Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 13 (2011), Issue 3 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoints

Original Papers

Short Paper

Review

Viewpoint

Opportunities and Challenges of Cloud Computing to Improve Health Care Services

Alex Mu-Hsing Kuo[1], PhD

[1] School of Health Information Science, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Alex Mu-Hsing Kuo, PhD
School of Health Information Science
University of Victoria
PO Box 3050 STN CSC
Victoria, BC, V8W 3P5
Canada
Phone: 1 250 4724300
Fax: 1 250 4724751
Email: akuo@uvic.ca

Abstract

Cloud computing is a new way of delivering computing resources and services. Many managers and experts believe that it can improve health care services, benefit health care research, and change the face of health information technology. However, as with any innovation, cloud computing should be rigorously evaluated before its widespread adoption. This paper discusses the concept and its current place in health care, and uses 4 aspects (management, technology, security, and legal) to evaluate the opportunities and challenges of this computing model. Strategic planning that could be used by a health organization to determine its direction, strategy, and resource allocation when it has decided to migrate from traditional to cloud-based health services is also discussed.

(J Med Internet Res 2011;13(3):e67) doi:10.2196/jmir.1867

KEYWORDS

Health care; electronic health record; cloud computing; bioinformatics; quality improvement

Introduction

Cloud computing refers to an on-demand, self-service Internet infrastructure that enables the user to access computing resources anytime from anywhere [1]. It is a new model of delivering computing resources, not a new technology. Examples of commonly used non-health care applications include Microsoft Hotmail and Google Docs, while some better known applications in health care include Microsoft HealthVault and Google Health platform (recently discontinued [2]). However, compared with conventional computing, this model provides three new advantages: massive computing resources available on demand, elimination of an up-front commitment by users, and payment for use on a short-term basis as needed [3]. Several articles, forums, and blogs have reported its applications in industry, business, transportation, education, and national security [4-7].

Health care, as with any other service operation, requires continuous and systematic innovation in order to remain cost effective, efficient, and timely, and to provide high-quality services. Many managers and experts predict that cloud computing can improve health care services, benefit health care research, and change the face of information technology (IT) [8-13]. For example, Schweitzer [10], Haughton [11], and Kabachinski [12] believe that cloud computing can reduce electronic health record (EHR) startup expenses, such as hardware, software, networking, personnel, and licensing fees, and therefore will encourage its adoption. Research by Rosenthal et al shows that the biomedical informatics community, especially consortiums that share data and applications, can take advantage of the new computing paradigm [13]. As indicated in the paper by Anderson et al, data-handling problems, complexity, and expensive or unavailable computational solutions to research problems are major issues in biomedical research data management and analysis [14]. Several informatics innovations have demonstrated that cloud computing has the potential to overcome these difficulties [15-21].

Despite the many benefits associated with cloud computing applications for health care, there are also several management, technology, security, and legal issues to be addressed. The aim of this paper is to discuss the concept of cloud computing, its current applications in health care, the challenges and opportunities, and how to implement strategic planning when the organization has decided to move to the new model of service.

Cloud Computing: A New Economic Computing Model

Cloud computing is still a developing paradigm, and its definition, attributes, and characteristics will evolve over time. Vaquero et al studied more than 20 definitions and tried to extract a consensus definition as well as a minimum definition containing the essential characteristics. Based on the study, they defined cloud computing as follows [22]:

> *Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically re-configured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customizedService-Level Agreements.*

From a service point of view, cloud computing includes 3 archetypal models: software, platform, and infrastructure [1,23-25].

(1) *Software as a service (SaaS):* The applications (eg, EHRs) are hosted by a cloud service provider and made available to customers over a network, typically the Internet.

(2) *Platform as a service (PaaS):* The development tools (eg, operation systems) are hosted in the cloud and accessed through a browser. With PaaS, developers can build Web applications without installing any tools on their computer, and then deploy those applications without any specialized administrative skills.

(3) *Infrastructure as a service (IaaS):* The cloud user outsources the equipment used to support operations, including storage, hardware, servers, and networking components. The provider owns the equipment and is responsible for housing, running, and maintaining it. The user typically pays on a per-use basis.

