Rainey J, Luxenberg M, Greenesid L, Christenson M, Kavanaugh A. Evaluation of the QUITPLANSM Programs: Final "Synthesis" Report: QUITPLAN Helpline, QUITPLAN Centers, QUITPLAN at Work, and quitplan.com. Minneapolis, MN: ClearWay Minnesota; 2004.
34. 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(Suppl 1):S43-S48. [Medline: [17491170](#)]
35. ; SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine Tob Res* 2002 May;4(2):149-159. [Medline: [12028847](#)] [doi: [10.1080/14622200210123581](#)]
36. Hyland A, Li Q, Bauer JE, Giovino GA, Steger C, Cummings KM. Predictors of cessation in a cohort of current and former smokers followed over 13 years. *Nicotine Tob Res* 2004 Dec;6 Suppl 3(Suppl 3):S363-S369. [Medline: [15799599](#)]
37. Perkins K. Sex differences in nicotine versus nonnicotine reinforcement as determinants of tobacco smoking. *Exp Clin Psychopharmacol* 1996;4(2):166-177. [doi: [10.1037/1064-1297.4.2.166](#)]
38. Wetter DW, Kenford SL, Smith SS, Fiore MC, Jorenby DE, Baker TB. Gender differences in smoking cessation. *J Consult Clin Psychol* 1999 Aug;67(4):555-562. [Medline: [10450626](#)] [doi: [10.1037/0022-006X.67.4.555](#)]
39. Gritz ER, Nielsen IR, Brooks LA. Smoking cessation and gender: the influence of physiological, psychological, and behavioral factors. *J Am Med Womens Assoc* 1996 Jan;51(1-2):35-42. [Medline: [8868547](#)]
40. Tillgren P, Haglund BJ, Lundberg M, Romelsjö A. The sociodemographic pattern of tobacco cessation in the 1980s: results from a panel study of living condition surveys in Sweden. *J Epidemiol Community Health* 1996 Dec;50(6):625-630. [Medline: [9039380](#)]
41. Maziak W, Hense HW, Döring A, Keil U. Ten-year trends in smoking behaviour among adults in southern Germany. *Int J Tuberc Lung Dis* 2002 Sep;6(9):824-830. [Medline: [12234139](#)]
42. Kouvonen A, Kivimäki M, Virtanen M, Pentti J, Vahtera J. Work stress, smoking status, and smoking intensity: an observational study of 46,190 employees. *J Epidemiol Community Health* 2005 Jan;59(1):63-69 [FREE Full text] [Medline: [15598729](#)] [doi: [10.1136/jech.2004.019752](#)]
43. Albertsen K, Hannerz H, Borg V, Burr H. The effect of work environment and heavy smoking on the social inequalities in smoking cessation. *Public Health* 2003 Nov;117(6):383-388. [Medline: [14522152](#)] [doi: [10.1016/S0033-3506\(03\)00103-3](#)]
44. El-Bastawissi A, McAfee T, Zbikowski SM, Hollis J, Stark M, Wassum K, et al. The uninsured and Medicaid Oregon tobacco user experience in a real world, phone based cessation programme. *Tob Control* 2003 Mar;12(1):45-51 [FREE Full text] [Medline: [12612361](#)] [doi: [10.1136/tc.12.1.45](#)]
45. Whelan AM, Cooke CA, Sketris IS. The impact of socioeconomic and demographic factors on the utilization of smoking cessation medications in patients hospitalized with cardiovascular disease in Nova Scotia, Canada. *J Clin Pharm Ther* 2005 Apr;30(2):165-171. [Medline: [15811170](#)] [doi: [10.1111/j.1365-2710.2004.00622.x](#)]
46. Balfour D, Benowitz NL, Fagerstrom K, Kunze M, Keil U. Diagnosis and treatment of nicotine dependence with emphasis on nicotine replacement therapy. A status report. *Eur Heart J* 2000 Mar;21(6):438-445. [Medline: [10671484](#)] [doi: [10.1053/euhj.1999.1949](#)]
47. Prochaska JO. A stage paradigm for integrating clinical and public health approaches to smoking cessation. *Addict Behav* 1996 Nov;21(6):721-732. [Medline: [8904938](#)] [doi: [10.1016/0306-4603\(96\)00031-7](#)]
48. Hymowitz N, Cummings KM, Hyland A, Lynn W, Pechacek TF, Hartwell T. Predictors of smoking cessation in a cohort of adult smokers followed for five years. *Tob Control* 1997;6(Suppl 2):S57-S62. [Medline: [9583654](#)]

49. Cokkinides V, Ward E, Jemal A, Thun M. Under-use of smoking-cessation treatments: results from the National Health Interview Survey, 2000. *Am J Prev Med* 2000 Jan;28(1):119-122. [Medline: [12817354](#)] [doi: [10.1016/j.amepre.2004.09.007](#)]
50. Grandes G, Cortada JM, Arrazola A, Laka JP. Predictors of long-term outcome of a smoking cessation programme in primary care. *Br J Gen Pract* 2003 Feb;53(487):101-107 [[FREE Full text](#)] [Medline: [12817354](#)]
51. Ferguson JA, Patten CA, Schroeder DR, Offord KP, Eberman KM, Hurt RD. Predictors of 6-month tobacco abstinence among 1224 cigarette smokers treated for nicotine dependence. *Addict Behav* 2003 Sep;28(7):1203-1218. [Medline: [12915164](#)] [doi: [10.1016/S0306-4603\(02\)00260-5](#)]
52. Couper MP, Peytchev A, Strecher VJ, Rothert K, Anderson J. Following up nonrespondents to an online weight management intervention: randomized trial comparing mail versus telephone. *J Med Internet Res* 2007 Jun 13;9(2):e16 [[FREE Full text](#)] [Medline: [17567564](#)]
53. Hall SM, Delucchi KL, Velicer WF, Kahler CW, Ranger-Moore J, Hedeker D, et al. Statistical analysis of randomized trials in tobacco treatment: longitudinal designs with dichotomous outcome. *Nicotine Tob Res* 2001 Aug;3(3):193-202. [Medline: [11506764](#)] [doi: [10.1080/14622200110050411](#)]
54. Tomson T, Björnström C, Gilljam H, Helgason A. Are non-responders in a quitline evaluation more likely to be smokers? *BMC Public Health* 2005 May 23;5(1):52 [[FREE Full text](#)] [Medline: [15910682](#)] [doi: [10.1186/1471-2458-5-52](#)]
55. An LC, Schillo BA, Kavanaugh AM, Lachter RB, Luxenberg MG, Wendling AH, et al. Increased reach and effectiveness of a statewide tobacco quitline after the addition of access to free nicotine replacement therapy. *Tob Control* 2006 Aug;15(4):286-293 [[FREE Full text](#)] [Medline: [16885577](#)] [doi: [10.1136/tc.2005.014555](#)]

Abbreviations

NRT: nicotine replacement therapy

submitted 14.03.07; peer-reviewed by R Borland, B Bock; comments to author 15.05.07; revised version received 14.08.07; accepted 14.08.07; published 30.09.07.

Please cite as:

Saul JE, Schillo BA, Evered S, Luxenberg MG, Kavanaugh A, Cobb N, An LC

Impact of a Statewide Internet-Based Tobacco Cessation Intervention

J Med Internet Res 2007;9(3):e28

URL: <http://www.jmir.org/2007/3/e28/>

doi: [10.2196/jmir.9.4.e28](#)

PMID:

© Jessie E Saul, Barbara A Schillo, Sharrilyn Evered, Michael G Luxenberg, Annette Kavanaugh, Nathan Cobb, Lawrence C An. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 30.9.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

Toward a Model for Field-Testing Patient Decision-Support Technologies: A Qualitative Field-Testing Study

Rhodri Evans¹, MRCGP; Glyn Elwyn¹, FRCGP, PhD; Adrian Edwards¹, MRCP, MRCGP, PhD; Eila Watson², PhD; Joan Austoker³, PhD; Richard Grof⁴, PhD

¹Department of Primary Care and Public Health, Cardiff University, Wales, UK

²School of Health and Social Care, Oxford Brookes University, Oxford, UK

³Cancer Research UK Primary Care Education Research Group, Division of Public Health and Primary Health Care, University of Oxford, Oxford, UK

⁴Centre for Quality of Care Research (WOK), Radboud University, Nijmegen, Netherlands

Corresponding Author:

Rhodri Evans, MRCGP

Department of Primary Care and Public Health

Cardiff University, Neuadd Meirionnydd

Heath Park, Cardiff CF14 4YS

Wales, UK

Phone: +44 2920687195

Fax: +44 2920687219

Email: evansr10@cardiff.ac.uk

Abstract

Background: Field-testing is a quality assurance criterion in the development of patient decision-support technologies (PDSTs), as identified in the consensus statement of the International Patient Decision Aids Standards Collaboration. We incorporated field-testing into the development of a Web-based, prostate-specific antigen PDST called Prosdex, which was commissioned as part of the UK Prostate Cancer Risk Management Programme.

Objectives: The aim of this study was to develop a model for the future field-testing of PDSTs, based on the field-testing of Prosdex. Our objectives were (1) to explore the reactions of men to evolving prototypes of Prosdex, (2) to assess the effect of these responses on the development process, and (3) to develop a model for field-testing PDSTs based on the responses and their effect on the development process.

Methods: Semistructured interviews were conducted with the men after they had viewed evolving prototypes of Prosdex in their homes. The men were grouped according to the prototype viewed. Men between 40 and 75 years of age were recruited from two family practices in different parts of Wales, United Kingdom. In the interviews, the men were asked for their views on Prosdex, both as a whole and in relation to specific sections such as the introduction and video clips. Comments and technical issues that arose during the viewings were noted and fed back to the developers in order to produce subsequent prototypes.

Results: A total of 27 men were interviewed, in five groups, according to the five prototypes of Prosdex that were developed. The two main themes from the interviews were the responses to the information provided in Prosdex and the responses to specific features of Prosdex. Within these themes, two of the most frequently encountered categories were detail of the information provided and balance between contrasting viewpoints. Criticisms were encountered, particularly with respect to navigation of the site. In addition, we found that participants made little use of the decision-making scale. The introduction of an interactive contents page to prototype 2 was the main change made to Prosdex as a result of the field-testing. Based on our findings, a model for the field-testing of PDSTs was developed, involving an exploratory field-testing stage between the planning stage and the development of the first prototype, and followed by the prototype field-testing stage, leading to the final PDST.

Conclusions: In the field-testing of Prosdex, a Web-based prostate-specific antigen PDST, the responses of interviewed men were generally favorable. As a consequence of the responses, an interactive contents page was added to the site. We developed a model for the future field-testing of PDSTs, involving two stages: exploratory field-testing and prototype field-testing.

(*J Med Internet Res* 2007;9(3):e21) doi:[10.2196/jmir.9.3.e21](https://doi.org/10.2196/jmir.9.3.e21)

KEYWORDS

Field-testing; patient decision-support technologies; prostate-specific antigen (PSA); prostatic neoplasms; informed choice; decision support techniques; patient education; patient participation; consumer health informatics; Internet

Introduction

Field-testing is increasingly recognized as an important step in the quality assurance of patient decision-support technologies (PDSTs), interventions commonly known as decision aids. This was underlined by the International Patient Decision Aids Standards (IPDAS) Collaboration consensus statement on PDST quality, the product of a Delphi process involving all major stakeholder groups, at the end of which nine domains of PDST quality criteria were agreed upon [1]. One of these domains was systematic developmental process, which incorporated the criterion of field-testing in order to show that a decision aid was acceptable to patients [1]. IPDAS, however, did not define field-testing, and, more broadly, the PDST/decision aid literature gives very little guidance in this respect [2]. Furthermore, there are potentially two processes encapsulated in field-testing: (1) the development of a prototype with users, and (2) the “live” testing of a refined prototype.

In 2002, we were commissioned to develop a Web-based, prostate-specific antigen (PSA) PDST, called Prosdex, and included field-testing as part of the development process [3]. Prosdex formed part of the UK Prostate Cancer Risk Management Programme strategy, led by the National Cancer Screening Programmes, which had, at its heart, the concept of informed choice in PSA testing [4]. According to the strategy, UK men interested in the PSA test would be provided with information to enable them, with their family doctor, to make an informed decision. Prosdex was developed in order to present this information in the format of a Web-based, multimedia, interactive PDST. This opportunity for users to explore the information presented on PSA explains the full name of Prosdex: Prostate-Specific Antigen Decision Explorer [3].

Prosdex presents evidence-based information about prostate cancer and PSA testing, encouraging users to weigh the pros and cons of testing for themselves. Much of the information came from an earlier, paper-based decision aid for PSA testing commissioned as part of the UK Prostate Cancer Risk

Management Programme and approved by its Scientific Reference Group [4]. Of particular importance in that decision aid were the statistical/epidemiological data, which allowed us to present some of the more controversial issues, such as the validity of the PSA test. For instance, we stated in Prosdex that two thirds of men with a raised PSA test do not have prostate cancer. The development of Prosdex was also underpinned by a systematic review of PSA decision aids, undertaken not only to garner information on extant PDSTs, but also to explore their effects. We found that the evaluations of PSA decision aids demonstrated, fairly consistently, an improvement in knowledge about PSA testing and prostate cancer; in contrast, however, no clear effect was found on PSA testing itself [5]. The findings were broadly similar to those of a Cochrane review of the effect of PDSTs that considered a range of clinical domains [6]. This review found that patients who use PDSTs participate more, know more, have more realistic expectations of benefits and harms, and are more likely to receive an option with outcomes they most value [6,7].

Narrative is also employed in Prosdex to present information. Specifically, there are 25 video clips of enacted patient experiences about the PSA test and subsequent investigations/treatments. The transcripts for these clips were obtained from a qualitative study of men’s experiences of PSA testing [8]. Informed choice is actively encouraged in Prosdex through structured decision support in the form of a decision-making scale. The link to this functionality lies on the top right of each page, thereby allowing users to weigh the impact of the information in that particular page on their decision-making process. Specifically, they are able to indicate whether they are for, against, or undecided about PSA testing on the basis of that information. Each decision is then added, or “stacked,” in the decision summary to produce a cumulative result for the pages viewed. Prosdex has been designed to cater to the needs of users with visual and hearing difficulties. Consequently, there is a voice-over option to which the website defaults, but which can be switched off; there are also subtitles for the video clips.

Figure 1a. Prosdex screenshots

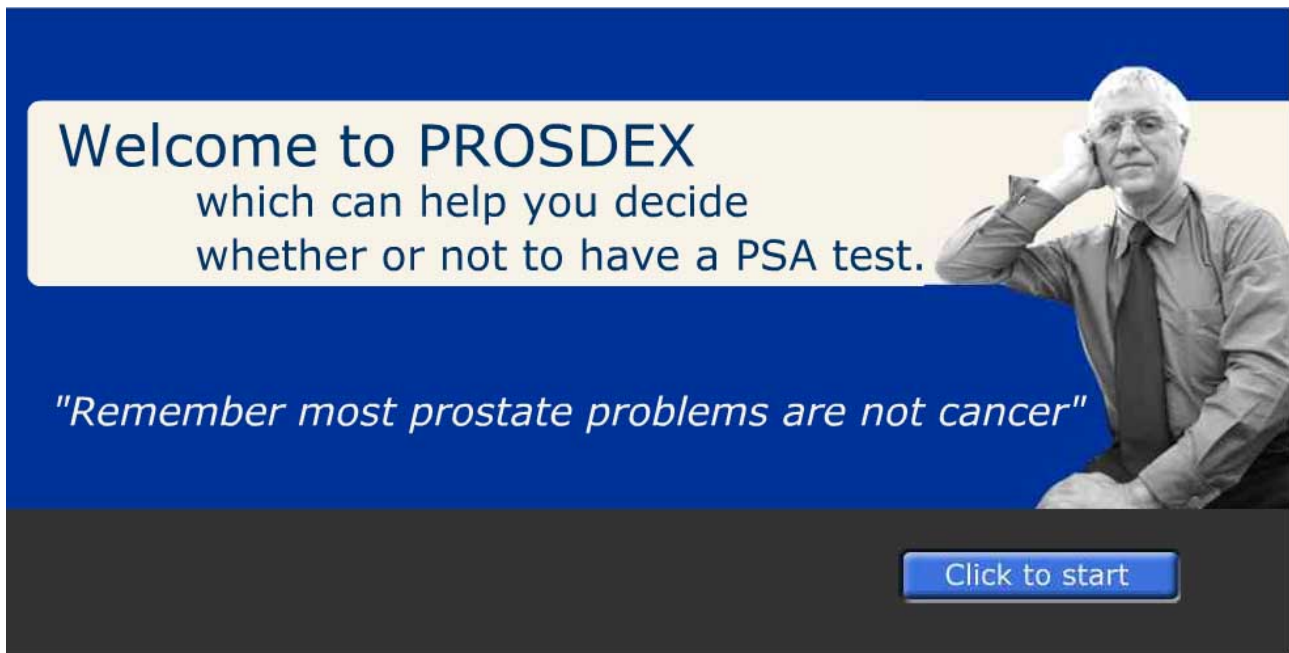
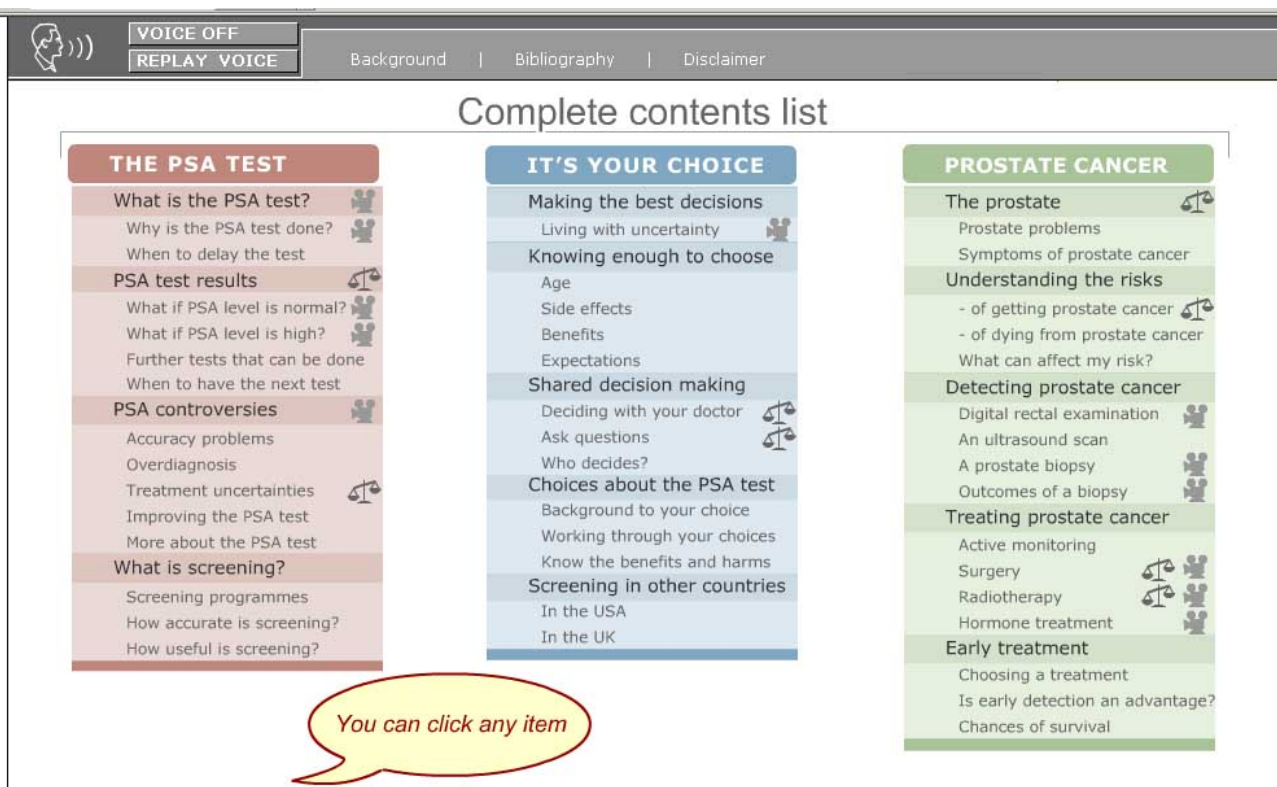


Figure 1b. Prosdex screenshots



In this study, we attempted to capture the process of field-testing Prosdex by interviewing men who used it. Furthermore, by introducing evolving versions of Prosdex, we hoped that the series of interviews would help us, as developers, to identify strengths and weaknesses and modify the prototype. Beyond this, however, we wanted to explore the role of field-testing in the development of PDSTs. Specifically, our aim in this study was to develop a model for future field-testing of PDSTs. Our objectives were: (1) to explore the responses of men to evolving prototypes of Prosdex, (2) to assess the effect of these responses

on the development process, and (3) to develop a model for field-testing PDSTs based on the responses and their effect on the development process.