To deploy cloud computing, the US National Institute of Standards and Technology (NIST) listed 4 models (see Figure 1) [1,26]:

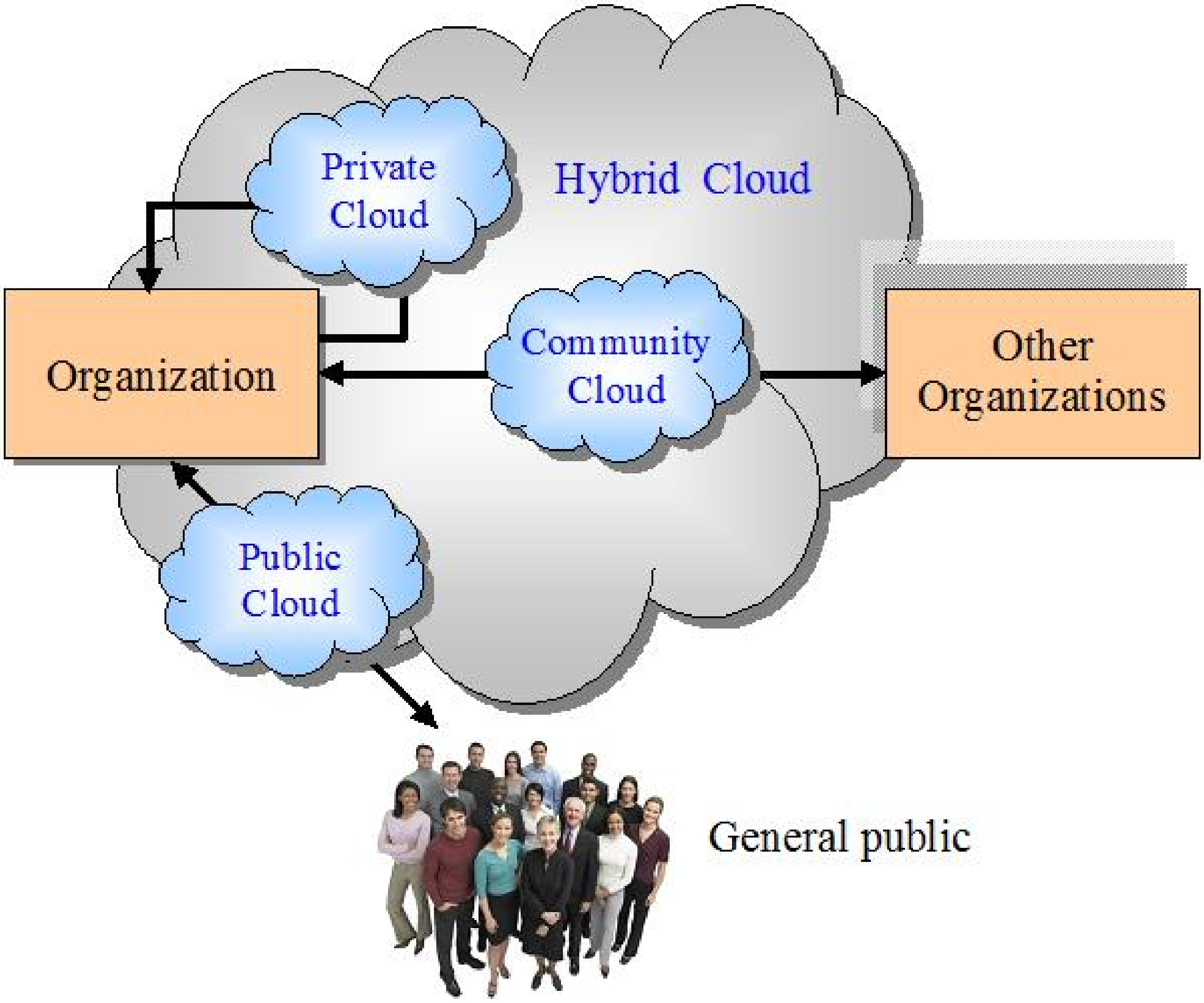
(1) *Public cloud:* A cloud service provider makes resources (applications and storage) available to the general public over the Internet on a pay-as-you-go basis. For example, the Amazon Elastic Compute Cloud (EC2) allows users to rent virtual computers on which to run their own applications. EC2 runs within Amazon's network infrastructure and data centers and allows customers to pay only for what they use with no minimum fee.

(2) *Private cloud:* A cloud infrastructure is operated solely for a single organization. In other words, the proprietary network or the data center supplies hosted services to a certain group of people. For example, Microsoft Azure enables customers to build the foundation for a private cloud infrastructure using Windows Server and System Center family of products with the Dynamic Data Center Toolkit.

(3) *Community cloud:* The cloud infrastructure is shared by several organizations with common concerns (eg, mission, security requirements, policy, and compliance considerations). For example, the Google GovCloud provides the Los Angeles City Council with a segregated data environment to store its applications and data that are accessible only to the city's agencies.

(4) *Hybrid cloud:* The cloud infrastructure comprises 2 or more clouds (private, public, or community). In this infrastructure, an organization provides and manages some resources within its own data center and has others provided externally. For example, IBM collaborates with Juniper Networks to provide a hybrid cloud infrastructure to enterprises to seamlessly extend their private clouds to remote servers in a secure public cloud [27].

Figure 1. The cloud computing deployment models.



Status and Adoption of Cloud Computing in Health Care

Many previous studies reported the potential benefits of cloud computing and proposed different models or frameworks in an attempt to improve health care service [28-35]. Among them, Rolim et al proposed a cloud-based system to automate the process of collecting patients' vital data via a network of sensors connected to legacy medical devices, and to deliver the data to a medical center's "cloud" for storage, processing, and distribution. The main benefits of the system are that it provides users with 7-days-a-week, real-time data collecting, eliminates manual collection work and the possibility of typing errors, and eases the deployment process [36]. Nkosi and Mekuria described a cloud computing protocol management system that provides multimedia sensor signal processing and security as a service to mobile devices. The system has relieved mobile devices from executing heavier multimedia and security algorithms in delivering mobile health services. This will improve the utilization of the ubiquitous mobile device for societal services and promote health service delivery to marginalized rural communities [37]. Rao et al reported a pervasive cloud initiative called Dhatri, which leveraged the power of cloud computing and wireless technologies to enable physicians to access patient health information at anytime from anywhere [38]. Koufi et al described a cloud-based prototype emergency medical system for the Greek National Health Service integrating the emergency system with personal health record systems to provide physicians with easy and immediate access to patient data from anywhere and via almost any computing device while containing costs [39].

Numerous of articles and resources also reported the successful application of cloud computing in bioinformatics research [15-21,40,41]. For example, Avila-Garcia et al proposed a framework based on the cloud computing concept for colorectal cancer imaging analysis and research for clinical use [18]. Bateman and Wood used Amazon's EC2 service with 100 nodes to assemble a full human genome with 140 million individual reads requiring alignment using a sequence search and alignment by hashing (SSAHA) algorithm [19]. Kudtarkar et al also used Amazon's EC2 to compute orthologous relationships for 245,323 genome-to-genome comparisons. The computation took just over 200 hours and cost US $8,000, approximately 40% less than expected [20]. Memom et al applied cloud computing to evaluate the impact of G-quadruplexes on Affymetrix arrays [21]. The Laboratory for Personalized

Medicine of the Center for Biomedical Informatics at Harvard Medical School took the benefits of cloud computing to develop genetic testing models that managed to manipulate enormous amounts of data in record time [41].

Besides academic researchers, many world-class software companies have heavily invested in the cloud, extending their new offerings for medical records services, such as Microsoft's HealthVault, Oracle's Exalogic Elastic Cloud, and Amazon Web Services (AWS), promising an explosion in the storage of personal health information online. Also, the use of health cloud computing is reported worldwide. For example, the AWS plays host to a collection of health care IT offerings, such as Salt Lake City-based Spearstone's health care data storage application, and DiskAgent uses Amazon Simple Storage Service (Amazon S3) as its scalable storage infrastructure [42].

The American Occupational Network is improving patient care by digitizing health records and updating its clinical processes using cloud-based software from IBM Business Partners MedTrak Systems. The company now can provide faster and more accurate billing to individuals and insurance companies, shortening the average time to create a bill from 7 days to less than 24 hours, and reducing medical transcription costs by 80% [43].