Methods

A qualitative study design was employed using semistructured interviews and incremental prototypes of Prosdex. Men between the ages of 40 and 75 were recruited, the target age range for the UK Prostate Cancer Risk Management Programme. The

men were recruited from two family doctor practices in Wales, United Kingdom: one in a coastal/rural town and the other in a postindustrial town. The men had previously participated in a randomized controlled trial examining the effect of a brief patient decision aid—a written one-page leaflet given to the intervention group. All of the men in that trial completed a written questionnaire exploring their knowledge of and attitudes toward PSA testing and their intention to be tested [9]. At the end of the questionnaire, the men were asked to indicate whether they would be interested in participating in the qualitative study, and those who did so were sent invitation letters, information sheets, and consent forms by the research team.

Men who agreed to participate were contacted, and arrangements were made for them to view a stand-alone prototype of Prosdex on a laptop, in their homes, in the presence of one of the

researchers (RE), who sat behind them. The researcher did not impart any advice or answer questions on content matters during the viewing—questions were, however, addressed during the subsequent interview. Technical questions, particularly those relating to difficulties in using Prosdex (eg, navigational problems), were answered contemporaneously. In the event of the men being unable to use a computer mouse, the researcher performed this function for them, opening specified Web pages but not giving any direction on use. The men were asked to indicate when they had finished using Prosdex and were then given a 5- to 10-minute break before the interview.

In the interview, the men were initially asked for their opinion of Prosdex in general. Then, they were asked for their views on specific aspects of Prosdex, some of which are listed in [Table 1](#).

Table 1. Specific aspects of Prosdex discussed in the interview

	Aspect of Prosdex
1	Front page
2	Voice-over
3	Ease of use of the left-hand heading section
4	Section headings: clear or confusing
5	Decision scale and decision summary page
6	Video clips
7	Information in Prosdex: Presentation, and ease of use, in center panel Detail of information Legibility of information Explanation of information Format: text, video clips Satisfaction with the information Views on the presentation of different outcomes
8	Relative preference for leaflet or Prosdex
9	Suggestions for making Prosdex easier to use
10	Time taken to use Prosdex
11	Aspects that were most/least helpful

The transcribed interviews were coded independently by RE and GE with qualitative software, Atlas-ti (version 4.1), and using the technique of constant comparison [10,11]. The coded transcripts were then subjected to thematic analysis by RE and GE. Technical issues that arose during the viewings were noted by RE, and those, in addition to comments from the interviews, were fed back to the multimedia designer. Feedback occurred after a group of men had viewed each prototype, in order to maintain version control. This iterative development process resulted in evolving prototypes of Prosdex. The content of the site, however, stayed the same throughout. Finally, after analyzing the men's responses and subsequent changes to Prosdex, we developed a model for field-testing PDSTs.

Results

The results are presented in five sections: (1) characteristics of the interviewed men, (2) data from themes, (3) analysis of data, (4) outline of changes made to Prosdex, (5) a model for field-testing PDSTs.

Characteristics of the Interviewed Men

A total of 27 men were interviewed after using Prosdex, between September 2004 and February 2005, and they were grouped according to the prototype viewed. There were five groups; the group that used prototype 1 (7 men) was deliberately larger than the others in order to capture the majority of the technical problems before the production of further prototypes. The men viewed Prosdex for between 15 and 45 minutes.

Table 2. Characteristics of the interviewed men

Prosdx Prototype	Date of Viewing	Number of Men	Age Range	Median Age	Number Who Previously Had PSA Test	Number Unable to Use Computer Mouse
1	September 2004	7	49-70	55	0	2
2	October 2004	6	50-76	60	1	1
3	November 2004	5	42-70	60	3	0
4	January 2005	5	50-70	68	1	0
5	February 2005	4	43-68	58	1	1
Total		27	42-76		6	4

Data From the Themes

Two main themes were identified, and they are shown in [Table 3](#) and [Table 4](#), along with their categories and illustrative quotations. The respondent/man is identified according to the

prototype group; for instance, the third man to use and be interviewed about prototype 4 is P4,i3. It should be noted, however, that the quotations are presented in relation to the themes for the whole sample, not in relation to the developing prototypes.

Table 3. Theme 1: Responses to the Information Provided in Prosdex - Categories and quotations

Category	Quotations
detail	<p>(Q) Do you think there was enough information, or too much?</p> <p>P1,i2: <i>I didn't think there was too much. I think on the question on symptoms, I don't know whether it was possible to give any more information, because once you start giving instances or factors, I suppose it's impossible to be exhaustive in any case, and therefore you can only give a broad brush.</i></p> <p>P1,i3: <i>It made things a lot clearer, but I am slow on the uptake anyway. It takes a long time for things to sink in at the moment.</i></p> <p>P1,i4: <i>I was taking it all in, so I think there was enough to be honest. Maybe in time I will think about something, and I should have asked this or that; it's like everything else. I think there was enough for the first time to be honest.</i></p> <p>P1,i7: <i>Very informative. It raises some points which obviously concern you. The sexual activity aspect. It's very comprehensive; it spells everything out for you.</i></p> <p>P4,i2: <i>I would like to know more but I'm not sure, after having looked at the website, whether the information is actually in the public domain anyway. With the test being as inaccurate as it is.</i></p> <p>(Q) Did the information go into enough detail for you?</p> <p>P5,i1: <i>Oh certainly enough detail. There was definitely enough to make a decision.</i></p>
balance	<p>P1,i6: <i>It gave you the pluses and the minuses quite well.</i></p> <p>P3,i4: <i>I thought it was very informative; it told you the advantages and disadvantages, and the percentage of possibility of having the problem with the prostate, and not be detected, which you really don't want to hear that. What you want is a positive answer all the time, but obviously in life you can't have that.</i></p> <p>P4,i1: <i>There perhaps ought to be more emphasis on the fact that benign prostatic hypertrophy was a perfectly normal characteristic of an aging male population, but on the other hand there is the possibility that it might be either an aggressive or an unaggressive nature, and that initially people don't need to go any further than that.</i></p>
suggestions to improve the information	<p>P1,i2: <i>The one point that I did think could be improved was where it said, "What is the practice with regard to PSA in other countries?" And it only mentioned America, and ideally, I think it should compare to other European countries.</i></p> <p>P1,i3: <i>As I say, being a layman, not a lot of people know where the prostate is and all that. There could be a little bit more then.</i></p> <p>(Q) Was there any information that you would have liked, but you couldn't find?</p> <p>P1,i5: <i>The diagram was quite informative. I would have liked more detail. I would have liked more pictures as well.</i></p> <p>(Q) Was there any other information that was not there?</p> <p>P1,i7: <i>Possibly some statistics on tests that have been done, particularly as they used a comparison. They showed something about comparing the frequency of when these tests are carried out, say like in the USA, and they used a similar screening program for breast cancer for women. I think it would be interesting to see what sort of statistics have been gained.</i></p> <p>P2,i6: <i>But they didn't say if you're 75-85 how it would be likely to affect you or not affect, a purely selfish point of view. Having reached 75 now I want to know what are the prospects for me over the next ten years.</i></p> <p>P4,i3: <i>That phrase, "up to 1 in 5," that phrase doesn't mean anything. And I really think that that shouldn't be there.</i></p> <p>(Q) Are there any other types of information that you would like to see?</p> <p>P5,i3: <i>A bibliography would have been useful, if there were references to more detailed information.</i></p>

Table 4. Theme 2: Responses to Specific Features of Prosdex - Categories and Quotations

Feature	Category	Quotation
navigation		<p>P1,i6: <i>I think it could have been slightly better. I think it would have been better if you had gone directly from one section to the other, if you are guided better from one section to the other in a better way. I think it was going backwards and forwards all over the place. It could be a little bit confusing, and you could actually forget or miss bits.</i></p> <p>(Q) How long did it take you to feel comfortable using it?</p> <p>P2,i2: <i>Minutes. As soon as I worked out that you could take it in any order you wished. But I was quite happy to follow along with the program.</i></p> <p>P3,i3: <i>I personally found it very easy, but I would think maybe someone quite a bit older who didn't have computer skills probably would be a little overwhelmed.</i></p> <p>P4,i2: <i>Very easy to navigate round, and I understood it, so I would think 90% of the population could understand it.</i></p>
video clips	balance	<p>(Q) Do you think there were enough video clips, or too many?</p> <p>P1,i2: <i>I don't think there were too many. There were two videos where they were referring to similar symptoms about radiotherapy and diarrhea. No, I thought the balance was right.</i></p> <p>P1,i5: <i>I thought they were very good actually. Some of them were a bit disheartening, but it depends on people's pain level. I mean I have got quite a high pain level.</i></p>
	detail	<p>P2,i6: <i>They looked a little bit staged, like actors saying the words...just a little bit too rehearsed. And very brief, the comments were very brief. Could you condense those down into less choices but longer explanations?</i></p> <p>P5,i2: <i>Well, it's enough detail to talk about it, but would be better detail to actually see it. It would give you a better idea of what you've got to do and what you've got to go through. Like the operation.</i></p>
voice-over	clarity	<p>P1,i7: <i>Very clear, and an easy pace to listen to as well. It neither went too fast nor too slow.</i></p> <p>P2,i1: <i>I think the thing was, that you didn't know whether to listen to the voice or read the words, and then go back and hear the voice again. I wasn't sure about that. If I had to go through the program another time, I would get to know my way around it better, let's put it that way.</i></p> <p>P2,i5: <i>I found I was starting to read over it then waiting for the voice to catch up.</i></p> <p>P3,i3: <i>It saves my eyesight and it also slows me down. I would probably, if I was purely reading it, I would probably speed read it and skip quite a lot more. So I found the voice very, very helpful.</i></p>
decision-making scale	limited utility	<p>P1,i5: <i>On about four or five things, but generally I got too engrossed in the bit on the left reading through it all, and listening to it as well.</i></p> <p>(Q) Did you use that?</p> <p>P1,i6: <i>I didn't actually, because I was going through the rest of the info, so I didn't bother. Maybe I should have done, I'm sorry.</i></p> <p>P1,i6: <i>I think it may have been easier if the decision scale was at the bottom, underneath the section you are reading, as opposed to a little box on the top right. So as you go through it, click it, then go to the next page, click it on the bottom.</i></p> <p>P2,i1: <i>I feel that before you moved on, if there was some sort of audio or visual prompt so that if you haven't clicked on the decision box it prompts you - maybe a little pop up or a bleep or something to tell you that you hadn't ticked the decision box.</i></p> <p>P2,i2: <i>There wasn't any indication of where to use it. Whether you just had to use it at the end or at every page you'd read. I wasn't sure what to do.</i></p> <p>P2,i3: <i>I didn't actually go to that, because I knew what I'd put in.</i></p> <p>P4,i2: <i>I did it in my head. I'm used to making decisions, so I don't need a little Geiger counter to tell me.</i></p>

Analysis of Data

Theme 1: Responses to the Information Provided in Prosdex

Three main categories were identified: (1) detail, (2) balance, (3) suggestions to improve the information.

Detail

In general, the men were happy with the amount of information provided (P1,i4), although there was an appreciation of the difficulty in deciding on the level of detail (P1,i2) and a realization of the weakness of the evidence base (P4,i2). Openness on sensitive issues was commended (P1,i7), and there was some evidence that the site helped to clarify some of the

complexities and uncertainties of PSA testing (P1,i3). This level of detail was noted, in some cases, to be helpful for the decision-making process (P5,i1).

Balance

Mostly positive comments were made about the balance of the information on the site. The presentation of uncertainty was commended (P1,i6), and there was an appreciation of the difficulties involved in presenting such information (P3,i4). Nonetheless, there were some dissenting comments in this respect; for instance, one man would have preferred a greater emphasis on the benign nature of most prostate conditions (P4,i1).

Suggestions to Improve the Information

Specific suggestions were made to improve on the information on the site. These included a desire for more background anatomical information (P1,i3), more diagrams (P1,i5), more age-specific information (P2,i6), and a preference for a bibliography (P5,i3). In addition, there were comments on the lack of information about other European countries (P1,i2), and some criticism of the presentation of the statistical information.

Theme 2: Responses to Specific Features of Prosdex

The four specific features that were discussed in greatest detail were (1) navigation of the site, (2) video clips, (3) voice-over, (4) decision-making scale.

Navigation of the Site

The navigation difficulties with prototype 1 resulted in the most significant criticism of Prosdex (P1,i6). Men using later prototypes were less critical of the navigation, almost certainly due to the interactive contents page developed after field-testing of prototype 1 (P2,i2).

Video Clips

The two main categories identified here were those of balance and detail. In terms of balance, the responses were positive (eg, regarding our presentation of contrasting opinions and experiences) (P1,i2). There was also an appreciation of the difficulty in striking such a balance, particularly when dealing with sensitive issues (P1,i5).

With respect to detail, there were two specific criticisms. One man expressed a desire for more graphic detail in relation to the descriptions of prostate investigations and treatments (P5,i2). Another man expressed a preference for less choice of video clips for a particular issue, and for greater detail in those clips (P2,i6).

Voice-Over

The category of note here was clarity, and, in this respect, the responses were mixed.

Only one man (P2,i4) decided to switch off the voice-over using the button provided. Of those who left the voice-over on, some gave positive responses (P1,i7); in particular, one man found the process of reading to be made easier (P3,i3) with the voice-over. In terms of negative responses, one man found the voice confusing (P2,i1), and another found the voice-over to restrict his use of Prosdex (P2,i5).

Decision-Making Scale

The significant category here was limited utility, a consequence of the men making little use of the decision-making scale (P1,i6; P2,i3; P4,i2). The reasons given for the minimal use of the scale varied. For one man it related to the positioning of the scale on the screen (P1,i6); for another, it seemed to be caused by a limited understanding of when to use the scale (P2,i2). As the viewing of Prosdex progressed, one man focused on the content and stopped using the scale (P1,i5). One solution offered for this was audiovisual prompts/reminders to use the scale (P2,i1).

Outline of Changes Made to Prosdex During the Field-Testing

Navigation

The major change made to Prosdex during the course of field-testing was to improve the navigation of the site. As previously noted, men using prototype 1 found it difficult to keep a record of which pages they had viewed (P1,i6). Consequently, for prototype 2, an interactive contents page was developed that not only indicated to the men which pages they had visited, but also allowed them to navigate directly to sections of interest. This change improved the navigation significantly for the men, and no other amendments were deemed necessary in this respect.

Content

The content, both text and video, remained unchanged in Prosdex since the responses regarding this were generally positive, in particular about the detail and balance of the information on the site. As highlighted above, there were some specific suggestions, and these were considered by the developers. It was decided, however, that either the information was, in fact, already present in the site, or that the requested content would have overwhelmed sections that were already very detailed. An example of this was the request for a pan-European comparison of PSA screening (P1,i2). Our decision to keep the comparison at a UK/USA level was made in order to provide the UK target audience with a relevant comparison of different practices.

Voice-Over

Despite the mixed responses to the voice-over functionality, it was retained in Prosdex. As previously noted, only one of the respondents (P2,i4) asked for the voice-over to be switched off, and only two of the respondents (P2,i1 and P2,i5) stated that the voice-over affected their reading of the text. Furthermore, as developers of a publicly available health information site, we were obliged to make arrangements for visually-impaired users or those with reading difficulties. Finally, we were confident that the criticisms raised could be addressed by the clearly marked option on the site to switch off the voice-over.

Decision-Making Scale

The decision-making scale was also retained in Prosdex despite its low usage in the field-testing. Our reason for doing so was based on the original design for Prosdex, one of the key features of which was a tool for interactive decision making. There was also no evidence that the scale interfered with other components of the site, and it was agreed that some users might find it to be of benefit.

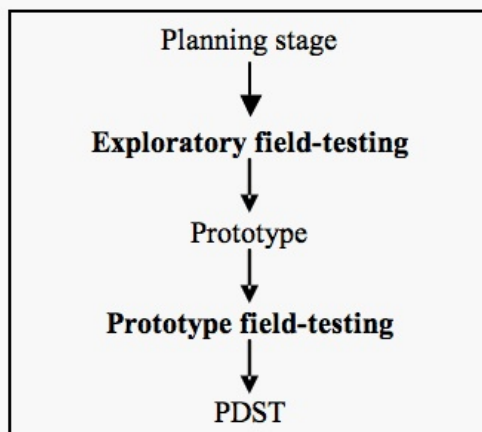
A Model for Field-Testing PDSTs

For the purposes of developing a model for field-testing PDSTs, we reflected on the qualitative data from the men's responses and on the changes made to Prosdex. We found that PDST field-testing was composed of two distinct processes: (1) a process of user involvement in the development of the PDST, and (2) user trials of one or more prototypes. Consequently, for the model, we divided field-testing into two stages (Figure 2). In the first stage, which we defined as exploratory field-testing,

users would be asked to look at specific components of the PDST early in its development, before the construction of the first prototype, thus allowing users to influence key decisions early on. In the second stage, which we defined as prototype

field-testing, users would be shown successive prototypes, as in this study, but with reference to changes made during the development process.

Figure 2. Proposed model for field-testing PDSTs



Discussion

Summary of Main Findings

The two main themes from the interviews were the responses to the information provided in Prosdex and the responses to specific features of Prosdex. Within these themes, two of the most frequently encountered categories were detail of the information provided and balance between contrasting viewpoints. Criticisms were, however, encountered, particularly with respect to the navigation of the site. In addition, we found that the men made little use of the decision-making scale.

The introduction of an interactive contents page to prototype 2 was the main change made to Prosdex as a result of the field-testing. Other aspects of the site, notably the content, voice-over, and decision-making scale, were not changed, for two reasons. First, the collective responses did not justify radical amendments such as removing specific features. Second, there were factors other than the men's responses to consider in the development process, notably PDST quality criteria. For example, one of the reasons for retaining the decision-making scale was that values clarification is an internationally recognized quality criterion for PDSTs [1].

Finally, based on our findings, a model for the field-testing of PDSTs was developed, comprising two stages: exploratory field-testing and prototype field-testing.

Limitations of the Study

Only two family practices were used to recruit men for this study. It would have been desirable to recruit men from a greater number of practices to ensure a broader socioeconomic and geographic population distribution. Another limitation was the fact that the men had previously participated in a randomized controlled trial of a brief PSA paper decision aid in which they all had completed a written questionnaire. However, we would argue that this study differed significantly in that it focused on

the details and technical aspects of a specific PDST, Prosdex, which was not featured in the trial. A qualitative methodology, semistructured interviews, was employed in this study. Arguably, however, the study design was descriptive, using qualitative techniques and employing a relatively technical, specific interview schedule, which, to an extent, is in accordance with the model stage of the complex intervention framework, as developed by Campbell et al [12].