The US Department of Health & Human Services' Office of the National Coordinator for Health Information Technology recently chose Acumen Solutions' cloud-based customer relationship management and project management system for the selection and implementation of EHR systems across the United States. The software enables regional extension centers to manage interactions with medical providers related to the selection and implementation of an EHR system [44].

Telstra and the Royal Australian College of General Practitioners announced the signing of an agreement to work together to build an eHealth cloud. Telstra is one of the leading telecommunications providers in Australia; the College is the largest general practice representative body in Australia with more than 20,000 members and over 7000 in its National Rural Faculty. The eHealth cloud will host health care applications including clinical software, decision-support tools for diagnosis and management, care plans, referral tools, prescriptions, training, and other administrative and clinical services [45].

In Europe, a consortium including IBM, Sirrix AG security technologies, Portuguese energy and solution providers Energias de Portugal and EFACEC, San Raffaele Hospital (Italy), and several European academic and corporate research organizations contracted Trustworthy Clouds—a patient-centered home health care service—to remotely monitor, diagnose, and assist patients outside of a hospital setting. The complete lifecycle, from prescription to delivery to intake to reimbursement, will be stored in the cloud and will be accessible to patients, doctors, and pharmacy staff [46].

Health Cloud Computing Opportunities and Challenges

Recent research indicates that 75% of chief information officers reported that they will need and use cloud computing in the near future [47,48]. The forecast, conducted by Mark Beccue, suggested that the number of people subscribing to mobile cloud applications will rise from 71 million to nearly a billion by 2014 [49]. In health sectors, many organizations, managers, and experts believe that the cloud computing approach can also improve services and benefit research [8-13]. In addition, a report by the European Network and Information Security Agency (ENISA) stated that this new computing model is set to see massive global investment in many sectors, including health care [50]. The report also estimated that, by 2013, US $44 billion will be spent worldwide on cloud computing, potentially providing huge benefits to health care.

As with any innovation, cloud computing should be rigorously evaluated before its widespread adoption. Few research papers have systematically studied the impact of cloud computing on health care IT in terms of its opportunities and challenges. This study reviews the literature and evaluates the opportunities and challenges from the viewpoint of management, technology, security, and legality (see Table 1).

Table 1. Cloud computing opportunity and challenge summary

Aspects	Opportunities	Challenges
Management	Lower cost of new IT[a] infrastructure	Lack of trust by health care professionals
	Computing resources available on demand	Organizational inertia
	Payment of use on a short-term basis as needed	Loss of governance
		Uncertain provider’s compliance
Technology	Reduction of IT[a] maintenance burdens	Resource exhaustion issues
	Scalability and flexibility of infrastructure	Unpredictable performance
	Advantage for green computing	Data lock-in
		Data transfer bottlenecks
		Bugs in large-scale distributed cloud systems
Security	More resources available for data protection	Separation failure
	Replication of data in multiple locations increasing data security	Public management interface issues
	Dynamically scaled defensive resources strengthening resilience	Poor encryption key management
		Privilege abuse
Legal	Provider’s commitments to protect customer’s data and privacy	Data jurisdiction issues
	Development of guidelines and technologies to enable the construction of trusted platforms by not-for-profit organizations	Privacy issues
	Fostering of regulations by government for data and privacy protection	

[a] Information technology.

Management Aspect

Opportunity

The principle advantage of cloud computing is its low cost. For example, Amazon charges only US $0.1 per hour for 1.0-GHz × 86 instruction set architecture “slices” of EC2. Amazon S3 charges US $0.12 to $0.15 per gigabyte-month, with additional bandwidth charges of US $0.10 to $0.15 per gigabyte to move data into and out of AWS over the Internet [51]. An organization can easily get a cost-effective and on-premise IT solution through cloud computing without the need to purchase or evaluate hardware or software, or to hire internal IT staff to maintain and service in-house infrastructure [20,41,51]. The result is that the organization can focus on critical tasks without having to incur additional costs with regard to IT staffing and training.

Also, the cloud computing approach speeds deployment while maintaining vital flexibility (ie, rapid elasticity and ubiquitous access to health resources). This capability means that, as demand changes, hospitals and other health care providers do not need to adjust their infrastructures to accommodate the changes.

Challenges

The main challenges include lack of trust in data security and privacy by users, organizational inertia, loss of governance, and uncertain provider’s compliance.

Trust is at the heart of the resistance that many customers have to the cloud [52]. Concerns arise when their sensitive data and mission-critical applications move to a cloud computing paradigm where providers cannot guarantee the effectiveness of their security and privacy controls [53].

Cultural resistance (ie, organizational inertia) to share data and change traditional ways of working is a common management challenge to adopting cloud computing.

In some cases, a service level agreement may not offer a commitment to allow the client to audit its data. The loss of data governance could have a severe impact on a cloud user’s strategy and therefore on the capacity to meet its mission and goals.

Finally, if a provider cannot meet the requisite compliance norms (eg, applicable laws, regulations, standards, contracts, or policy changes), then a customer’s investment may be at risk. In some cases, certain customer services (eg, credit card transactions) cannot be used [54].

Technology Aspect

Opportunity

Smaller hospitals, medical practices, and laboratories typically do not have internal IT staff to maintain and service in-house infrastructure for mission-critical applications such as EHRs. Therefore, eliminating the new infrastructure cost and the IT maintenance burdens can remove many obstacles to EHR adoption [10,55]. For bigger health organizations, placing data storage or IT application needs in the hands of a cloud provider essentially shifts the IT management burden to a third-party provider. From an IT management's point of view, cloud computing can increase the scalability, flexibility, and cost effectiveness of infrastructure.

Also, cloud computing has advantages for so-called *green computing*—the more efficient use of computer resources to help the environment and promote energy saving. Usage of ready-made computing resources tailored to an organization's needs certainly helps it to reduce electricity expenses. While it saves on electricity, it also saves on resources required to cool off computers and other components. This reduces the emission of dangerous materials into the environment [56].

Challenges

Several technical challenges related to the use of cloud computing include resource exhaustion, unpredictability of performance, data lock-in, data transfer bottlenecks, and bugs in large-scale distributed cloud systems.

Low cost and computing resources available on demand are two key features of cloud computing. However, the market is becoming crowded with large providers. Because of high competition, many cloud providers overcommit computing resources (eg, central processing unit [CPU] allocation, storage space, applications) to attract customers. In order to maintain the profit, they cut corners in the value-delivery system. For example, they may limit access to the cloud resources, or use out-of-date hardware or software or deploy older CPU technology. Unfortunately, most cloud customers are unable to govern the virtual architecture, and the providers usually do not permit an audit by the customers. The result is variable leading to unpredictable performance in the service [57]. This difference between the customer's expectation and what the provider can really deliver presents a major technical challenge for the cloud customer to provide high-quality service to its own users.

Data lock-in is also an important challenge. In some cases, cloud users may have to move data or services to another provider or back to an in-house IT environment because the provider ceases business or service operations. For example, Google decided to discontinue its Google Health service on January 1, 2012. Users have a year to download their health data [2]. Unfortunately, most cloud infrastructures provide very little capability on data, application, and service interoperability [51]. This makes it difficult for a customer to migrate from one provider to another, or move data and services back to an in-house IT environment.

Some cloud users (eg, biomedical research laboratories) may need to frequently upload to or download very large amounts of data from the cloud. Application users may find that there is a data transfer bottleneck because of physical networking bandwidth limitation. Another specific technical risk is that of bugs in large-scale distributed cloud systems. When compared with in-house IT systems, the errors in these very large distributed infrastructures are more difficult to debug [51].

Security Aspect

Opportunity

Perhaps the strongest resistance to the adoption of cloud computing in health IT centers relates to data security [58]. Nevertheless, compared with locally housed data, this model is not necessarily less secure. In some cases, it typically improves security because cloud providers (eg, Microsoft, Google, Amazon) are able to devote huge resources to solving security issues that many customers cannot afford, in contrast to the destruction of many medical records and legal documents in the Japan 9.0 magnitude earthquake or the New Orleans Hurricane Katrina disaster.

All kinds of security measures, such as in hardware, software, human resources, and management costs, are cheaper when implemented on a large scale. Most cloud providers replicate users' data in multiple locations. This increases data redundancy and independence from system failure and provides a level of disaster recovery. In addition, a cloud provider always has the ability to dynamically reallocate security resources for filtering, traffic shaping, or encryption in order to increase support for defensive measures (eg, against distributed denial-of-service attacks). The ability to dynamically scale defensive resources on demand has obvious advantages for resilience [54].