The validity of the study's findings is potentially open to criticism as no formal measures were employed in this respect. For instance, there was no triangulation, using data from other methods such as surveys [11]. Such an approach would, however, have been impractical in our opinion due to the dependency in this study on the presence of a researcher to facilitate the viewing of the PDST. Moreover, we would contend that the observational data from these viewings provided a degree of corroboration. For instance, the comments from group 1 on the navigational difficulties accorded with the researcher's observations. Finally, respondent validation was not used as the men's responses were dependent on their immediate recollections and views of Prosdex [11]. Corroborating these responses with the results at a later date would not, in our opinion, have been a reliable method.

Comparisons With Previous Work

As noted earlier, we previously undertook a systematic review of evaluations of PDSTs on the topic of PSA [5]. In contrast, there are, to date, no studies that specifically consider the field-testing of PSA PDSTs. There are, however, such studies in other clinical areas, although most of these focus on the usability and acceptability of prototype PDSTs, corresponding to the prototype field-testing stage of our proposed model. For example, Irwin et al found in a pilot study that a decision aid for women with breast cancer was described as helpful by most of the users [13]. Feldman-Stewart and colleagues field-tested a PDST designed for men with early stage prostate cancer with

a group of “surrogate patients”—men, without prostate cancer, of the same age as the target group of the PDST [14]. It was observed that the men were able to understand the information provided and that most were able to express treatment choices. In a noncancer setting, Lalonde et al found high levels of acceptability for a PDST aimed at improving the knowledge of patients with hypertension/hyperlipidemia [15]. Finally, and significantly, in the context of a multimedia Web-based PDST such as Prosdex, Diefenbach and Butz field-tested a multimedia interactive education system for prostate cancer patients and found high levels of acceptability [16].

The importance of prototype field-testing was highlighted by O’Donnell and colleagues in a review of the implementation of patient decision aids in clinical practice [7]. One of the significant barriers for implementation was described as “usability for diverse patients.” Specifically, the authors noted the lack of evidence on the assessment of the readability of PDSTs—a weakness shared by Prosdex—though they welcomed the finding, in the Cochrane review inventory, that most PDSTs were developed for general audiences (eg, grade 8 reading level) [6]. O’Donnell et al suggested further research on how PDSTs could improve the decision quality for people who vary by demographic characteristics. This is an important statement as it extends the potential scope of prototype field-testing. Moreover, there is a strong argument that our proposed second stage of field-testing only becomes valid if it has taken into consideration the diversity of the target audience.

There is an even greater research deficit for the exploratory field-testing of PDSTs. In one of the few studies available, Sawka et al described the development of a decision aid for choice of surgical treatment for breast cancer [17]. Notably, the study involved a needs assessment stage during which focus groups were held involving women with a previous diagnosis of breast cancer, and which considered issues such as information the women wished they had received at diagnosis. Subsequently, the decision aid was developed in conjunction with a steering group that revised various drafts of the aid. Finally, in a pilot study, almost all of the women responded positively to the decision aid. This twin approach of needs assessment and pilot study forms a strong basis for the development of a decision aid and, moreover, corresponds, in our opinion, to our proposed two-stage model for field-testing PDSTs.

The paucity of research into field-testing has implications for developers of PDST quality criteria. As previously mentioned, field-testing is, at present, regarded as an important criterion in the IPDAS framework. Moreover, this framework gives direction for the development process of PDSTs. Arguably, components of that development process are very similar to the two stages of field-testing that we propose. This is particularly true of the exploratory stage, and it again raises the question of the definition of field-testing. What is certain, however, is that with such little understanding of this criterion, it is difficult to contend, at present, that firm assessments can be made against it [1].

Implications for Clinical Practice and Future Research

Our proposed two-stage model and, in particular, the exploratory field-testing stage, raises a number of challenges for developers of PDSTs, not least of which is the difficulty of accommodating it within the pressures of deadlines and budgets. There is also the challenge of balancing the opinions of users with those of experts/scientific reference groups, particularly in situations of clinical uncertainty/equipose. Arguably, the model is too simplistic in that it presupposes a linear progression from exploratory to prototype field-testing. In reality, more complex PDSTs might follow a different development path wherein the factual content, for instance, would require both exploratory and prototype field-testing in order to develop other features of the PDST, for example, videos of patient experiences. Moreover, the model does not take into account contextual factors, such as the influence of family/friends and health professionals, which could have an impact on the utilization of PDSTs in a natural setting. Nevertheless, the principle of two-stage field-testing for PDSTs, whether applied in parts or as a whole, still holds true in our opinion; we suggest further research to test it and other future models of field-testing. In doing so, it is hoped that reviewers of PDSTs, and international standard groups such as IPDAS, will have at their disposal a clearer definition of field-testing.

Conclusions

In the field-testing of Prosdex, a Web-based PSA PDST, the responses of interviewed men were generally favorable. As a consequence of the responses, an interactive contents page was added to the site. We developed a model for the future field-testing of PDSTs involving two stages: exploratory field-testing and prototype field-testing.

Acknowledgments

We are grateful to all the men who participated in this study, and to the two family practices involved. We also wish to acknowledge the help of Mr. Benjamin Elwyn for transcribing the interviews. The study was supported financially by Cancer Research UK.

Conflicts of Interest

None declared.

References

1. Elwyn G, O’Connor A, Stacey D, Volk R, Edwards A, Coulter A, et al; International Patient Decision Aids Standards (IPDAS) Collaboration. Developing a quality criteria framework for patient decision aids: online international Delphi

- consensus process. *BMJ* 2006 Aug 26;333(7565):417 [FREE Full text] [Medline: [16908462](#)] [doi: [10.1136/bmj.38926.629329.AE](#)]
2. Kim YM, Kols A, Martin A, Silva D, Rinehart W, Prammawat S, et al. Promoting informed choice: evaluating a decision-making tool for family planning clients and providers in Mexico. *Int Fam Plan Perspect* 2005 Dec;31(4):162-171 [FREE Full text] [Medline: [16439343](#)]
 3. PROSDEx Home page. URL: <http://www.prosdex.com/> [accessed 2007 Jun 11] [WebCite Cache ID 5MzK3MmBO]
 4. UK National Screening Committee. Prostate Cancer Risk Management Programme. 2001. .
 5. Evans R, Edwards A, Brett J, Bradburn M, Watson E, Austoker J, et al. Reduction in uptake of PSA tests following decision aids: systematic review of current aids and their evaluations. *Patient Educ Couns* 2005 Jul;58(1):13-26. [Medline: [15950832](#)] [doi: [10.1016/j.pec.2004.06.009](#)]
 6. O'Connor AM, Stacey D, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2001(3):CD001431 [Update: *Cochrane Database Syst Rev* 2003;(2):CD001431]. [Medline: [11686990](#)]
 7. O'Donnell S, Cranney A, Jacobsen MJ, Graham ID, O'Connor AM, Tugwell P. Understanding and overcoming the barriers of implementing patient decision aids in clinical practice. *J Eval Clin Pract* 2006 Apr;12(2):174-181. [Medline: [16579826](#)] [doi: [10.1111/j.1365-2753.2006.00613.x](#)]
 8. Evans R, Edwards AGK, Elwyn G, Watson E, Grol R, Brett J, et al. "It's a maybe test": men's experiences of prostate specific antigen testing in primary care. *Br J Gen Pract* 2007 Apr;57(537):303-310. [Medline: [17394734](#)]
 9. Watson E, Hewitson P, Brett J, Bukach C, Evans R, Edwards A, et al. Informed decision making and prostate specific antigen (PSA) testing for prostate cancer: a randomised controlled trial exploring the impact of a brief patient decision aid on men's knowledge, attitudes and intention to be tested. *Patient Educ Couns* 2006 Nov;63(3):367-379. [Medline: [16875796](#)] [doi: [10.1016/j.pec.2006.05.005](#)]
 10. Muhr T. ATLAS.ti 4.1 - Short User's Manual. Berlin, Germany: Scientific Software Development; 1997.
 11. Mays N, Pope C. Qualitative research in health care. Assessing quality in qualitative research. *BMJ* 2000 Jan 1;320(7226):50-52 [FREE Full text] [Medline: [10617534](#)] [doi: [10.1136/bmj.320.7226.50](#)]
 12. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. *BMJ* 2007 Mar 3;334(7591):455-459. [Medline: [17332585](#)] [doi: [10.1136/bmj.39108.379965.BE](#)]
 13. Irwin E, Arnold A, Whelan TJ, Reyno LM, Cranton P. Offering a choice between two adjuvant chemotherapy regimens: a pilot study to develop a decision aid for women with breast cancer. *Patient Educ Couns* 1999 Jul;37(3):283-291. [Medline: [14528554](#)] [doi: [10.1016/S0738-3991\(98\)00117-7](#)]
 14. Feldman-Stewart D, Brundage MD, Van Manen L. A decision aid for men with early stage prostate cancer: theoretical basis and a test by surrogate patients. *Health Expect* 2001 Dec;4(4):221-234. [Medline: [11703496](#)] [doi: [10.1046/j.1369-6513.2001.00139.x](#)]
 15. Lalonde L, O'Connor AM, Drake E, Duguay P, Lowensteyn I, Grover SA. Development and preliminary testing of a patient decision aid to assist pharmaceutical care in the prevention of cardiovascular disease. *Pharmacotherapy* 2004 Jul;24(7):909-922. [Medline: [15303454](#)] [doi: [10.1592/phco.24.9.909.36104](#)]
 16. Diefenbach MA, Butz BP. A multimedia interactive education system for prostate cancer patients: development and preliminary evaluation. *J Med Internet Res* 2004 Jan 21;6(1):e3 [FREE Full text] [Medline: [15111269](#)] [doi: [10.2196/jmir.6.1.e3](#)]
 17. Sawka CA, Goel V, Mahut CA, Taylor GA, Thiel EC, O'Connor AM, et al. Development of a patient decision aid for choice of surgical treatment for breast cancer. *Health Expect* 1998 Jun;1(1):23-36. [Medline: [11281859](#)] [doi: [10.1046/j.1369-6513.1998.00003.x](#)]

Abbreviations

PSA: prostate-specific antigen

PDST: patient decision-support technology

IPDAS: International Patient Decision Aids Standards Collaboration

submitted 28.02.07; peer-reviewed by D Stacey, J Davison; comments to author 11.04.07; revised version received 28.05.07; accepted 28.05.07; published 13.07.07.

Please cite as:

Evans R, Elwyn G, Edwards A, Watson E, Austoker J, Grol R

Toward a Model for Field-Testing Patient Decision-Support Technologies: A Qualitative Field-Testing Study

J Med Internet Res 2007;9(3):e21

URL: <http://www.jmir.org/2007/3/e21/>

doi: [10.2196/jmir.9.3.e21](#)

PMID: [17627930](#)

© Rhodri Evans, Glyn Elwyn, Adrian Edwards, Eila Watson, Joan Austoker, Richard Grol. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 13.07.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

Stakeholder Perspectives on the Development of a Virtual Clinic for Diabetes Care: Qualitative Study

Natalie Armstrong, PhD; Hilary Hearnshaw, PhD, CMath; John Powell, MB, PhD, MRCPsych, FFPH; Jeremy Dale, MB, PhD, FRCGP

1

Corresponding Author:

Natalie Armstrong, PhD
Health Sciences Research Institute
Warwick Medical School
University of Warwick
Coventry CV4 7AL
United Kingdom
Phone: +44 24 765 74505
Fax: +44 24 765 28375
Email: Natalie.Armstrong@warwick.ac.uk

Abstract

Background: The development of the Internet has created new opportunities for health care provision, including its use as a tool to aid the self-management of chronic conditions. We studied stakeholder reactions to an Internet-based “virtual clinic,” which would allow people with diabetes to communicate with their health care providers, find information about their condition, and share information and support with other users.

Objective: The aim of the study was to present the results of a detailed consultation with a variety of stakeholder groups in order to identify what they regard as the desirable, important, and feasible characteristics of an Internet-based intervention to aid diabetes self-management.

Methods: Three focus groups were conducted with 12 people with type 1 diabetes who used insulin pumps. Participants were recruited through a local diabetes clinic. One-on-one interviews were conducted with 5 health care professionals from the same clinic (2 doctors, 2 nurses, 1 dietitian) and with 1 representative of an insulin pump company. We gathered patient consensus via email on the important and useful features of Internet-based systems used for other chronic conditions (asthma, epilepsy, myalgic encephalopathy, mental health problems). A workshop to gather expert consensus on the use of information technology to improve the care of young people with diabetes was organized.

Results: Stakeholder groups identified the following important characteristics of an Internet-based virtual clinic: being grounded on personal needs rather than only providing general information; having the facility to communicate with, and learn from, peers; providing information on the latest developments and news in diabetes; being quick and easy to use. This paper discusses these characteristics in light of a review of the relevant literature. The development of a virtual clinic for diabetes that embodies these principles, and that is based on self-efficacy theory, is described.

Conclusions: Involvement of stakeholders is vital early in the development of a complex intervention. Stakeholders have clear and relevant views on what a virtual clinic system should provide, and these views can be captured and synthesized with relative ease. This work has led to the design of a system that is able to meet user needs and is currently being evaluated in a pilot study.

(*J Med Internet Res* 2007;9(3):e23) doi:[10.2196/jmir.9.3.e23](https://doi.org/10.2196/jmir.9.3.e23)

KEYWORDS

Virtual clinic; diabetes; Internet; stakeholder consultation; consumer health informatics; focus groups

Introduction

The development of the Internet has created new opportunities for health care provision, including its use as a tool to aid the self-management of chronic conditions. In the United Kingdom,

there are over 2 million people diagnosed with diabetes, and Internet-based interventions may represent a way of helping them self-manage their condition from home. There have been several recent studies of Internet-based interventions with patients carried out in Europe [1-4] and North America [5-8].

A report of a small study that pilot-tested the feasibility of allowing patients with type 2 diabetes to co-manage their condition from home [9] showed proof of concept. The virtual clinic concept studied here is such a system, aimed at people receiving care for diabetes from the UK National Health Service (NHS), and it is enhanced by being based on the behavioral theory of self-efficacy [10]. The theory suggests that to enhance self-efficacy (the confidence an individual has that he or she can achieve a particular objective) an intervention should increase autonomy, reduce negative perceptions of being different, offer vicarious learning and modelling opportunities from peers, encourage setting of achievable goals, and give rewards for such achievements. Enhanced self-efficacy, in turn, increases the implementation of successful self-management. A systematic literature review of behavioral interventions for adolescents with type 1 diabetes found that those interventions that were theoretically based were significantly more effective than those that were not [11].

The aim of this research was to consult stakeholder groups and explore their perspectives on the desirable, important, and feasible characteristics of an Internet-based virtual clinic system for people who have type 1 diabetes and use insulin pumps, and to flag any potential worries or concerns. This builds on a recently conducted survey of potential users [12]. Gathering stakeholder perspectives is a vital first stage in order to ensure that the system developed meets the needs of future users [13] and that it achieves its full potential in their eyes [14]. Currently, much of the work to gather user feedback takes place following development of the system in question. In contrast, the objective of the study reported here was to undertake detailed consultation with a variety of stakeholder groups and feed their comments directly into the process of developing the system. It was envisaged that the virtual clinic would provide people living with diabetes, and their health professionals, access to the records of their condition (including uploaded blood glucose readings), a messaging facility, information and advice for patients, and a peer-to-peer support area, thereby meeting many of the criteria identified as desirable within Internet-based diabetes management [15].

Methods

In order to gather detailed comments from the stakeholder groups identified, a qualitative approach was taken. Several elements were used, including focus groups, individual interviews, email consensus gathering, and an expert workshop.

Participants in both the focus groups and interviews received an information sheet about the study and gave their written consent to take part. The elements of the study were approved by an NHS Local Research Ethics Committee.

Focus Groups

Three focus groups were held with people who had type 1 diabetes and who used insulin pumps to control their disease. These participants were recruited from a local diabetes clinic. Each focus group had between 3 and 5 participants. The total number of participants was 12 (2 male and 10 female), and all were over 16 years of age. All participants had responded to a

written invitation to take part in the research, indicating at least a basic level of literacy. All those who responded to our written invitations were subsequently asked to take part in one of the focus groups. Although focus groups were held during the day and evening, a small number of those willing to take part could not attend any of the sessions. These people were invited to submit their views by email, which were included in the email consensus discussed below. Focus groups were conducted at the education center within the hospital at which the clinic was based, a convenient and familiar location. No health professionals were present at the focus group sessions. At each session, a short demonstration was given to familiarize participants with the concepts we were planning to develop. This comprised the following elements: (1) a demonstration of an existing system that could record physiological data (such as blood glucose readings) and facilitate patient/health professional messaging, and (2) a description of some additional features we thought may be of interest to potential users, including downloadable information, links to other relevant websites, and peer-to-peer discussion boards. The facilitator then used the topic guide to focus the discussion on the following: participants' initial reactions to the concept, the most and least important/useful elements, whether they would be likely to use such a system, factors that may facilitate or hinder use, what the benefits may be, and any concerns they may have. Sessions were audio-recorded and lasted an average of 1 hour.

Individual Interviews

Individual interviews were carried out with 5 health professionals from the same clinic (2 doctors, 2 specialist nurses, 1 dietician) and with 1 representative from an insulin pump manufacturer/supplier who had emerged from the focus group discussions as an important source of information and support. The same demonstration was given as for the focus groups, and the interviews included similar questions. All interviews were audio-recorded and lasted between 30 minutes and 1 hour.

Consensus Gathering by Email

A process of email consensus gathering was carried out with patients using Internet-based systems for other chronic conditions in order to gain their views on the potential of the virtual clinic and their experience of using other systems. A message was posted on discussion boards aimed at patients with long-term conditions, including asthma, epilepsy, and mental health problems, and also on a discussion board for insulin pump users in order to consult further with the target group for the intervention being developed. Users were invited to contribute their views by email.

Expert Workshop

An expert workshop was held at which invited delegates from the United Kingdom discussed the role of information technology (IT) in diabetes care. Delegates had a range of backgrounds, with expertise in areas such as diabetes management, the use of IT in health care, and patient involvement in long-term conditions. They included academic researchers (with backgrounds as diverse as primary care, psychology, psychiatry, engineering, and sociology), diabetologists, people living with diabetes, and a representative

from the charity Diabetes UK. Delegates discussed the proposed system in small groups and completed questionnaires addressing the same issues as in the focus groups and interviews.

Data Analysis

Qualitative data analysis was undertaken by the first and second author. Focus group and interview transcripts, email responses from users of other systems, and the questionnaires completed during the expert workshop were all preliminarily analyzed independently by each author using a thematic analysis approach. Recurrent themes were identified as they emerged from the data, rather than on the basis of researcher preconceptions. Following this, the two authors held an analysis meeting at which the emergent themes identified by each were compared and discussed. Following agreement on these, each author undertook a full analysis of approximately half the transcripts, with a small overlap to allow comparison of theme interpretation and allocation of data extracts to particular themes.