Challenges

There are many data security risks in the use of IT, such as hacker attacks, network breaks, natural disasters, separation failure, public management interface, poor encryption key management, and privilege abuse. Specific risks to cloud computing are separation failure, public management interface, poor encryption key management, and privilege abuse.

Cloud computing is usually accessible to many different customers. If the provider fails to separate the resources, it could cause very serious security risks. For example, a customer requests to delete data stored in the virtual infrastructure; as with most operating systems, this may not result in true erasing of the data immediately. The data are still stored on the disk but are just not available [54]. In the multiple tenancies environment, hardware resources are reused by other customers. In this case, a third party could have access to another customer's "deleted" data. This presents a higher risk to the cloud customers than with dedicated hardware.

The public management interface is cloud computing's other Achilles' heel. As indicated in the ENISA's cloud computing risk summary [54]:

> *The customer management interfaces of public cloud providers are Internet accessible and mediate access to larger sets of resources (than traditional hosting providers) and therefore pose an increased risk especially when combined with remote access and Web browser vulnerabilities*

Strong encryption with key management is one of the core mechanisms that cloud computing systems use to guard against data loss and theft. However, a poor key management procedure may cause loss of encryption keys, disclosure of secret keys or passwords to malicious parties, or unauthorized use for authentication.

Lastly, as cloud use increases, employees may increasingly become targets for criminal organizations. If the malicious insider is a system administrator, then he or she could use his or her privileges to steal critical data.

Legal Aspect

Opportunity

Data and privacy protection are essential to building the customer trust needed for cloud computing to reach its full potential. If the providers adopt better and clearer policies and practices, users would be better able to assess the related risks they face. Fortunately, many main providers have commitments to develop best policies and practices to protect customers' data and privacy [59-61]. Besides providers' commitments to this protection, some organizations, such as the Cloud Security Alliance, have developed a comprehensive guide to deal with security and privacy issues [62]. The Trusted Computing Group (http://www.trustedcomputinggroup.org/), a not-for-profit organization, suggests a set of hardware and software technologies to enable the construction of trusted platforms. Governments also play a critical role by fostering widespread agreement regulations for both users and providers [63-66].

Challenges

The use of cloud computing presents many legal issues such as contract law, intellectual property rights, data jurisdiction, and privacy [67-71]. Among them, data jurisdiction and privacy issues are major concerns.

In the cloud, physical storages could be widely distributed across multiple jurisdictions, each of which may have different laws regarding data security, privacy, usage, and intellectual property [70,71]. For example, the US Health Insurance Portability and Accountability Act (HIPAA) [63] restricts companies from disclosing personal health data to nonaffiliated third parties, and the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (PATRIOT) Act [72] gives the US government the right to demand data if it declares conditions as being an emergency or necessary to homeland security. Similarly, the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) [64] limits the powers of organizations to collect, use, or disclose personal information in the course of commercial activities. However, a provider may, without notice to a user, move the user's information from jurisdiction to jurisdiction. Data in the cloud may have more than one legal location at the same time, with differing legal consequences.

Cloud computing is a shared resource and multitenancy environment for capacity, storage, and network. The privacy risk of this type of environment includes the failure of mechanisms for separating storage, memory, routing, and even reputation between different tenants of the shared infrastructure. The centralized storage and shared tenancy of physical storage space means the cloud users are at higher risk of disclosure of their sensitive data (eg, health records) to unwanted parties [54].

Poor breach notification is also an important privacy issue [73]. For example, the PIPEDA proposed a new requirement for organizations to report material data breaches to the Privacy Commissioner of Canada and to notify individuals where there is a risk of harm [64]. Unfortunately, the breach notification does not really protect a customer's privacy. A recent survey shows that consumers who have received data breach notifications within the past year are at a much greater risk for fraud than the typical consumer [74].