Results

Six key themes, some with subthemes, were identified from the data and are described in detail below:

1. communication between patients and health professionals
2. presentation of patient data and permanency of the record
3. the importance and value of peer support
4. awareness of the personal nature of diabetes
5. how an Internet-based system would fit with the current provision of care
6. an Internet-based system may not be suitable for all people with diabetes

Communication Between Patients and Health Professionals

The facility to send and receive secure messages through an Internet-based system was largely welcomed by all the stakeholder groups consulted, although there were some concerns. For health professionals, one of the key benefits was that the means through which they communicated with their patients outside of their face-to-face appointments would be more standardized. Having an asynchronous messaging system was seen to be beneficial as both parties could check their messages when it was convenient for them and fit this into their other activities, thereby saving a lot of time. This concurs with previous research [16,17]. For example, one health professional (HP) illustrated the time taken in trying to respond to a patient enquiry that had been left on the clinic answer phone:

I rang somebody and she said, "I'm right in the middle of shopping in [town]. Can you ring me back?" So, there's a lot of time that you spend sort of ringing them. Or, I've got a school teacher, and she's rang four times today and I have rang her back several times, and you know eventually [that] you get to talk but it might be a couple of days, so I think quite a lot of our patients feel that it's easier to email us. [HP6]

Similarly, a focus group participant (FGP) explained how she had largely given up trying to seek advice through the clinic

helpline as she found having to leave a message and waiting to be called back simply unworkable:

It's so hard.... You have to leave a message, they ring you back, you're in an open office so you can't talk, so you just bash your way through it and hope it's going to come out right at the end. [FGP3]

Being able to send a message and receive a reply was therefore welcomed by patients, so long as the system was adequately supported and health professionals responded to messages in a timely manner [18]. There were potential concerns raised by some health professionals that the number of patient contacts would increase and that the workload of the consultants would similarly increase since most patient messages had previously been filtered by the specialist nurses. These were fairly minor concerns though and did not outweigh the perceived benefits.

Presentation of Patient Data and Permanency of the Record

One of the key advantages of the proposed Internet-based system identified by health professionals was the benefit of having ready access to a patient's blood glucose readings. The advantages were two-fold: (1) having up-to-date results readily accessible, and (2) having results in a standard format. At present, health professionals were often faced with results in various formats:

Some of them [records kept by patients] might be too scribbled and not really very clear.... But, yes, scraps of paper might have, sometimes, [the time of day]...written in the corner, and you're trying to see the time of the day and sort of trying to make them in your head into [a] date profile to some extent. So, they can be a bit difficult, yes. [HP2]

The permanency of the record was also seen as beneficial by health professionals since, in contrast to a paper diary, the readings would not leave with the patient and would still be available for review and reflection at a later date. The potential for use as a teaching and training tool was also raised. In contrast, patients were more divided on the benefits of uploading their results. Those who found recording their readings problematic welcomed anything that would make it easier and more convenient, but for others the system was seen to offer nothing new. However, all patients recognized the benefit of the health professional having access to their readings when dealing with queries.

The Importance and Value of Peer Support

In common with other research [5,19,20], one of the most valuable elements of the system identified by patients, and also recognized by health professionals, was peer-to-peer support. Patients in particular identified two key ways in which such support would be useful. First, being able to pick up tips and suggestions for managing their diabetes was viewed as beneficial, even for those who had had diabetes for some time:

But there are so many things that you can come across for the first time, and the one I had to seek advice on was the flu injection last year, which caused chaos, and I thought, "Well, is this the flu injection or is

there something else that I'm missing?".... But again, a [discussion] board like this, just to push the question in and see what response you get back... [FGP6]

Second, being able to communicate with someone who understood what they were going through was considered very valuable by patients, particularly for those who did not know others with diabetes, thereby demonstrating the potential of Internet-based communication to move beyond the individual's usual sphere of contacts.

Within my life, around work, around home, and around socializing, I know no one else with diabetes, so to be able to get on...[the] Internet and sort of like to be able to have a chat with somebody... Even like you say, sometimes it might just be for a moan, but sometimes you can get so frustrated with it that you just want to be able to take it all out.... You want to be able to talk to someone, but finding... You talk to your partner and you talk to your friends, but they don't always understand what you're saying. [FGP9]

The health professionals also recognized the potential benefits of peer-to-peer support for their patients, particularly the ways in which Internet-based provision extended the group sessions currently offered at the clinic.

We have group sessions...and they are teaching each other...from their experience.... And they'll be able to do that on a regular basis, on a daily basis, rather than on a three-monthly basis, and they'll all be there in the chat room, potentially, rather than just the people who turned up to the group clinic. [HP1]

However, there were some minor concerns among health professionals that the peer-to-peer support could lead to the "propagation of myths" and to individuals passing on what had worked for them to others for whom it may be inappropriate. It was generally felt that monitoring the topics under discussion would be sufficient to counter these concerns. The issue of patients' awareness and appreciation of the individual nature of diabetes is discussed below. A recent systematic review of the effects of online peer-to-peer interactions [21] found no evidence to support concerns of causing harm, although it also failed to find robust evidence of any health benefit.

Awareness of the Personal Nature of Diabetes

As stated above, one of the concerns health professionals had about the peer-to-peer support was that patients using it may not realize that what had worked for one person may not work for others:

Patients frequently want to propagate what has been good for them, and it might not be good for everybody, if you know what I'm saying. They might have found some particular way of dealing with situations...but it doesn't always apply to everybody. [HP1]

However, it was clear from the focus groups with patients that they were very well aware of how diabetes affected them personally and that this may differ markedly between individuals. The knowledge and experience they had built up from managing their diabetes on a day-to-day basis were substantial, and they spoke at length about the need to evaluate

suggestions posted on discussion boards in the light of their personal situation. The following is a typical example:

Whatever I read on there [the discussion board] may be useful, but I know it's not individually designed for me. So, overall it can be a useful guide when you start out, but you really do have to know your own system, don't you. [FGP8]

This ability on the part of potential users to appraise the information posted on a discussion board and evaluate whether it is likely to be useful to them, together with existing research demonstrating that most information posted on boards of this kind is accurate (or very quickly corrected by other participants) [22], indicates that the concerns voiced by health professionals are likely to be largely unfounded.

How an Internet-Based System Would Fit With the Current Provision of Care

The general consensus among the focus group participants and health professionals was that the proposed Internet-based system would fit well with existing clinic provision. The system would provide a useful means of communication and support between routine clinic appointments and would bring benefits in terms of increased standardization, more efficient means of communication, and extended scope of current group consultations through the use of Internet-based asynchronous discussion boards and real-time chat rooms. Some concerns were raised about a potential increase in the demand of health professionals' time for dealing with messages received from patients, and this was identified as a potential factor that could hinder usage. However, the general consensus was that, in the local clinic at least, any additional time would be minimal and may well be offset through the time savings brought about by other benefits. This is one of the issues to be explored in the planned pilot study.

An Internet-Based System May Not Be Suitable for All

The vast majority of participants in the stakeholder consultation raised the point that an Internet-based system of the type proposed would not be suitable or appropriate for all people with diabetes. The main issue centered on the need for potential users to be comfortable with the required technology and to have a computer with Internet access. It was suggested that this may limit potential users to those who are younger and more familiar with technology, as supported by research by Giménez-Pérez et al [23]. However, a recent study by McKay et al [6] has shown that novice computer users will participate in an Internet-based system to assist their self-management. A further concern, raised particularly by a health professional, was that the system could be unsuitable for those with a poor understanding of their condition or for whom diabetes is only one of many health problems.

Discussion

The research was successful in consulting stakeholder groups to explore their perspectives on the desirable, important, and feasible characteristics of an Internet-based system for diabetes. Involvement of stakeholders is a vital early stage in the

development of a complex intervention, yet all too often their views are only gathered in later stages. Stakeholders have clear and relevant views on what such a system should provide, and these views can be captured and synthesized with relative ease. In this case, we identified 6 themes that are supported by findings from previous studies on the perceived convenience of an asynchronous messaging system [16,17] and the importance and value of peer-to-peer support [5,19,20]. The themes also contribute to debate about who is likely to use Internet-based systems and suggest that patients are eminently capable of assessing advice posted by peers on a discussion board and relating it to their personal situation. By basing the focus groups and interviews in an existing diabetes clinic, the research was able to suggest that an Internet-based system is likely to fit well with existing care provision as well as to explore the likely impact on health professionals' time.

The main strength of this study is the involvement of a diverse range of stakeholders in detailed consultation during the development stage of an Internet-based intervention. The diverse range of methods used to involve these stakeholders (ie, interviews, focus groups, email consensus gathering, and an expert workshop) is also a key strength. In particular, the use

of focus groups and a workshop means that we allowed participants to discuss and develop each others' perspectives. However, there are also some limitations to the study. Patients were self-selecting volunteers, so they may have different views from those who did not volunteer to participate. It is noticeable, for example, that more women than men volunteered to take part. There is also a potential issue concerning how the perspectives of patients using insulin pumps relate to those of patients who control their type 1 diabetes through other means. We are aware that much of the feedback from patients does not specifically relate to their use of an insulin pump; rather, it is concerned with more generic issues such as communication with their health professionals and the value placed on peer support. However, we acknowledge that further stakeholder consultation with the target group(s) would be necessary before the intervention could be developed for diabetes patients who do not use insulin pumps.

This consultation demonstrated that an Internet-based system is attractive to the stakeholders consulted during the course of this study and has led to the design of a system that is able to meet their needs. This system has now been developed and is being evaluated in a pilot study.

Acknowledgments

The authors wish to thank all those who participated in the research and Voluntis for allowing access to their software for demonstration. The expert workshop was supported by a grant from the Diabetes Foundation. The study was funded through a Warwick Medical School Postdoctoral Fellowship.

Conflicts of Interest

None declared.

Multimedia Appendix

Presentation of this research (ppt) [[PPT file \(MS Powerpoint\), 65 KB - jmir_v9i3e23_app1.ppt](#)]

References

1. Ferrer-Roca O, Franco Burbano K, Cárdenas A, Pulido P, Diaz-Cardama A. Web-based diabetes control. *J Telemed Telecare* 2004;10(5):277-281. [Medline: [15494085](#)]
2. Bellazzi R, Larizza C, Montani S, Riva A, Stefanelli M, D'Annunzio G, et al. A telemedicine support for diabetes management: the T-IDDM project. *Comput Methods Programs Biomed* 2002 Aug;69(2):147-161. [Medline: [12100794](#)] [doi: [10.1016/S0169-2607\(02\)00038-X](#)]
3. Gómez EJ, Hernando ME, García A, Del Pozo F, Cermeño J, Corcoy R, et al. Telemedicine as a tool for intensive management of diabetes: the DIABTel experience. *Comput Methods Programs Biomed* 2002 Aug;69(2):163-177. [Medline: [12100795](#)] [doi: [10.1016/S0169-2607\(02\)00039-1](#)]
4. Cavan DA, Everett J, Plougmann S, Hejlesen OK. Use of the Internet to optimize self-management of type 1 diabetes: preliminary experience with DiasNet. *J Telemed Telecare* 2003;9 Suppl 1(S1):S50-S52. [Medline: [12952722](#)] [doi: [10.1258/135763303322196330](#)]
5. McKay HG, Feil EG, Glasgow RE, Brown JE. Feasibility and use of an Internet support service for diabetes self-management. *Diabetes Educ* 1998;24(2):174-179. [Medline: [9555356](#)] [doi: [10.1177/014572179802400207](#)]
6. McKay HG, Glasgow RE, Feil EG, Boles SM, Barrera M. Internet-based diabetes self-management and support: initial outcomes from the diabetes network project. *Rehabil Psychol* 2002;47(1):31-48.
7. Glasgow RE, Boles SM, McKay HG, Feil EG, Barrera M. The D-Net diabetes self-management program: long-term implementation, outcomes, and generalization results. *Prev Med* 2003 Apr;36(4):410-419. [Medline: [12649049](#)] [doi: [10.1016/S0091-7435\(02\)00056-7](#)]

8. McMahon GT, Gomes HE, Hickson Hohne S, Hu TMJ, Levine BA, Conlin PR. Web-based care management in patients with poorly controlled diabetes. *Diabetes Care* 2005 Jul;28(7):1624-1629 [FREE Full text] [Medline: [15983311](#)] [doi: [10.2337/diacare.28.7.1624](#)]
9. Goldberg HI, Ralston JD, Hirsch IB, Hoath JI, Ahmed KI. Using an Internet comanagement module to improve the quality of chronic disease care. *Jt Comm J Qual Saf* 2003 Sep;29(9):443-451. [Medline: [14513667](#)]
10. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977 Mar;84(2):191-215. [Medline: [847061](#)] [doi: [10.1037/0033-295X.84.2.191](#)]
11. Hampson SE, Skinner TC, Hart J, Storey L, Gage H, Foxcroft D, et al. Behavioral interventions for adolescents with type 1 diabetes: how effective are they? *Diabetes Care* 2000 Sep;23(9):1416-1422 [FREE Full text] [Medline: [10977043](#)] [doi: [10.2337/diacare.23.9.1416](#)]
12. Lowe P, Hearnshaw H, Griffiths F. Attitudes of young people with diabetes to an Internet-based virtual clinic. *J Telemed Telecare* 2005;11 Suppl 1(S1):59-60. [Medline: [16035996](#)] [doi: [10.1258/1357633054461840](#)]
13. 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 [FREE Full text] [Medline: [15142919](#)] [doi: [10.1136/bmj.328.7449.1159](#)]
14. Kerr C, Murray E, Stevenson F, Gore C, Nazareth I. Internet interventions for long-term conditions: patient and caregiver quality criteria. *J Med Internet Res* 2006;8(3):e13 [FREE Full text] [Medline: [16954123](#)] [doi: [10.2196/jmir.8.3.e13](#)]
15. Mazzi CP, Kidd M. A framework for the evaluation of Internet-based diabetes management. *J Med Internet Res* 2002 Jan 10;4(1):e1 [FREE Full text] [Medline: [11956033](#)] [doi: [10.2196/jmir.4.1.e1](#)]
16. Tjora A, Tran T, Faxvaag A. Privacy vs usability: a qualitative exploration of patients' experiences with secure Internet communication with their general practitioner. *J Med Internet Res* 2005;7(2):e15 [FREE Full text] [Medline: [15998606](#)] [doi: [10.2196/jmir.7.2.e15](#)]
17. Lin CT, Wittevrongel L, Moore L, Beaty BL, Ross SE. An Internet-based patient-provider communication system: randomized controlled trial. *J Med Internet Res* 2005;7(4):e47 [FREE Full text] [Medline: [16236699](#)] [doi: [10.2196/jmir.7.4.e47](#)]
18. Car J, Sheikh A. Email consultations in health care: 2--acceptability and safe application. *BMJ* 2004 Aug 21;329(7463):439-442 [FREE Full text] [Medline: [15321903](#)] [doi: [10.1136/bmj.329.7463.439](#)]
19. Iafusco D, Ingenito N, Prisco F. The chatline as a communication and educational tool in adolescents with insulin-dependent diabetes: preliminary observations. *Diabetes Care* 2000 Dec;23(12):1853. [Medline: [11128367](#)] [doi: [10.2337/diacare.23.12.1853b](#)]
20. Zrebiec JF, Jacobson AM. What attracts patients with diabetes to an internet support group? A 21-month longitudinal website study. *Diabet Med* 2001 Feb;18(2):154-158. [Medline: [11251681](#)] [doi: [10.1046/j.1464-5491.2001.00443.x](#)]
21. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 15;328(7449):1166 [FREE Full text] [Medline: [15142921](#)] [doi: [10.1136/bmj.328.7449.1166](#)]
22. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. *BMJ* 2006 Apr 22;332(7547):939-942 [FREE Full text] [Medline: [16513686](#)] [doi: [10.1136/bmj.38753.524201.7C](#)]
23. Giménez-Pérez G, Gallach M, Acera E, Prieto A, Carro O, Ortega E, et al. Evaluation of accessibility and use of new communication technologies in patients with type 1 diabetes mellitus. *J Med Internet Res* 2002 Dec 20;4(3):E16 [FREE Full text] [Medline: [12554547](#)] [doi: [10.2196/jmir.4.3.e16](#)]

submitted 19.03.07; peer-reviewed by K Nazi, J Car; comments to author 08.05.07; revised version received 13.06.07; accepted 23.07.07; published 09.08.07.

Please cite as:

Armstrong N, Hearnshaw H, Powell J, Dale J

Stakeholder Perspectives on the Development of a Virtual Clinic for Diabetes Care: Qualitative Study

J Med Internet Res 2007;9(3):e23

URL: <http://www.jmir.org/2007/3/e23/>

doi: [10.2196/jmir.9.3.e23](#)

PMID:

provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

An Evaluation of Personal Health Information Remnants in Second-Hand Personal Computer Disk Drives

Khaled El Emam¹, BEng, PhD; Emilio Neri², BEng; Elizabeth Jonker³, BA

¹University of Ottawa and Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada

²TrialStat Corporation, Ottawa, Canada

³Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada

Corresponding Author:

Khaled El Emam, BEng, PhD

CHEO Research Institute

401 Smyth Road

Ottawa, Ontario K1H 8L1

Canada

Phone: ++1 613 797 5412

Fax: ++1 613 741 0397

Email: kelemam@uottawa.ca

Abstract

Background: The public is concerned about the privacy of their health information, especially as more of it is collected, stored, and exchanged electronically. But we do not know the extent of leakage of personal health information (PHI) from data custodians. One form of data leakage is through computer equipment that is sold, donated, lost, or stolen from health care facilities or individuals who work at these facilities. Previous studies have shown that it is possible to get sensitive personal information (PI) from second-hand disk drives. However, there have been no studies investigating the leakage of PHI in this way.

Objectives: The aim of the study was to determine the extent to which PHI can be obtained from second-hand computer disk drives.

Methods: A list of Canadian vendors selling second-hand computer equipment was constructed, and we systematically went through the shuffled list and attempted to purchase used disk drives from the vendors. Sixty functional disk drives were purchased and analyzed for data remnants containing PHI using computer forensic tools.

Results: It was possible to recover PI from 65% (95% CI: 52%-76%) of the drives. In total, 10% (95% CI: 5%-20%) had PHI on people other than the owner(s) of the drive, and 8% (95% CI: 7%-24%) had PHI on the owner(s) of the drive. Some of the PHI included very sensitive mental health information on a large number of people.

Conclusions: There is a strong need for health care data custodians to either encrypt all computers that can hold PHI on their clients or patients, including those used by employees and subcontractors in their homes, or to ensure that their computers are destroyed rather than finding a second life in the used computer market.

(*J Med Internet Res* 2007;9(3):e24) doi:[10.2196/jmir.9.3.e24](https://doi.org/10.2196/jmir.9.3.e24)

KEYWORDS

Privacy; confidentiality; security; data disclosure

Introduction

The adoption of electronic medical records is growing [1-5]. Concurrently, a majority of patients, and the public in general, are concerned about unauthorized disclosure and use of their personal health information (PHI) in an era of the electronic medical record [6-10].

Concern about privacy has caused some members of the public to not be totally honest with their health care provider [10]. A

survey in the United States found that as many as 15% of adults have changed their behavior to protect their privacy [6]. Those behavior changes include going to another doctor, paying out-of-pocket when insured to avoid disclosure, not seeking care to avoid disclosure to an employer, giving inaccurate or incomplete information on medical history, and asking a doctor not to write down the health problem or to record a less serious or embarrassing condition. More than a quarter of teens indicated that they would not seek out health care if they had concerns about the confidentiality of their information [11]. In

a survey of US physicians, nearly 87% reported that a patient had asked that information be kept out of their record, and nearly 78% of physicians said that they had withheld information from a patient's record due to privacy concerns [12]. Similar behaviors have been reported in Canada. A survey estimated that 12% of Canadians have withheld information from a health care provider because of concerns over whom the information might be shared with or how it might be used [13], and an estimated 735000 Canadians decided not to see a health care provider for the same reasons [14]. Such behavior changes can reduce the accuracy of health data [15-18].

Due to inaccurate data, patient safety may be jeopardized: clinicians may make treatment errors [19] or make errors in ordering medications [20]. Furthermore, researchers may underestimate disease prevalence [21], and health system managers may underestimate compliance with standards of care such as vaccination guidelines [22]. Health care organizations may be fined if they report inaccurate data to government agencies [23].

While federal and provincial health care privacy legislation, such as the Health Insurance Portability and Accountability Act in the United States and the Personal Health Information Privacy Act in Ontario, do motivate organizations to properly protect PHI, we do not know the extent to which that has actually been effective in eliminating inadvertent disclosures of PHI.

One relatively easy way to get personal information (PI) about other people is through the disk drives available on the second-hand computer market [24-26]. These computers may have been deliberately resold by their owners (individuals or organizations), donated to good causes (eg, charities or schools) who subsequently sold them, or may have ended up on the second-hand market after they were lost or stolen. In this study we examine the data remnants in second-hand disk drives to determine the extent to which PHI is inadvertently leaking from data custodians. To our knowledge, there have been no studies that have attempted to assess the extent to which PHI can be inappropriately disclosed in this way.

Methods

One approach to evaluate the extent to which PHI is leaking from data custodians is to count the number of security breaches that are publicized in the media. This, however, has a number of disadvantages: (1) not all security breaches involve PHI (eg, many are of financial data) and media reports may not make the distinction, (2) not all security breaches result in PHI being disclosed (eg, the data was encrypted) and such details usually do not appear in media reports, and (3) only some US states and only one Canadian province have breach notification laws [27]. Therefore, it is plausible that many breaches never get reported in the media.

Consequently, in this study we estimate the prevalence of PI and PHI leaks through second-hand disk drives. Our measure is the proportion (percentage) of second-hand disk drives available on the reseller market with PI and PHI on them. We purchased functional computer disk drives from the second-hand computer equipment reseller market across Canada and

examined their contents using digital forensic tools. All nonfunctional drives were returned and replaced.

This study was conducted in the winter of 2006/07. Ethics approval was obtained from the Research Ethics Board at the Children's Hospital of Eastern Ontario (CHEO) Research Institute.

Number and Type of Drives

Our focus was on drives that would be used by individual end users (ie, in their desktop machines and laptops). This means that we excluded drives that were used in servers. Hence, we focused on smaller disk drives with a capacity range of 10 GB to 40 GB.

There is no previous research on PHI leaks in second-hand disk drives; therefore, we relied on data remnant studies of PI to determine the number of disk drives needed to estimate the prevalence of PHI [24-26]. We expected the proportion of drives that leak PHI to be smaller than the proportion leaking PI since there is less health information collected and stored electronically compared to other types of PI (eg, financial and legal information). We therefore expected the proportion of drives with PHI to be closer to the lower end of PI, which is 18% [26]. The size of the 95% confidence interval in previous studies that analyzed more than 12 functional drives ranged from $\pm 9\%$ to $\pm 11\%$ [26]. We then selected an interval value in the middle: $\pm 10\%$, which ensures that the precision of our estimates is within the expected range for this type of study. The minimal number of drives to obtain such a confidence interval is 57. Consequently, we aimed to get data from 60 functional disk drives.

Identifying Vendors

A comprehensive list of 125 Canadian second-hand computer equipment vendors was identified from telephone directories, contacts and experts in the computer industry, Canadian vendors selling on eBay during the study period, local business directory searches, and a Google search to find "used computer equipment in Canada." The results were reviewed to form a list of 40 credible potential vendors distributed geographically across the country. We went down the shuffled list and systematically contacted these vendors via telephone and/or email for more information on their inventory. Used disk drives were purchased either in person, over the phone, or the Web, and were picked up personally or shipped to our lab. We limited the maximum number of drives per vendor to 10 so as to ensure a wider distribution of sources.

After contact, some vendors were excluded for a number of reasons:

- They would not sell individual disk drives. Due to cost constraints, we could only purchase stand-alone drives rather than fully configured second-hand computers from which the drives could be physically extracted.
- Some vendors in rural areas did not want to ship the equipment across the country or did not accept payments by phone or over the Internet.
- Some vendors did not have disk drives within the stipulated size range in stock at the time of the study.

We were able to purchase equipment from 12 different vendors (multiple sites for retail chains were counted as multiple vendors).

Data Recovery

Functioning drives can be classified as blank, recoverable, or securely wiped. Blank drives were readable but there was no data on them at all, current or deleted. Data on drives that have been formatted or repartitioned can be quite easily recovered [24]. Files that are deleted are also recoverable since a delete does not actually remove the data from the drive but only removes the entry from the file directory. We used a commercial software package called Recover My Files (GetData Pty Ltd, Hurstville, NSW, Australia, version 3.98, build 5282) to retrieve the data from the drives and recover the files that have been deleted [28]. The same tool was used to recover data from formatted and repartitioned disk drives. Further information about data recovery is included in the Multimedia Appendix.

It is possible to use tools that implement a specific secure delete algorithm that ensures that the data cannot be recovered. The DOD 5220.22-M standard is a US Department of Defense standard providing specifications for clearing and sanitizing electronic data storage devices [29-31]. There are some commercial and freely available tools that implement that standard (eg, see [32]). It is not possible to extract the data from such drives. For drives that were not blank and that were not recoverable, we used a hexadecimal editor to read the patterns of data on the disk. Disks that have been wiped using this approach either have a single character (usually a zero) written to the disk or have a characteristic pattern of alternate ones and zeros followed by a random character written to the disk.

Data Coding and Analysis

All data from the recovered drives were stored on DVDs. A search of the files on each DVD was performed in order to isolate files that may have contained PI and PHI. The DVDs were searched for Microsoft Word documents, Excel documents, PowerPoint documents, Outlook files, Access database files, Adobe (PDF) documents, and text files. All discovered files were manually reviewed and a summary of the discovered PI and PHI was completed for each disk. An attempt was also made to identify the owner(s) of each disk drive from the information it contained.

PI was defined as identifying information (eg, name, address, social insurance number) about an individual or individuals *plus* other sensitive but nonmedical information (eg, financial information, personal correspondence, divorce documents, legal records). PHI was defined as identifying information about an individual or individuals *plus* any sensitive physical or mental health information. By this definition, if a drive only had a list of names and addresses but no sensitive information associated with them, then such data would not be considered PI or PHI. This definition is somewhat conservative because one can argue that a list of names and addresses suggests that all of these individuals were associated in some way; hence, association information would be revealed. Therefore, our results would be considered a lower bound on the prevalence of PI and PHI in the second-hand drives we analyzed.

Four ratings were made for each drive depending on the type of information it contained: (1) PI about the owner, (2) PI about other individuals apart from the owner, (3) PHI about the owner, and (4) PHI about other individuals apart from the owner. If a drive was clearly owned by multiple people (eg, members of a family), then they were all considered the owner in our coding. This means that, for example, if PI existed on any one of them, then the drive was considered to contain PI on the owner.

When considering whether information about the owner was really PI or not, we needed to decide what to do about *work products*. A work product is the output of an individual's professional or employment responsibilities. For example, a physician's prescription record would be considered a work product of the physician, irrespective of whether the patient is identifiable or not. The Federal Privacy Commissioner of Canada does not consider work products to be PI [33]. However, the European Commission had a different interpretation and considers information on and *relating to* an individual regardless of the position or capacity of the individual, such as a prescription record written by a physician, a communication of PI [34]. Given the uncertainty across jurisdictions, we treated information deemed to be work products as PI in one analysis and not PI in another analysis and present both sets of results.

Two independent individuals rated the drives. The first rater analyzed all of the 60 drives. The second rater analyzed a subset of the drives to ensure that the coding was reliable. Where there was disagreement, the two raters met to discuss their rationale and reach a consensus rating.

To determine how many drives the second rater needed to analyze, we performed a power analysis for using the Kappa statistic [35]. Given that there are no precedents for the interrater reliability of data extraction from second-hand drives, we relied on generally accepted benchmarks for Kappa values. Hartmann notes that Kappa values should exceed 0.6 [36]. Landis and Koch provide a more general benchmark, where values between 0.4 and 0.6 are considered moderate agreement [37]. A similar benchmark is provided by Altman [38]. Fleiss suggests that values between 0.4 and 0.75 represent intermediate to good agreement [39]. To err on the conservative side, we assumed that our value of Kappa would be at least 0.5, which would be considered a moderate level of agreement according to the above benchmarks. At that level of agreement, and 80% power to reject a null hypothesis comparing Kappa to agreement by chance, the second rater needed to code only 32 drives [40,41].

The final results are presented in terms of the percentage of disk drives containing PI and PHI on owners and other individuals, with the 95% confidence interval [42]. Interrater agreement is presented in terms of the Kappa statistic and its 95% confidence interval.

Special Protocols

Three special protocols were put in place for this study:

1. Some disk drives were expected to contain inappropriate/obscene material (eg, pornography). We therefore did not explicitly look through image files (file extensions .gif, .jpg, .psd, .tif, and .bmp). Also, one member of our research team initially screened the drives for files

and directories with suggestive names and flagged these particular drives as potentially containing such materials. No searching of files with suggestive names was done on flagged drives.

2. If any illegal materials were discovered (eg, child pornography), that information was to be passed on to the police.
3. If there were cases of disclosure of particularly sensitive PI or PHI for a large number of individuals, then they were to be reported to the appropriate (federal or provincial) privacy commissioner for follow-up.

Results

All of the 60 drives were from personal computers and ran the Windows operating system. Repartitioning and formatting are two common approaches for manipulating the drives. However,

as noted earlier, much data can be recovered despite these efforts. There were 35 drives that were repartitioned or formatted, and 5 that had had nothing done to them (all data were readily available). Therefore, data was potentially recoverable from 67% (95% CI: 54%-77%) of the drives.

A significant percentage of drives (28%; 95% CI: 19%-41%) were wiped using the DOD 5220.22-M standard. Three of the drives (5%) were completely blank and there had not been any data written to them. Five of the drives had pornographic files on them. Two of the drives were referred to the provincial privacy commissioner's office due to the sensitivity of the data that was found. No cases were referred to the police.

We were able to retrieve data from 39 drives (one of the repartitioned drives had no data on it). This represents 65% (95% CI: 52%-76%) of the total. Overall, we extracted 57 DVDs of data from the 39 drives.

Table 1. Contingency table with marginal totals and percentages showing the status of purchased drives distributed by province of purchase*

Province	Repartitioned	Formatted	DoD 5220.22-M Standard	Data Readily Available	Blank	Total
Ontario	19	11	5	4	3	42 (70%)
Quebec		5				5 (8%)
Alberta			12			12 (20%)
British Columbia				1		1 (2%)
Total	19 (32%)	16 (27%)	17 (28%)	5 (8%)	3 (5%)	60 (100%)

*For store chains, we considered the location of the specific store that we purchased from. The actual owners of the disk drives may be located in a different province or country. Four of the drives bought from Ontario belonged to US-based entities: 2 of them were state government departments, 1 was a municipal department, and 1 belonged to an individual.

There were 7 vendors in Ontario, 1 in Quebec, 3 in Alberta, and 1 in British Columbia. The distribution of drives by province is shown in [Table 1](#). There are relatively more drives purchased from Ontario. Data were extracted from the drives from 6 of the Ontario vendors. All of the drives purchased from Alberta had been securely cleared such that no data were recoverable.

After our data analysis, we contacted the 3 Alberta vendors to understand why they all had used secure methods for deleting data on the drives they resell. They all stated that they had internal standard operating procedures for doing secure deletes

on the second-hand disk drives that they sell because they do not want to get involved in any data breaches.

We were able to determine the address of the drive owner for 26 of the 39 drives with recoverable data. All of the 26 disk drives came from urban areas (using the Canada Post and the Canadian Medical Association definitions of a rural postal code, which is only one of multiple possible definitions [43]). The owners for 22 of these 26 drives were in the same province as the vendor. The other 4 were US-owned but were sold by Ontario vendors.

Table 2. Claims made by the vendors of the drives from which we were able to extract data

Vendor Statement About Wiping Drives	Count
“Like new condition”	1
Verbally stated that the drives were formatted	1
“Recertified to factory settings”	1
None	5

The 8 vendors from whom we bought drives that had data on them did not actually make any claims that the data would be

removed in a secure way (see [Table 2](#)). Therefore, they were not in breach of any agreements that they had made.

Table 3. Percentage of drives with recoverable files and percentage of total drives with available personal data

	Owner PI (A)*	Owner PI (B)*	Other PI	Owner PHI	Other PHI
Percentage of Recovered (95% CI)	72% (28/39) (56%-83%)	62% (24/39) (46%-75%)	56% (22/39) (41%-71%)	13% (5/39) (6%-27%)	15% (6/39) (7%-28%)
Percentage of Total (95% CI)	47% (28/60) (35%-59%)	40% (24/60) (29%-53%)	37% (22/60) (26%-49%)	8% (5/60) (7%-24%)	10% (6/60) (5%-20%)
Kappa† (95% CI)	0.8 (0.6-1)	0.6 (0.33-0.88)	0.78 (0.54-1)	0.76 (0.45-1)	0.795 (0.52-1)

* (A) indicates that work products were considered as PI, and (B) indicates that work products were not considered as PI.

† Interrater agreement Kappa scores and their 95% confidence intervals.

A summary of the type of data that was uncovered in these drives is shown in [Table 3](#). All Kappa scores were above our threshold of moderate agreement. The vast majority of drives with data had PI about their owners (close to half of all drives) and about others. Examples of PI found on the disk drives included:

- personal budgets, salary information, tax returns and completed tax filing forms
- letters regarding personal relationships
- information on life insurance policies and inheritances
- payroll records of employees, including addresses, dates of birth, and social insurance numbers
- email correspondence regarding employees and their actions
- police record checks
- divorce documents

A considerable percentage of the total drives had PHI about the owners (8%) and other people apart from the owner (10%). The vast majority of that information was in correspondence (eg, Word documents, PDF documents, and email). Examples of PHI found on the disk drives included:

- psychological assessments of adults and children, correspondence related to custody cases involving children, affidavits, and social history of abuse victims
- medical certificates
- letters regarding alcohol addictions of other individuals (not the owners of the drive)
- reports from a registered nurse about other individuals' health problems, cases of abuse, children's health, and medication lists
- correspondence regarding the placement of adults and children in long-term care facilities

The 6 disk drives with PHI about other people came from 3 different vendors; half of them were for personal use and the other half belonged to organizations.

Discussion

We found that PI, including PHI, was recoverable from 65% of the drives we purchased; 8% of the disk drives had PHI of the owners, and 10% of all drives had PHI on other people. Half of the latter set came from organizations that were directly entrusted with that information and the other half from people working for such organizations (eg, a nurse who took some information home to work on it).

Approximately 8% of personal computers in use worldwide are second-hand machines [44], out of a total installed base of 980 million in 2007 [45]. To the extent that our findings are generalizable, an 8%-10% prevalence of PHI in second-hand disk drives in use paints a disturbing picture about the inappropriate disclosure of PHI. The second-hand computer market is expected to grow significantly in the next few years [46,47] and, with it, the opportunity for further such disclosure of PHI.

In a previous data remnants study done in the United States [24], 158 drives were bought. Of these, 129 were successfully imaged. Approximately 9% were wiped. It was possible to extract data from many of the remaining drives (38%). A similar international study provided the results presented here in [Table 4](#). PI was recoverable from 60% of North American drives, although the sample size was quite small. Based on data from the United Kingdom and Australia, the range of drives with PI is 18% to 49%. While our findings on PI are consistent with previous studies, giving them some face validity, previous studies did not consider PHI.

Table 4. Summary of findings from an international data remnants study [26]

	UK and Australia (2005)	UK (2006)	Australia (2006)	Germany (2006)	North America (2006)
Total Drives	116	200	53	40	24
Faulty Drives	13 (11%)	87 (43%)	3 (6%)	30 (75%)	12 (50%)
Wiped*	17 (16%)	55 (49%)	18 (36%)	4 (40%)	1 (8%)
Had PI*	51 (49%)	35 (31%)	9 (18%)	3 (30%)	7 (60%)

*The percentage of these disk drives that were not faulty.

Prevalence of PHI

Our results indicate that not as much health information is leaking as other types of information, such as financial and legal information. Why is relatively less PHI available electronically on these drives?

In Canada, the use of computers and the Internet is quite common. The majority of the population has access to a home computer [48], and most citizens have access to the Internet [49,50]. However, this does not mean that they have easy electronic access to their own PHI.

There are a number of ways that individuals can get electronic access to their own health data. For example, individuals may request their medical records from the institution that provided them with care. In practice, very few hospitals provide medical records electronically or make them accessible [51]. Another study found that a very small percentage of members of an integrated delivery system used eHealth services when provided to them [52].

It is more likely that PHI will exist in correspondence, such as email. The proportion of US Internet users who reported communicating over the Internet with their health care provider in 2005 was 10% [53]; a European survey found that 4% have approached their family doctor over the Internet, and about 7% of email users in the United States exchange emails with physicians or health professionals [54]. The proportion of physicians who report communicating by email with their patients varies from 3.6% to 24% [55-57]. About a quarter of patients correspond via email with family members [50]. PHI may also be exchanged electronically with peers [58]. This is consistent with our findings in that most of the PHI we found was in correspondence rather than in electronic medical records or database files.

One would expect that as electronic medical records become more widely deployed, more PHI will be available to patients electronically and hence the risk of inappropriate disclosure of PHI will increase over time. The disclosure risk is highest with care and service providers who would have extensive electronic correspondence with and documentation on many patients and clients on their work and personal computers.

Practices for Securing Data on Disk Drives

There is clearly a need for organizations and individuals, and certainly in Ontario and to some extent Quebec, to take actions to reduce the risk of personal data leaking from second-hand disk drives. A disk can leave its custodian in three ways: it is destroyed, it is given away, or it is lost or stolen.

The safest way to dispose of a disk drive is to properly destroy it. Approximately 38% of all used personal computers, including their drives, are destroyed [59]. While there are a large number of techniques that an individual or organization can potentially employ to destroy equipment [60], many of them require specialized equipment or resources and it is therefore not practical for most users to do it themselves. However, destruction of equipment can be outsourced to specialized vendors.

If equipment will be donated or resold, the risks of PHI leaks remain high. Donated equipment may end up in foreign second-hand markets, as demonstrated by a recent case of British computers ending up in Africa [61]. Approximately 6% of second-hand computers are exported [59]. If not exported, local resellers will not necessarily wipe data from the drives they acquire [62].

There are three general approaches that can be pursued to protect data on equipment before it is given away: de-identification of data, encryption of data, and the use of secure delete technology. Such approaches should be applied on all computers that will hold PHI, including the personal computers of staff and contractors who may take data home to work off-site.

Any PHI on a disk drive ought to be de-identified at the earliest opportunity. De-identification ensures that the risk of finding out the identity of the individuals about whom the data pertains is low [63]. This means removing or masking directly identifying data and applying other anonymization techniques to protect indirectly identifying data [64]. However, there will be many cases when data need to be identifiable to be useful. Hence, additional techniques would also need to be used.

Another way to protect data is by using encryption technology. Encryption can be used to create specific virtual drives, and all sensitive information can be stored on the virtual drives. Unless the password used is weak or the encryption algorithm is compromised, it would be extremely difficult to extract the information. However, this is generally not enough. Many programs will store their data, temporary files, cached files, backup files, and registry values outside the encrypted virtual drive. Quite a significant amount of information can be left in these files. Most users would not know to change the settings of their applications to only use the encrypted disk drive, and sometimes that option is not available. Therefore, if one really wants to protect data, this would probably not be the best approach unless one possesses a great deal of technical expertise (to change the setting of the applications to force them to use the encrypted drive).

The best encryption technology to use is whole disk encryption [65] that is invoked before the operating system, during system boot, starts to operate. This ensures that all data on the drive (temporary, backup, and data) are encrypted. Fortunately, this type of technology is becoming more generally available in common operating systems and hardware. Therefore, one would expect that, in the next few years it will be much more widely deployed and will significantly reduce the risks we have identified. Specific stand-alone products are listed by the Privacy Commissioner of Ontario in a fact sheet [65].