Cloud Computing Strategic Planning

When a health organization considers moving its service into the cloud, it needs strategic planning to examine the new model's benefits and risks, assess its capabilities to achieve the goal, and identify strategies designed for its implementation. Several references are available for establishing a cloud strategic plan. For example, Marks and Lozano [75] describe the cloud computing adoption life cycle method involving 9 stages to help users begin a cloud project. These are proof of concept/pilot project, strategy and roadmap, modeling and architecture, implementation planning, implementation, expansion, integration, collaboration, and maturity.

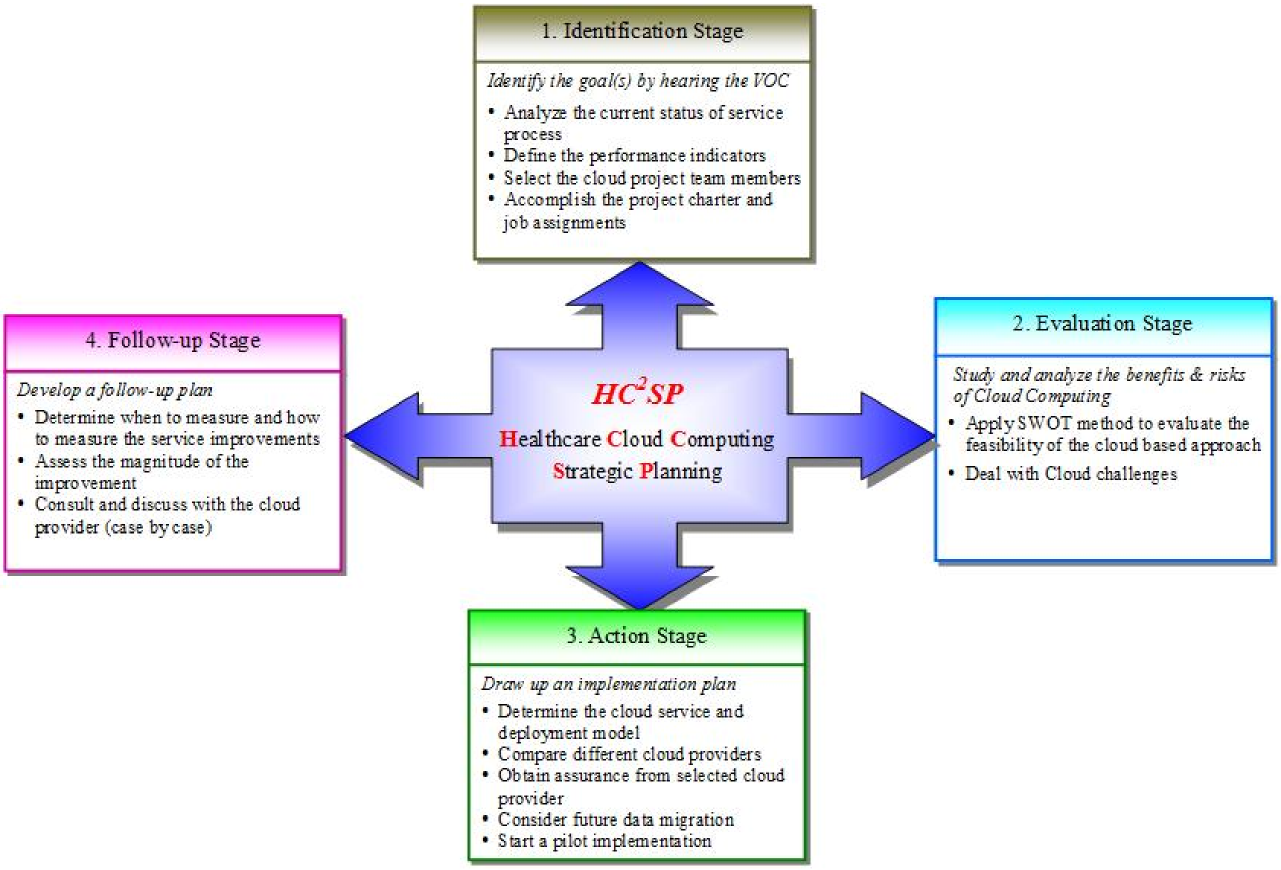
The Project Management Institute, a not-for-profit membership association for the project management profession, published a white paper on cloud computing that can be used as a reference for any cloud project manager. The paper provides 8 key steps for implementing cloud computing, as well as 2 case studies that support the method [76].