The second technology one can use is secure delete. This allows one to delete all of the data on the drive so that they are not recoverable (such as when using the DoD 5220.22-M standard). Secure delete by itself, however, is not enough. One needs to perform a more general disk wipe. Software for wiping disks usually performs a secure delete as well as removing all of the temporary, backup, and cached files from the system.

A recent study noted that commercial software for wiping disks tends to be quite unreliable [66]. In one case, the software did

not even attempt a secure delete because of a software bug. The difficulty with wiping software is that the program needs to determine where each application keeps its information. This is difficult to do for a very large number of applications that change often. It has been argued that because the market for privacy tools is small (and hence the vendors have limited resources), such vendors will not be able to keep up-to-date with the application and operating system changes [66]. Therefore, while the use of wiping software is reassuring, it may not actually be sufficient to protect personal data on disk drives.

Even if an organization does not resell or donate its equipment, theft and loss are real risks. For instance, a recent survey noted that 47% of organizations reported theft of a laptop or mobile computing device [67]. Some recent health care examples: (1) the theft of 2 laptop computers containing the names, birth dates, addresses, PHI, and insurance information for 3000-4000 patients was reported after a break-in at a rehabilitation clinic [68], (2) a laptop computer that contained 51 assessment reports was stolen from the car of a psycho-educational consultant working for the school board [69], and (3) a laptop computer containing data on 2900 patients participating in clinical trials was stolen from a researcher's car [70]. All of the techniques described above that are used to protect data when equipment is donated or sold should therefore be considered even if there is no intention to part with the equipment. One cannot control loss or theft events.

In summary then, it is best to properly destroy equipment when it is no longer in use. Even if that is not possible or desirable, it is still advisable to have full drive encryption to be activated as soon as the computers are purchased. With full drive encryption, there is minimal risk (unless the passwords used are weak) if the disk drive is given away, lost, or stolen at a later date.

Acknowledgments

This work was partially funded by the Office of the Privacy Commissioner of Canada through its contributions program and by Bell Information and Communications Technology Inc through an ORNEC (Ontario Research Network for e-Commerce) grant. This work was approved by the Research Ethics Board of the Children's Hospital of Eastern Ontario Research Institute. We thank Mary Ocampo for her help with coding the data on the disk drives.

Conflicts of Interest

None declared.

Multimedia Appendix

Details of the data recovery process [[PDF file \(Adobe Acrobat\), 49 KB - jmir_v9i3e24_app1.pdf](#)]

References

1. Irving R. 2002 Report on Information Technology in Canadian Hospitals. Thornhill, ON: Canadian Healthcare Technology; 2003.
2. Healthcare Information and Management Systems Society (HIMSS). 15th Annual HIMSS Leadership Survey Sponsored by Superior Consultant Company. Healthcare CIO Results 2004.
3. Andrews JE, Pearce KA, Sydney C, Ireson C, Love M. Current state of information technology use in a US primary care practice-based research network. *Inform Prim Care* 2004;12(1):11-18. [Medline: [15140348](#)]

Limitations

In this study we only examined one source of leakage of PHI from data custodians. However, our results indicate that this is an important source of very sensitive PHI. As more information is stored and exchanged electronically, the risks from such leakage are bound to only increase unless current practices change.

We explicitly limited the study to Canadian vendors since this geographic location has not been studied before and because Canada has strong federal privacy laws. Hence, one would have expected that the ability to find PHI would be quite low—which was not the case.

The representativeness of the 60 drives of all drives in the second-hand market is a concern. We would argue that this sample underestimates the problem for a number of reasons. First, we excluded large disk drives, which eliminated data from servers. Servers would potentially contain large databases of PI and/or PHI. Second, some vendors (eg, those in rural locations) became suspicious of our motives for purchasing disk drives (“Why is someone from Ottawa buying a single drive from rural Alberta?”) and therefore refused to sell. We suspect that vendors with drives containing un-wiped PI and/or PHI were less likely to sell us the equipment. Therefore, our results should be considered a lower bound on the extent to which PI and PHI leaks through second-hand drives.

We did not specifically seek disk drives with health information on them. Had we done so, the PHI proportions would likely have been higher. However, that would not have provided a realistic assessment of the risk. Second-hand equipment vendors do not specialize by domain (ie, there were no vendors that specialize in selling only used equipment from health care facilities). Had we specifically requested equipment from health care institutions, it may have sounded like a suspicious request and dissuaded the vendors from completing the transaction.

4. Bower A. The Diffusion and Value of Healthcare Information Technology. Pittsburgh, PA: RAND Corporation; 2004.
5. Fonkych K, Taylor R. The State and Pattern of Health Information Technology Adoption. Pittsburgh, PA: RAND Corporation; 2005.
6. ; California HealthCare Foundation. Medical Privacy and Confidentiality Survey Summary and Overview. Oakland, CA: California HealthCare Foundation; 1999. URL: <http://www.chcf.org/documents/ihealth/survey.pdf> [accessed 2007 Sep 11] [WebCite Cache ID 5Q2YtO8yH]
7. ; Harris Interactive. Health information privacy (HIPAA) notices have improved public's confidence that their medical information is being handled properly. 2005. URL: <http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=894> [accessed 2007 Sep 11] [WebCite Cache ID 5Q2YQkeWM]
8. Grimes-Gruczka T, Gratzner C; The Institute for the Future. Ethics Survey of Consumer Attitudes about Health Web Sites. Oakland, CA: California HealthCare Foundation; 2000.
9. Willison DJ, Keshavjee K, Nair K, Goldsmith C, Holbrook AM; Computerization of Medical Practices for the Enhancement of Therapeutic Effectiveness investigators. Patients' consent preferences for research uses of information in electronic medical records: interview and survey data. *BMJ* 2003 Feb 15;326(7385):373 [FREE Full text] [Medline: 12586673] [doi: 10.1136/bmj.326.7385.373]
10. Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. *BMJ* 2001 Feb 3;322(7281):279-282 [FREE Full text] [Medline: 11157532] [doi: 10.1136/bmj.322.7281.279]
11. Cheng TL, Savageau JA, Sattler AL, Dewitt TG. Confidentiality in health care. A survey of knowledge, perceptions, and attitudes among high school students. *JAMA* 1993 Mar 17;269(11):1404-1407. [Medline: 8441216] [doi: 10.1001/jama.269.11.1404]
12. ; Association of American Physicians and Surgeons. New poll: doctors lie to protect patient privacy. 2001 Jul 31. AAPS Online. URL: <http://www.aapsonline.org/press/nrnewpoll.htm> [accessed 2007 Sep 11] [WebCite Cache ID 5NzvR0MX2]
13. ; EKOS Research Associates. Rethinking the Information Highway. . 2003.
14. Saravamuttoo M. Privacy: changing attitudes in a tumultuous time. Sixth Annual Privacy and Security Workshop; November 3-4, 2005; Toronto, ON. .
15. Gostin LO, Turek-Brezina J, Powers M, Kozloff R. Privacy and security of health information in the emerging health care system. *Health Matrix Clevel* 1995;5(1):1-36. [Medline: 10141742]
16. Hodge JG, Gostin LO, Jacobson PD. Legal issues concerning electronic health information: privacy, quality, and liability. *JAMA* 1999 Oct 20;282(15):1466-1471. [Medline: 10535438] [doi: 10.1001/jama.282.15.1466]
17. Woodward B. The computer-based patient record and confidentiality. *N Engl J Med* 1995 Nov 23;333(21):1419-1422. [Medline: 7477127] [doi: 10.1056/NEJM199511233332112]
18. Goldman J. Testimony before the subcommittee on health of the committee on ways and means on "Patient Confidentiality." 1998 Mar 24. .
19. Leape LL, Bates DW, Cullen DJ, Cooper J, Demonaco HJ, Gallivan T, et al. Systems analysis of adverse drug events. ADE Prevention Study Group. *JAMA* 1995 Jul 5;274(1):35-43. [Medline: 7791256] [doi: 10.1001/jama.274.1.35]
20. Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc* 2004;11(2):104-112 [FREE Full text] [Medline: 14633936] [doi: 10.1197/jamia.M1471]
21. Johnson N, Mant D, Jones L, Randall T. Use of computerised general practice data for population surveillance: comparative study of influenza data. *BMJ* 1991 Mar 30;302(6779):763-765 [FREE Full text] [Medline: 2021767]
22. Wilton R, Pennisi AJ. Evaluating the accuracy of transcribed computer-stored immunization data. *Pediatrics* 1994 Dec;94(6 Pt 1):902-906. [Medline: 7971009]
23. Davidson B, Lee Y, Wang R. Developing data production maps: meeting patient discharge data submission requirements. *Int J Healthcare Technol Manage* 2004;6(2):223-240. [doi: 10.1504/IJHTM.2004.004978]
24. Garfinkel S, Shilat A. Remembrance of data passed: a study of disk sanitization practices. *IEEE Secur Priv* 2003:17-27. [doi: 10.1109/MSECP.2003.1176992]
25. Jones A, Mee V, Meyler C, Gooch J. Analysis of data recovered from computer disks released for resale by organizations. *J Inform Warfare* 2005;4(2):45-53.
26. Jones A, Valli C, Sutherland I, Thomas P. The 2006 analysis of information remaining on disks offered for sale on the second hand market. *J Digit Forensics Secur Law* 2006;1(3):23-36.
27. ; Canadian Internet Policy and Public Interest Clinic. Approaches to Security Breach Notification: A White Paper. Ottawa, ON: Canadian Internet Policy and Public Interest Clinic; 2007. URL: http://www.cippic.ca/documents/bulletins/BreachNotification_9jan07-print.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5QwUai54N]
28. GetData. Recover My Files. URL: <http://www.recovermyfiles.com/> [accessed 2007 Sep 11] [WebCite Cache ID 5QCzWP9C4]
29. ; Department of Defense - Department of Energy - Nuclear Regulatory Commission - Central Intelligence Agency. DoD 5220.22-M National Industrial Security Program Operating Manual (NISPOM). Washington, DC: US Government Printing Office; 1995. URL: <http://download.ehealthinformation.ca/cite/nispo-522022m-1995.pdf> [accessed 2007 Sep 11] [WebCite Cache ID 5QJEDu55S]

30. ; Department of Defense. DoD 5220.22-M: National Industrial Security Program Operating Manual (NISPOM). Washington, DC: Department of Defense; 2006. URL: <http://www.dss.mil/isp/odaa/documents/nispom2006-5220.pdf> [accessed 2007 Sep 11] [WebCite Cache ID 5QExtKegU]
31. ; Updated DSS Clearing and Sanitization Matrix. 2007. DSS. URL: http://www.dss.mil/isp/odaa/documents/clearing_and_sanitization_matrix.pdf [WebCite Cache ID 5QJEQOMOT]
32. Russinovich M. SDelete v.1.51. 2006 Nov 1. Microsoft Technet. URL: <http://www.microsoft.com/technet/sysinternals/Security/SDelete.msp> [accessed 2007 Sep 11] [WebCite Cache ID 5QeDFx2bn]
33. Radwanski G. PIPED Act Case Summary #15. 2001 Oct 2. Office of the Privacy Commissioner of Canada. URL: http://www.privcom.gc.ca/media/an/wn_011002_e.asp [accessed 2007 Sep 11] [WebCite Cache ID 5QCtVSGEC]
34. ; The Working Party on the Protection of Individuals With Regard to the Processing of Personal Data. Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data. Adopted on 20th June. 2007. European Commission. URL: http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5Q2YBu0CR]
35. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;XX(1):37-46. [doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
36. Hartmann DP. Considerations in the choice of interobserver reliability estimates. *J Appl Behav Anal* 1977;10(1):103-116 [FREE Full text]
37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)] [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)]
38. Altman D. *Practical Statistics for Medical Research*. London, UK: Chapman and Hall; 1991.
39. Fleiss J. *Statistical Methods for Rates and Proportions*. Hoboken, NJ: Wiley; 1981.
40. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005 Mar;85(3):257-268. [Medline: [15733050](https://pubmed.ncbi.nlm.nih.gov/15733050/)]
41. Flack V, Afifi A, Lachenbruch P. Sample size determinations for the two rater kappa statistic. *Psychometrika* 1988;53(3):321-325. [doi: [10.1007/BF02294215](https://doi.org/10.1007/BF02294215)]
42. Agresti A, Coull B. Approximate is better than "exact" for interval estimation of binomial proportions. *Am Stat* 1998;52(2):119-126. [doi: [10.2307/2685469](https://doi.org/10.2307/2685469)]
43. Pong R, Pitblado J. Don't take geography for granted! Some methodological issues in measuring geographic distribution of physicians. *Can J Rural Med* 2001;6(2):103-112.
44. Escherich M. Thriving secondary PC market puts old PCs to good use. Gartner Inc. 2005.
45. Shiffler GIII. Forecast: PC installed base worldwide, 2003-2011. Gartner Inc. 2007.
46. Escherich M, Shao J, Maita K. Mature regions fuel supply of used PCs.. Gartner Inc - Dataquest. 2005.
47. Escherich M, Teng A. Thirst for technology drives used PC demand in emerging markets.. Gartner Inc. 2005.
48. ; Organization for Economic Development and Cooperation. OECD Key ICT indicators: households with access to a home computer. 2006. OECD. URL: <http://www.oecd.org/dataoecd/19/46/34083096.xls> [accessed 2007 Sep 11] [WebCite Cache ID 5QLh1C9bv]
49. ; Organization for Economic Development and Cooperation. OECD Key ICT indicators: households with access to the Internet in selected OECD countries. 2006. OECD. URL: <http://www.oecd.org/dataoecd/19/45/34083073.xls> [accessed 2007 Sep 11] [WebCite Cache ID 5QLgvVT4D]
50. Fox S, Fallows D. Internet health resources: health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Washington, DC: Pew Internet & American Life Project; 2003. URL: http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf [WebCite Cache ID Nba]
51. Fioriglio G, Szolovits P. Copy fees and patients' rights to obtain a copy of their medical records: from law to reality. *AMIA Annu Symp Proc* 2005:251-255 [FREE Full text] [Medline: [16779040](https://pubmed.ncbi.nlm.nih.gov/16779040/)]
52. Hsu J, Huang J, Kinsman J, Fireman B, Miller R, Selby J, et al. Use of e-Health services between 1999 and 2002: a growing digital divide. *J Am Med Inform Assoc* 2005;12(2):164-171 [FREE Full text] [Medline: [15561786](https://pubmed.ncbi.nlm.nih.gov/15561786/)] [doi: [10.1197/jamia.M1672](https://doi.org/10.1197/jamia.M1672)]
53. 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] [Medline: [17627929](https://pubmed.ncbi.nlm.nih.gov/17627929/)] [doi: [10.2196/jmir.9.3.e20](https://doi.org/10.2196/jmir.9.3.e20)]
54. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. *BMC Public Health* 2007;7(1):53 [FREE Full text] [Medline: [17425798](https://pubmed.ncbi.nlm.nih.gov/17425798/)] [doi: [10.1186/1471-2458-7-53](https://doi.org/10.1186/1471-2458-7-53)]
55. Liebhaber AB, Grossman JM. Physicians slow to adopt patient e-mail. *Data Bull (Cent Stud Health Syst Change)* 2006 Sep(32):1-2. [Medline: [16995287](https://pubmed.ncbi.nlm.nih.gov/16995287/)]
56. Grant RW, Campbell EG, Gruen RL, Ferris TG, Blumenthal D. Prevalence of basic information technology use by U.S. physicians. *J Gen Intern Med* 2006 Nov;21(11):1150-1155. [Medline: [16879417](https://pubmed.ncbi.nlm.nih.gov/16879417/)] [doi: [10.1111/j.1525-1497.2006.00571.x](https://doi.org/10.1111/j.1525-1497.2006.00571.x)]
57. Brooks RG, Menachemi N. Physicians' use of email with patients: factors influencing electronic communication and adherence to best practices. *J Med Internet Res* 2006;8(1):e2 [FREE Full text] [Medline: [16585026](https://pubmed.ncbi.nlm.nih.gov/16585026/)] [doi: [10.2196/jmir.8.1.e2](https://doi.org/10.2196/jmir.8.1.e2)]

58. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 15;328(7449):1166 [FREE Full text] [Medline: [15142921](#)] [doi: [10.1136/bmj.328.7449.1166](#)]
59. Escherich M, Smulders C. Market focus: worldwide secondary PC market sizing methodology. Gartner Inc. 2005.
60. ; Communications Security Establishment. Clearing and Declassifying Electronic Data Storage Devices. Ottawa, ON: Government of Canada; 2006. URL: <http://www.cse-cst.gc.ca/documents/publications/gov-pubs/itsg/itsg06.pdf> [accessed 2007 Sep 11] [WebCite Cache ID 50IbcPcnJ]
61. UK bank details sold in Nigeria. BBC News. 2006. URL: <http://news.bbc.co.uk/2/hi/business/4790293.stm> [accessed 2007 Sep 11] [WebCite Cache ID 5QczZlZLg]
62. Sullivan B. I just bought your hard drive. MSNBC News. 2006. URL: http://redtape.msnbc.com/2006/06/one_year_ago_ha.html [accessed 2007 Sep 11] [WebCite Cache ID 5QwUjM9pR]
63. 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] [Medline: [17213047](#)] [doi: [10.2196/jmir.8.4.e28](#)]
64. El Emam K, Jonker E, Sams S, Neri E, Neisa A, Gao T, et al. Pan-Canadian De-Identification Guidelines for Personal Health Information. Report Prepared for the Office of the Privacy Commissioner of Canada. 2007. URL: <http://www.ehealthinformation.ca/documents/OPCReportv11.pdf> [accessed 2007 Sep 11] [WebCite Cache ID 5Ow1Nko5C]
65. Cavoukian A. Office of the Privacy Commissioner of Ontario. Encrypting Personal Health Information on Mobile Devices. Fact Sheet. 2007 May. URL: http://www.ipc.on.ca/images/Resources/up-fact_12e.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5Qd1NU27A]
66. Geiger M, Cranor L. Scrubbing stubborn data: an evaluation of counter-forensic privacy tools. *IEEE Secur Priv* 2006;4(5):16-25. [doi: [10.1109/MSP.2006.132](#)]
67. Gordon LA, Loeb MP, Lucyshyn W, Richardson R. Eleventh Annual CSI/FBI Computer Crime and Security Survey. Computer Security Institute. 2006.
68. Personal Health Information Protection Act Report HI-050047-1: A Physiotherapy and Rehabilitation Centre. Toronto, ON: Office of the Privacy Commissioner of Ontario. 2006. URL: http://www.ipc.on.ca/images/Findings/up-I_050047_1.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5OF67opQF]
69. Personal Health Information Protection Act Report HI-050044-1: A Psychologist Working for a School Board. Toronto, ON: Office of the Privacy Commissioner of Ontario. 2006. URL: http://www.ipc.on.ca/images/Findings/up-HI_050044_1_Report.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5OF7qHLuW]
70. Order HO-004. Toronto, ON: Office of the Privacy Commissioner of Ontario. 2007. URL: http://www.ipc.on.ca/images/Findings/up-3ho_004.pdf [accessed 2007 Sep 11] [WebCite Cache ID 5OFOzaj1O]

Abbreviations

PHI: personal health information

PI: personal information

submitted 25.05.07; peer-reviewed by P Abdel Malik, P Chhanabhai; comments to author 08.06.06; revised version received 12.08.07; accepted 06.09.07; published 30.09.07.

Please cite as:

El Emam K, Neri E, Jonker E

An Evaluation of Personal Health Information Remnants in Second-Hand Personal Computer Disk Drives

J Med Internet Res 2007;9(3):e24

URL: <http://www.jmir.org/2007/3/e24/>

doi: [10.2196/jmir.9.3.e24](#)

PMID:

© Khaled El Emam, Emilio Neri, Elizabeth Jonker. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>, 30.09.2007). Except where otherwise noted, articles published in the Journal of Medical Internet Research are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Original Paper

Effect of Emailed Messages on Return Use of a Nutrition Education Website and Subsequent Changes in Dietary Behavior

W Gill Woodall¹, PhD; David B Buller², PhD; Laura Saba³, PhD; Donald Zimmerman⁴, PhD, BS; Emily Waters³; Joan M Hines⁵, MPH; Gary R Cutter⁶, PhD; Randall Starling¹, PhD

¹University of New Mexico, Albuquerque, NM, USA

²Klein Buendel, Inc, Golden, CO, USA

³University of Colorado at Denver and Health Sciences Center, Denver, CO, USA

⁴Colorado State University, Fort Collins, CO, USA

⁵The Children's Hospital, Denver, CO, USA

⁶University of Alabama, Birmingham, AL, USA

Corresponding Author:

David B Buller, PhD

Klein Buendel, Inc

1667 Cole Blvd, Suite 225

Golden, CO 80401

USA

Phone: +1 303 565 4340

Fax: +1 303 565 4320

Email: dbuller@kleinbuendel.com

Abstract

Background: At-risk populations can be reached with Web-based disease prevention and behavior change programs. However, such eHealth applications on the Internet need to generate return usage to be effective. Limited evidence is available on how continued usage can be encouraged.

Objective: This analysis tested whether routine email notification about a nutrition education website promoted more use of the website.

Methods: Adults from six rural counties in Colorado and New Mexico, United States (n = 755) participating in a randomized trial and assigned to the intervention group (n = 380) received, over a period of 4 months, email messages alerting them to updates on the website, along with hyperlinks to new content. Update alerts were sent approximately every 5 weeks (each participant received up to 4 messages). Log-ons to the website were the primary outcome for this analysis.

Results: A total of 23.5% (86/366) of the participants responded to at least one email, and 51.2% (44/86) of these participants responded to half of the email messages by logging on to the website. Significantly more log-ons occurred on email notification days compared to all other days (OR = 3.71, 95% CI = 2.72-5.06). More log-ons also occurred just after the notification but declined each day thereafter (OR = 0.97, 95% CI = 0.96-0.98 one day further from mass email). Non-Hispanics (OR = 0.46, 95% CI = 0.26-0.84), older participants (OR = 1.04, 95% CI = 1.04-1.06), and those using the Internet most recently (OR = 0.62, 95% CI = 0.51-0.77) were more likely to log on. Responders to the messages had a more positive change in fruit and vegetable intake (mean change = +1.69) than nonresponders (+0.05), as measured with a food frequency assessment (adjusted Spearman partial correlation coefficient = 0.14, *P* = .049). Compared to nonresponders, responders were more likely to be non-Hispanic (*P* = .01), older (*P* < .001), and had used the Internet more recently (*P* < .001).

Conclusions: Messages sent by email appeared to promote a modest short-lived increase in use of a disease prevention website by some adults. Those who responded to the messages by logging on to the website may have been influenced to improve their diet.

(*J Med Internet Res* 2007;9(3):e27) doi:[10.2196/jmir.9.3.e27](https://doi.org/10.2196/jmir.9.3.e27)

KEYWORDS

Internet; diet; adult; behavior change; nonresponse; electronic mail; attrition; usage

Introduction

Evidence is mounting that Internet technology can be used to reach at-risk populations with Web-based disease prevention and behavior change programs. However, attrition and low use of these websites appears to be a common occurrence when they are used in community settings where participants are not “required” to use them. Surveys indicate that people are using the Internet to obtain health information; however, some immediate acute need, most commonly a real or potential health problem, seems to be the primary driver rather than the need for preventive information [1]. Unfortunately, strategies to improve exposure to health communication have not been investigated as often as strategies to influence health behavior [2].

It is a well-established principle in media effects research that audience activity determines the impact of media messages [3-5]. Exposure to media messages is fundamental to achieving effects, whether it be for news coverage, entertainment content, or a health communication program [2,4,6-8]. Audience members are selective in their choice of media and the content within media [9], and selective exposure has been demonstrated in new media such as the Internet [10].

In the case of the Internet, the choice of websites is essentially endless. The study website must compete with a large number of health-related websites for participants’ attention, not to mention the more than 350 million listings on Google related to nutrition, fruits and vegetables, cancer prevention, and diet—the topics of the current project—as of August 2007. In this situation, it may be that effective methods of increasing website use involve reminding adults about the website, providing hyperlinks that make it easy to visit with a single mouse click, and showing users what new content has been recently published to the website.

Beyond the conceptual issues surrounding why and how often individuals use the Internet, attrition and low website use threaten the internal validity of medical Internet research [11]. The experimental comparison in a randomized trial can be compromised when a substantial number of study participants do not visit the website being tested. Researchers often must rely on association between website usage, which is self-selected, and disease prevention outcomes rather than the randomly assigned comparison between treatment groups. In these cases, the possibility that a third variable(s) determines both greater website use and changes in prevention behavior undermines conclusions that websites are effective. It is also possible that website use is so infrequent in a trial that no association can be detected between treatment group or website use and health behaviors. In both scenarios, a potentially effective website is considered a failure without giving it a chance to demonstrate its full benefits.

In a recent community-based trial evaluating a nutrition education program, periodic messages were sent by email to notify participants of new content posted to the website and to remind them to visit the site. This paper reports the analyses of the effect of these email notifications on promoting visits to the website.

Methods

Context

The data reported in this paper were collected as part of a randomized trial to evaluate the effects of a nutrition education website promoting the consumption of fruits and vegetables. The main trial procedures have been reported elsewhere [12]. Relevant portions of these procedures are summarized here, and the content and methods for delivering the email messages to study participants and assessing usage of the website are described in full.

Population and Sample

A sample of 755 adults in a six-county rural region spanning the border between Colorado and New Mexico in the United States participated in the main trial. Participants had to live at least 6 months or more in the region at pretest, be 18 years of age or older, and consent to participate. Data from the 380 participants randomized to the intervention group were analyzed for this paper.

Participants were recruited by 12 community outreach trainers from June 2002 to January 2004. Community outreach trainers were local adults recruited and trained to locate potential participants, introduce the study, obtain informed consent, conduct the pretest interview (using computer-assisted interviewing software), and assign participants to study condition (using a randomization computer program written by the project biostatistician). Community outreach trainers were blind to condition during recruitment, enrollment, and pretesting. When necessary, they also could provide basic computer and Internet skills training to participants. Community outreach trainers provided each participant with a unique user identification code (ID) to access the website, and this ID was used to track website usage. Project staff observed community outreach trainers during the trial and confirmed that they performed project tasks as trained [12].

Trial Procedures

The main trial involved a randomized pretest-posttest controlled design, with individuals randomized to receive immediate access to the study website (intervention group) or to receive delayed access to it after the posttest (control group). At 4 months post-randomization, participants completed the posttest by telephone with professional telephone interviewers blind to the study condition.

Nutrition Education Website: “5 a Day, the Rio Grande Way”

The “5 a Day, the Rio Grande Way” website contained content on the health benefits of fruits and vegetables; instructions for buying, storing, and preparing fruits and vegetables; and ways to increase fruits and vegetables in the family diet, particularly with children. The following were also provided on the website: information on gardening, recipes, and fruits and vegetables in season; a community directory that included organizations that sold fruits and vegetables or supplies for gardening in the six-county region; and a listing of health resources on the Internet related to fruit and vegetable consumption. The selection

and organization of content areas on the website were based on social cognitive theory [13] and diffusion of innovations model [14] and were guided by expert advice and results from focus groups on nutrition and health information [15], evaluation of alternative message formats [16], and usability testing on the initial website structure with local residents [17].

Content of the Reminder Messages

A total of 12 different messages were created and sent to participants announcing new content on the website. In each message, universal record locators (URLs) were provided, linking to the areas in the website containing the new content. These email notifications were designed to alert participants that new, important, and relevant information had been added to the site. Participants were invited to take a look at this new content, which, in some cases, was seasonal in nature. The messages were personally addressed to each participant, and the participant's username and password were provided in the messages. The participants were reminded to contact the community outreach trainer who recruited them if they had any questions about the website or problems accessing it.

Procedures for Delivering Reminder Messages

Approximately every 5 weeks, participants were sent a reminder message by email. The messages were sent to the email addresses provided by participants at trial enrollment. If emails were returned as undeliverable, the community outreach trainers attempted to contact participants and obtain an updated email address. The messages were sent from the central offices of the study. Corresponding to the number of updates on the "5 a Day, the Rio Grande Way" website during the 4-month observation period, most participants received 2 (35.7%, 136/380), 3 (32.1%, 122/380), or 4 (27.1%, 103/380) messages highlighting the updates. There were five participants who received only one message. (There were 14 participants not eligible to receive any messages because they were part of a pilot test of recruitment strategies.)

Measures

Website use was recorded with a custom-made program running on the Web server. Each usage session was identified and linked to a unique participant using their ID. Number of log-ons to the website, total time spent on the website, and visits to various website features were recorded and totaled for the 4-month period. Also, the date each reminder message was emailed was recorded, and the number of reminder messages each participant was eligible to receive within the 4-month (120-day) period was determined. In the present analysis, log-ons to the website were analyzed because they indicated that a participant had visited the website. Also, log-ons displayed a high correlation with time spent using the website [12].

Change in dietary behavior was measured with a validated food frequency measure of fruit and vegetable intake conducted at pretest and posttest [18]; an item was added to assess consumption of common regional foods such as red chile, green chile, and salsa. Responses were converted to servings per day following Thompson et al [18]. Additional pretest and posttest measures assessed attitudes toward cancer and its prevention, support from others to eat fruits and vegetables, involvement

in purchasing and preparing foods, experience using computers and the Internet, perceived self-efficacy for using the Internet, and social and demographic characteristics as potential predictors of response to the email reminders. Participants reported on their frequency of exhibiting 13 eating and meal preparation practices. Two composites with adequate reliability were identified using a principal components analysis with a varimax rotation: eating habits (pretest Cronbach alpha = .75, posttest Cronbach alpha = .69) and access habits (pretest Cronbach alpha = .66, posttest Cronbach alpha = .58).

Statistical Analysis

Three analyses were performed in this paper to explore different aspects of the effect of mass emails on use of the experimental website and the response to it: (1) effect of email messages on log-on rates, (2) characteristics of participants who responded to the email messages by logging on, and (3) association of response to the email messages (by logging on) with dietary change. These analyses are explained in greater detail below.

Effect of Email Messages on Log-Ons to the Website

To explore whether users were prompted to log on to the website after a mass email was sent, a generalized linear mixed model for a binomial outcome (a form of generalized estimating equations) with a logit link function [19] using PROC GENMOD in SAS was utilized (SAS version 9.1, SAS Institute Inc, Cary, NC, USA). In this model, subjects had an observation for each day between their pretest and posttest. The outcome of interest is a 0/1 variable, where 1 indicates that the subject logged on to the website that day. The independent variable of interest was a 0/1 indicator for whether or not a mass email was sent that day. The generalized linear model allows for a first-order autoregressive covariance structure to account for within-subject correlations where time (in days) is measured from the date of the participant's pretest.

Characteristics of Participants Who Responded to Email Messages

In addition to determining the effect that email alerts had on log-ons to the website, it was important to explore what participant characteristics were associated with response. Subjects were classified as responders if at any point during their study period (4 months) they visited the "5 a Day, the Rio Grande Way" website within 5 days of a mass email being sent. Associations between variables with this classification of response were examined using a logistic regression; odds ratios were derived to provide a measure of effect size. A multivariate model for predicting response status was created using a forward stepwise selection procedure. Results from this multivariate model are reported as factor effects using a Wald chi-square to test for significance.

Association of Response to Email Messages With Dietary Change

The final analysis explored the relationship of the main outcome of the original study — change in fruit and vegetable consumption — to the response to mass emails. Using the same classification of responder as in the previous section, change in fruit and vegetable intake from pretest to posttest was correlated

with responder status. In a separate analysis, the proportion of mass emails that the participant responded to was correlated with change in fruit and vegetable intake. All correlations utilized a nonparametric Spearman partial correlation, adjusting for variables in the final multivariate model predicting response status. A nonparametric test was chosen due to the non-normal distribution (excessive skewing) of the change in fruit and vegetable consumption.

Results

Description of Sample

Table 1 presents the demographic characteristics of participants randomized to the intervention group. They were predominantly female and married, ranged in age from 18 to 86 years, lived in multi-person households with children, and were long-term residents of the study region at enrollment. Nearly two thirds self-identified as Hispanic, and almost one tenth were Native Americans.

Hispanics had used the Internet less frequently (0 times = 20.9%, 50/239; 1-4 times = 10.5%, 50/239; 5-10 times = 5.0%, 12/239; > 10 times = 63.6%, 152/239) than non-Hispanics (0 times = 10.8%, 13/120; 1-4 times = 5.8%, 7/120; 5-10 times = 2.5%, 3/120; > 10 times = 80.8%, 97/120; $n = 379$, $\chi^2_3 = 11.19$, $P = .01$). However, among current users of the Internet, there was no significant difference by Hispanic origin in time spent on the Internet in a typical day (Hispanic mean = 7.0 hours, SD = 10.6; non-Hispanic mean = 7.2 hours, SD = 9.1; $t_{1,245} = 0.17$, $P = .87$), but fewer Hispanics had an email account (73.7%, 112/152 current Hispanic users; 89%, 88/99 non-Hispanic users; $\chi^2_1 = 8.56$, $n = 251$, $P = .003$).

About one third of participants had a high school degree or less education, and another one third had a 2-year or 4-year college degree or postgraduate education. Most of the participants had prior experience using the Internet, but nearly one third had used it 10 times or less (Table 1). Two thirds reported currently using the Internet at enrollment.

Table 1. Demographic characteristics and Internet experience of participants in the intervention group of the randomized trial (n = 380)

Demographics	No.	%
Gender		
Male	47	12.4
Female	333	87.6
Age		
20 to 29	133	35.0
30 to 49	131	34.5
50 or older	104	27.3
Refused/missing	12	3.2
Hispanic origin		
Of Hispanic origin	246	64.7
Not of Hispanic origin	130	34.2
Refused/missing	4	1.1
Race		
American Indian / Alaska Native	36	9.5
Asian	3	0.8
Black	2	0.5
Native Hawaiian / Pacific Islander	1	0.3
White	133	35.0
None of these	173	45.5
Refused/missing	32	8.4
Education		
Eleventh grade or less	49	12.9
High school graduate / GED	82	21.6
Trade school or some college education	133	35.0
2-year or 4-year college degree	78	20.5
Postgraduate	38	10.0
Current marital status		
Married or living with someone	209	55.0
Widowed	22	5.8
Separated or divorced	54	14.2
Never been married	91	23.9
Refused/missing	4	1.1
Number of people in household (including subject)		
1	46	12.1
2	105	27.6
3	75	19.7
4	90	23.7
5 or more	64	16.9
Refused/missing	0	0.0
Number of minors in household		
0	157	41.3
1	77	20.3

Demographics	No.	%
2	96	25.3
3 or more	48	12.6
Refused/missing	2	0.5
Number of years subject resided in Upper Rio Grande Valley		
Less than 1	5	1.3
1 to 10	63	16.6
11 to 20	70	18.4
More than 20	232	61.1
Refused/missing	10	2.6
Number of times ever used Internet		
None	67	17.6
1 to 10	51	13.4
> 10	259	68.2
Don't know	3	0.8
Currently use Internet		
Yes	260	68.4
No	50	13.2
Never used	67	17.6
Refused/missing	3	0.8

Effect of Email Messages on Log-Ons to the Website

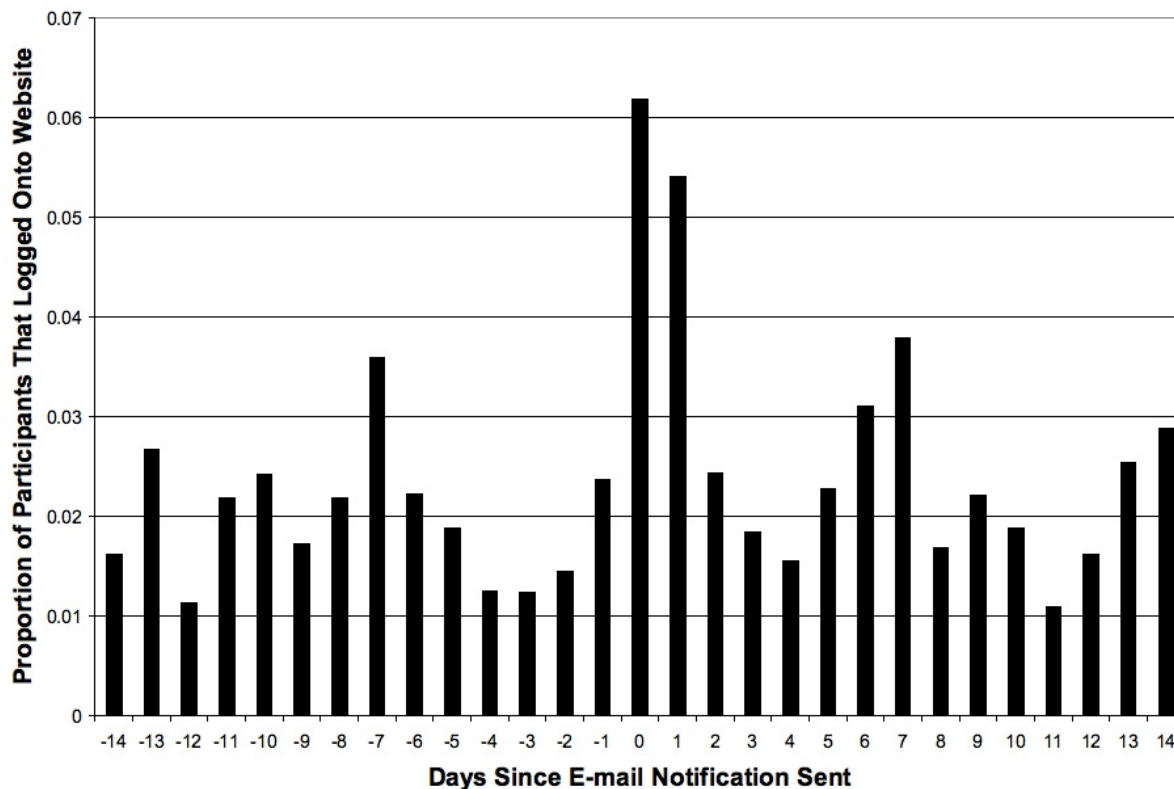
Overall, 23.5% (86/366) of participants responded to at least one of the email messages by logging on to the website within 5 days of when the message was sent. Only 6 participants responded to all of the messages sent to them. Of those who responded to at least one email, 51.2% (44/86) responded to half or more of the messages, while 48.8% (42/86) responded to fewer than half.

Participants were more likely to log on to the website on days when the messages were sent than on days when no email was sent (OR = 3.71, 95% CI = 2.72-5.06). This relationship remained evident when the analysis was expanded to include both the day the mass email was sent and the day after (OR = 4.11, 95% CI = 3.26-5.19).

However, [Figure 1](#) shows that the effect of the messages was short-lived. The number of log-ons to the website appeared to return to the frequency seen before sending the messages by 2 days after the messages were emailed. In fact, the odds ratio associated with a 1-day increase in time since last mass email is significantly less than 1 (OR = 0.97, 95% CI = 0.96-0.98), confirming this declining rate (OR = 0.85, 95% CI = 0.82-0.88, for the decline over 5 days after email message was sent).

Characteristics of Participants Who Responded to Email Messages

Several baseline characteristics (household size, number of children, education, race, Hispanic ethnicity, age, length of residence, gender), Internet experience (computer and Internet use, length and recency of Internet use, Internet self-efficacy, health information seeking self-efficacy, prior Internet training), and dietary patterns (readiness to change fruit and vegetable intake, diet responsibility, benefits of fruits and vegetables, dietary habits related to fruits and vegetables, comparison to peer intake of fruits and vegetables, family and friend support for eating fruits and vegetables) were tested for their association with visiting the website within 5 days of the email messages. The multivariate model predicting response contained three variables: (1) Hispanic ethnicity, (2) time since last use of the Internet (1 = within the past month, 2 = 1 to 5 months ago, 3 = 6 to 12 months ago, 4 = more than 12 months ago, 5 = never used Internet / not currently using Internet), and (3) age ([Table 2](#)). Non-Hispanic participants, older participants, and participants who has used the Internet more recently were more likely to log on to the website on the day of and the 4 days after the email messages were sent.

Figure 1. Average number of log-ons on 14 days prior to and after sending email messages to participants**Table 2.** Pretest characteristics significantly associated with responding to the email messages in final multivariate model

Characteristic	Odds Ratio *	95% CI	χ^2	P value
Hispanic ethnicity	0.46	0.26-0.84	6.49	.011
Amount of prior Internet use [†]	0.62	0.51-0.77	19.68	< .001
Older age	1.04	1.02-1.06	17.51	< .001

* Odds ratio indicates amount of change in response to email messages associated with being Hispanic (as opposed to non-Hispanic); one-unit increase on Internet use scale, and 1-year increase in age.

[†] Scale is 1 (within the past month), 2 (1 to 5 months ago), 3 (6 to 12 months ago), 4 (more than 12 months ago), and 5 (never used Internet / not currently using Internet); odds ratio indicates amount of change in response to email reminders associated with one-unit change on this scale of Internet use.

Association of Response to Email Messages With Dietary Change

Responding to the email messages was associated with dietary improvements. Participants who visited the website within 5 days after the message was sent reported larger increases in fruit and vegetable intake than those who did not (Spearman partial correlation coefficient = 0.14, $P = .049$; adjusted for ethnicity, age, and time since last use of the Internet). However, the partial correlation between the proportion of messages to which participants responded and increased daily intake of fruit and vegetables was nonsignificant ($\rho = 0.11$, $P = .12$).

Discussion

Sending messages that highlighted new content published to a nutrition education website appeared to prompt about one third of adults in a community setting assigned to use the website to log on and visit it, as it did in hospital contexts [20,21].

However, the effect of each message was short-lived, with number of log-ons returning to pre-message levels within 3 to 5 days. Moreover, many participants responded to less than half of the messages. Thus, this strategy is, at best, modestly successful at increasing exposure to a disease prevention website.

Selective Exposure to Disease Prevention Websites

Selective exposure to media arises because people have limited capacity to process messages. Attention is driven both by volitional processes under individuals' control and automatic cognitive orienting systems [22]. Theories such as the Cognitive Mediation Model hold that volitional selective exposure is driven by needs and motivations [6,8] (eg, by personal interest, a general surveillance motivation, or desire to gain information for future discussions with others). Email reminders may have appealed most to those individuals who were already interested in their health or health topics in general or who desired to

change their diet. Additional research is needed to explicate the motivations for using health websites.

Some topics or message formats in the emails and on the website may have been automatically attention-getting. Prior research showed that leads for online news stories that highlighted conflict and agony produced more selective exposure because danger-conveying signals or empathic sensitivities linked to emotional displays are inherently attention-getting [10]. One such “hardwired” frame in the current study might be personalizing messages by addressing the emails to users by name. Likewise, online messages containing attributes such as animation, ad position, and novelty may elicit an involuntary orienting response [23,24], and the promotion of “new” content in the email reminder may have automatically captured user attention. In future studies, various frames should be tested for improving selective exposure as well as multimedia features such as animation.

Selective Avoidance of Disease Prevention Websites

Selectivity means that people also can choose to avoid messages that do not interest them [4]. It may be that some participants had low commitment to the trial, low interest in their health and diet, and/or did not find the website attractive and chose to not visit it even when reminded by email. We were not able to detect whether individuals opened the email messages, so we could not directly measure avoidance. Software is now available that can detect when email messages containing hypertext links are opened [25], and it could be used to explore selective avoidance of online content.

Selective Exposure and Effectiveness of a Disease Prevention Website

In the Cognitive Mediation Model, exposure provokes attention and elaboration or message involvement, which can determine whether the message is effective. Consistent with this perspective, adults who responded to the email messages reported greater pretest-posttest increases in daily fruit and vegetable consumption than adults who did not respond. Internet websites have produced improvements in diet and diet-related behaviors in previous trials [20,26-31], and it appears that strategies that achieve greater exposure to such a website improve their influence [32]. It is notable that the website use and dietary improvements occurred in a multi-ethnic group of adults in rural community environments. Prior studies enrolled mostly white users, and only three were conducted in community environments [30,33,34].

It was surprising that participants who responded to more email reminders did not report larger improvements in fruit and vegetable intake. The motivation to return to the website may have been linked to the same motivations that led to change in dietary behavior (ie, interest in health, plans to change diet), but once there, responders may not have always processed the information effectively. Some processing tendencies for online content (eg, scanning rather than reading carefully) can interfere with learning and may offset gains from increased elaboration of information in the networked structure common on websites [35] (although networked structure may improve understanding [36]). Thus, media exposure is, at best, an imperfect predictor

of media effects. Some participants who spent more time on the website may have learned more information and skills and were more persuaded to change their diets, but others may have spent more time because they were not learning enough or were experiencing difficulty altering their diets. Like all media, understanding how the Internet is effective requires understanding both the determinants of exposure and processing of the information once exposure happens.

The Promise and Pitfalls of Email for Online Disease Prevention Communication

Email reminders may be an especially attractive means of promoting return use of a website. Email is an online function, used by nearly all Internet users [37]. While used equally by all age cohorts, email is the predominant online feature used by older adults [37] relative to other online functions, perhaps explaining why older participants responded most to the email reminders. Email reminders may be especially cost-effective in rural areas, where users are spread over large geographic areas.

However, the multicultural and rural character of the population in the study area made the likelihood of use of an Internet-based intervention somewhat less than might be expected in other areas of the United States. Along with being somewhat behind the general curve of Internet adoption [38], rural residents are frequently underserved in terms of health and medical information and stand to gain from the use of disease prevention websites. They may have more interest in online health information than urban users [38], although in the current sample, Hispanics were less frequent users of the Internet. The challenge remains to meet their needs with effectively designed Internet communication. These findings should generalize to suburban and urban users, too, because email is very popular in all regions of the country [38].

Still, ominous trends in the world of email may limit or diminish its effectiveness. The increasing amount of email advertising, or spam, has created dissatisfaction among Internet users, particularly women, older adults, and more experienced users [39]. Filtering software and other user strategies present barriers to delivering reminders by email. Distraction can interfere with message receipt and effectiveness [4,40], and many Internet users already report that spam makes it difficult to find email messages they desire to read in their inbox, particularly in personal email accounts [39]. Email reminders may be more effective in permission-based marketing circumstances [41,42] where the website provider has an existing relationship with the users (eg, a health insurance provider) and users expect to receive it (but even this may not be acceptable to all users [39]). Internet users also may be experiencing email overload, and the email reminders simply added to the stress of an ever increasing demand to be accessible 24/7 through this medium [43]. Finally, email may be less effective as instant and text messaging and other communication media appear and users become more selective among online media (a phenomenon already witnessed in the young [44]).

Limitations to the Study

There were limitations in this study. Visits to the website during the days immediately following the email reminders were

assumed to be in response to them because we could not detect whether the email messages were opened. The pattern of an abrupt increase in log-ons on the day of transmission certainly is suggestive that most log-ons were in response to these messages. The dietary outcome measure has been validated in previous research, reducing the likelihood of a social desirability bias or demand effect. Still, people who responded to the email messages were self-selected and could have been predisposed to alter their diet or better able to do so. The analysis of change in fruit and vegetable intake is limited by the lower than desired follow-up rate for the survey (62% overall); follow-up was higher among older, married, more educated, white, non-Hispanic participants born outside the region who were living in smaller households with children and who had lived

for a shorter time in the region, but it did not differ by treatment group [12].

Conclusion

It is essential that health providers effectively position their websites to attract use from the intended audience. "Pushing" the content to users registered on a website through routine email messages may be one way of prompting its use. Further research is needed to determine how best to create and present these email messages and to understand the motivations that underlie selective exposure to health websites and the information processing that takes place as users read email messages and navigate websites. Exposure to Internet health communication is a necessary first step to demonstrate its effectiveness in experimental trials and to make providers' investment in this new communication technology pay off.

Acknowledgments

This project was supported by a grant from the National Cancer Institute (CA81864). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Cancer Institute.

Conflicts of Interest

None declared.

References

1. Fox S. Health information online. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/PPF/r/156/report_display.asp [accessed 2007 Aug 14] [WebCite Cache ID 5Ikjukf4g]
2. Slater MD. Operationalizing and analyzing exposure: the foundation of media effects research. *Journalism Mass Commun* 2004;81:168-183.
3. Hawkins RP, Pingree S. Activity in the effects of television on children. In: Bryant J, Zillman D, editors. *Perspectives on Media Effects*. Hillsdale, NJ: Lawrence Erlbaum and Associates; 1986:233-250.
4. Kim J, Rubin AM. The variable of audience activity on media effects. *Commun Res* 1997;24(2):107-135. [doi: [10.1177/009365097024002001](https://doi.org/10.1177/009365097024002001)]
5. Woodall WG. Information-processing theory and television news. In: Robinson JP, Levy MR, editors. *The Main Source: Learning from Television News*. Beverly Hills, CA: Sage; 1986:133-158.
6. Beaudoin CE, Thorson E. Testing the cognitive mediation model: the roles of news reliance and three gratifications sought. *Commun Res* 2004;31(4):446-471. [doi: [10.1177/0093650204266098](https://doi.org/10.1177/0093650204266098)]
7. Eveland WP. The cognitive mediation model of learning from the news. *Commun Res* 2001;28(5):571-601. [doi: [10.1177/009365001028005001](https://doi.org/10.1177/009365001028005001)]
8. Eveland WP, Shah DV, Kwak N. Assessing causality in the cognitive mediation model: a panel study of motivations, information processing, and learning during campaign 2000. *Commun Res* 2003;30(4):359-386. [doi: [10.1177/0093650203253369](https://doi.org/10.1177/0093650203253369)]
9. Zillman D, Bryant J. *Select Exposure to Communication*. Hillsdale, NJ: Lawrence Earlbaum Associates; 1985.
10. Zillman D, Chen L, Knobloch S, Callison C. Effects of lead framing on selective exposure to internet news reports. *Commun Res* 2004;31(1):58-81. [doi: [10.1177/0093650203260201](https://doi.org/10.1177/0093650203260201)]
11. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11 [FREE Full text] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)]
12. Buller DB, Woodall WG, Zimmerman D. Randomized trial on the 5 a Day, the Rio Grande Way website, a web-based program to improve fruit and vegetable consumption. *J Health Commun* (forthcoming).
13. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall; 1986.
14. Rogers EM. *Diffusion of Innovations*. 5th edition. New York: Free Press; 2003.
15. Buller DB, Woodall WG, Zimmerman DE, Heimendinger J, Rogers EM, Slater MD, et al. Formative research activities to provide Web-based nutrition education to adults in the Upper Rio Grande Valley. *Fam Community Health* 2001 Oct;24(3):1-12. [Medline: [11563940](https://pubmed.ncbi.nlm.nih.gov/11563940/)]

16. Slater MD, Buller DB, Waters E, Archibeque M, Leblanc M. A test of conversational and testimonial messages versus didactic presentations of nutrition information. *J Nutr Educ Behav* 2003;35(5):255-259. [Medline: [14521825](#)] [doi: [10.1016/S1499-4046\(06\)60056-0](#)]
17. Zimmerman DE, Akerelrea CA, Buller DB, Hau B, Leblanc M. Integrating usability testing into the development of a 5 a day nutrition Web site for at-risk populations in the American Southwest. *J Health Psychol* 2003;8(1):119-134. [doi: [10.1177/1359105303008001448](#)]
18. Thompson FE, Subar AF, Smith AF, Midthune D, Radimer KL, Kahle LL, et al. Fruit and vegetable assessment: performance of 2 new short instruments and a food frequency questionnaire. *J Am Diet Assoc* 2002 Dec;102(12):1764-1772. [Medline: [12487538](#)] [doi: [10.1016/S0002-8223\(02\)90379-2](#)]
19. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986 Mar;42(1):121-130. [Medline: [3719049](#)] [doi: [10.2307/2531248](#)]
20. Tate DF, Wing RR, Winnett RA. Using Internet technology to deliver a behavioral weight loss program. *JAMA* 2001 Mar 7;285(9):1172-1177 [FREE Full text] [Medline: [11231746](#)] [doi: [10.1001/jama.285.9.1172](#)]
21. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *JAMA* 2003 Apr 9;289(14):1833-1836 [FREE Full text] [Medline: [12684363](#)] [doi: [10.1001/jama.289.14.1833](#)]
22. Lang A. The limited capacity model of mediated message processing. *J Commun* 2000;50(1):46-67. [doi: [10.1111/j.1460-2466.2000.tb02833.x](#)]
23. Lang A, Borse J, Wise K, David P. Captured by the world wide web: orienting to structural and content features of computer-presented information. *Commun Res* 2002;29(3):215-245. [doi: [10.1177/0093650202029003001](#)]
24. Diao F, Sundar SS. Orienting response and memory for web advertisements: Exploring effects of pop-up window and animation. *Commun Res* 2004;31(5):537-567. [doi: [10.1177/0093650204267932](#)]
25. LISTSERV Maestro. 2005. L-Soft. URL: <http://www.lsoft.com/products/listserv.asp> [accessed 2007 Aug 14] [WebCite Cache ID 5IkjpXVoX]
26. Celio AA, Winzelberg AJ, Wilfley DE, Eppstein-herald D, Springer EA, Dev P, et al. Reducing risk factors for eating disorders: comparison of an Internet- and a classroom-delivered psychoeducational program. *J Consult Clin Psychol* 2000 Aug;68(4):650-657. [Medline: [10965640](#)] [doi: [10.1037/0022-006X.68.4.650](#)]
27. Onema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. *Ann Behav Med* 2005 Feb;29(1):54-63. [Medline: [15677301](#)] [doi: [10.1207/s15324796abm2901_8](#)]
28. Onema A, Brug J, Lechner L. Web-based tailored nutrition education: results of a randomized controlled trial. *Health Educ Res* 2001 Dec;16(6):647-660 [FREE Full text] [Medline: [11780705](#)] [doi: [10.1093/her/16.6.647](#)]
29. Papadaki A, Scott JA. The Mediterranean eating in Scotland experience project: evaluation of an Internet-based intervention promoting the Mediterranean diet. *Br J Nutr* 2005 Aug;94(2):290-298. [Medline: [16115365](#)] [doi: [10.1079/BJN20051476](#)]
30. White MA, Martin PD, Newton RL, Walden HM, York-crowe EE, Gordon ST, et al. Mediators of weight loss in a family-based intervention presented over the internet. *Obes Res* 2004 Jul;12(7):1050-1059 [FREE Full text] [Medline: [15292468](#)]
31. Winzelberg AJ, Eppstein D, Eldredge KL, Wilfley D, Dasmahapatra R, Dev P, et al. Effectiveness of an Internet-based program for reducing risk factors for eating disorders. *J Consult Clin Psychol* 2000 Apr;68(2):346-350. [Medline: [10780136](#)] [doi: [10.1037/0022-006X.68.2.346](#)]
32. Williamson DA, Martin PD, White MA, Newton R, Walden H, York-crowe E, et al. Efficacy of an internet-based behavioral weight loss program for overweight adolescent African-American girls. *Eat Weight Disord* 2005 Sep;10(3):193-203. [Medline: [16277142](#)]
33. Harvey-Berino J, Pintauro S, Buzzell P, Digiulio M, Casey Gold B, Moldovan C, et al. Does using the Internet facilitate the maintenance of weight loss? *Int J Obes Relat Metab Disord* 2002 Sep;26(9):1254-1260 [FREE Full text] [Medline: [12187404](#)] [doi: [10.1038/sj.ijo.0802051](#)]
34. Womble LG, Wadden TA, MCGuckin BG, Sargent SL, Rothman RA, Krauthamer-Ewing ES. A randomized controlled trial of a commercial internet weight loss program. *Obes Res* 2004 Jun;12(6):1011-1018 [FREE Full text] [Medline: [15229342](#)]
35. Eveland WP, Dunwoody S. An investigation of elaboration and selective scanning as mediators of learning from the web versus print. *J Broadcast Electron* 2002;46(1):34-53. [doi: [10.1207/s15506878jobem4601_3](#)]
36. Eveland WP, Cortese J, Park H, Dunwoody S. How website organization influences free recall, factual knowledge, and knowledge structure density. *Hum Commun Res* 2004;30(2):208-233. [doi: [10.1111/j.1468-2958.2004.tb00731.x](#)]
37. Fox S, Madden M. Generations online. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/PPF/r/170/report_display.asp [accessed 2007 Aug 14] [WebCite Cache ID 5Ikjnyrcy]
38. Bell P, Reddy P, Rainie L. Rural areas and the Internet: rural American's Internet use has grown, but they continue to lag behind others. Washington, DC: Pew Internet & American Life Project; 2004. URL: http://www.pewinternet.org/PPF/r/73/press_release.asp [accessed 2007 Aug 14]
39. Fallows D. Spam: how it is hurting email and degrading life on the Internet. Washington, DC: Pew Internet & American Life Project; 2003. URL: http://www.pewinternet.org/report_display.asp?r=102 [accessed 2007 Aug 14] [WebCite Cache ID 5IkFN30I]

40. Buller DB. The effects of distraction during persuasion: A meta-analytic review. *Commun Monogr* 1986;53:91-114.
41. Godin S. *Permission Marketing*. New York: Simon & Schuster; 2003.
42. Yater T. Customer may I? Getting the OK to sell. *Infoworld* 2001;23:23-48.
43. Thomas GF, King CL, Baroni B, Cook L, Keitelman M, Miller S, et al. Reconceptualizing e-mail overload. *J Bus Tech Commun* 2006;20(3):252-287.
44. Lev-Ram M. Ohmigod, teens are so over email!. *CNN Money*. 2006 Jul 26. URL: <http://money.cnn.com/2006/07/26/technology/thirdscreen0726.biz2/index.htm> [accessed 2006 Aug 17] [[WebCite Cache ID 51kjkL94l](#)]

submitted 01.06.06; peer-reviewed by M Campbell, S Kirk; comments to author 07.08.06; revised version received 03.08.07; accepted 12.08.07; published 30.09.07.

Please cite as:

Woodall WG, Buller DB, Saba L, Zimmerman D, Waters E, Hines JM, Cutter GR, Starling R
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
URL: <http://www.jmir.org/2007/3/e27/>
doi: [10.2196/jmir.9.3.e27](https://doi.org/10.2196/jmir.9.3.e27)
PMID:

© W Gill Woodall, David B Buller, Laura Saba, Donald Zimmerman, Joan M Hines, Gary R Cutter, Randall Starling. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>, 30.09.07). Except where otherwise noted, articles published in the *Journal of Medical Internet Research* are distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, including full bibliographic details and the URL (see "please cite as" above), and this statement is included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>