Stanoevska-Slabeva et al [77] also provide practical guidelines for moving traditional IT infrastructure toward clouds: initial analysis of demand and readiness for cloud computing, strategic decision to introduce cloud computing, pilot implementation, internal interconnection, inclusion of external resources, and continuous monitoring and evaluation.

The US Federal Health IT Strategic Plan [78], released in June 2008, can also be used for large government bodies to implement health cloud projects. The Plan charged the Office of the National Coordinator for Health Information Technology with a leadership role for the development and nationwide implementation of an interoperable health IT infrastructure to improve the quality and efficiency of health care. The strategic plan has 2 goals: patient-focused health care and population health, with 4 objectives under each goal. The objectives for both goals are privacy and security, interoperability, adoption, and collaborative governance. The Plan for achieving each goal is detailed through 43 strategies that describe the work needed to achieve each objective. Each strategy is associated with a milestone against which progress can be assessed and a set of illustrative actions to implement each strategy.

Besides the above-discussed strategic planning methods, this paper, based on a study [79], proposes a health care cloud computing strategic planning (HC^2SP) model that can be used by a health organization to determine its direction, strategy, and resource allocation to migrate from traditional health services to cloud-based services. The model includes 4 stages: identification, evaluation, action, and follow-up (see Figure 2).

Figure 2. Health care cloud computing strategic planning (HC^{2}SP) model (SWOT = strengths, weaknesses, opportunities, and threats; VOC = voice of customer).



Stage 1: Identification

In this HC^{2}SP model, the first stage is to analyze the current status of the health organization’s service process and identify the fundamental objective of service improvement by hearing the voice of the customer or the patients. The root causes analysis method can be applied to analyze the problems of the current service process. A typical hierarchy of causes would be expressed as follows [80]:

Problem #1: The process of patient admission to, or discharge from, hospital is too long. Why? There is too much unnecessary (duplicate) charting. Why? The paper-based charting system is inefficient. Why? There is lack of automated information systems such as EHR/EMR. Why? It involves a lot of up-front IT investments and maintenance.

The objective identification and its scope must be clarified so as to serve the end users (patients) more efficiently and effectively. In addition, the strategic planning team has to define health care service quality indicators and explain their purpose as well as the use of each indicator. This stage of the model provides the strategic planning team with a well-defined scope for the service problem being faced.

Stage 2: Evaluation

The second stage of the model is to evaluate the opportunities and challenges of adopting cloud computing. ENISA [54], the Cloud Security Alliance [62], and NIST [53] have developed comprehensive guides to evaluate the benefits and risks of adopting cloud computing. A potential user can also apply a strengths, weaknesses, opportunities, and threats (SWOT) analysis to evaluate the feasibility of the cloud-based approach [81].

Furthermore, the user needs to assess methods to handle the identified issues. Many references are available for this purpose (see Table 2 [51,53,62,71,82,83]). For example, Armbrust et al [51] report 10 major obstacles for cloud computing. Each obstacle is paired with opportunities (solutions), ranging from straightforward product development to major research projects. Buyya and Ranjan [82] discuss several cloud-federated management issues, such as data transfer bottlenecks, shared logging, and federation of distributed clusters. They also provide further references to handle the discussed issues. In addition, Kuo et al [83] propose an XML-based mediator to conquer data lock-in problems.

Table 2. Potential solutions to the cloud computing challenges

Challenges	Resources	Solution Summary
Management and technical issues	Armbrust et al [51]	Ten solutions to handle technical, policy, and business issues
	Buyya and Ranjan [82]	Further references to handle cloud-federated management issues
	Kuo et al [83]	XML-based mediator to handle data lock-in (interoperability) problems
Security and legal issues	Cloud Security Alliance [62]	Solutions to handle cloud governance and operation issues (12 domains)
	NIST[a] guidelines [53]	Precaution recommendations to deal with security and privacy issues
	Ward and Sipior [71]	Five strategies for handling data jurisdiction issues

[a] National Institute of Standards and Technology.

The Cloud Security Alliance [62] describes 12 domains of concerns for cloud computing. The domains are divided into 2 broad categories: governance and operations. Solution recommendations are also provided for each domain. The NIST Guidelines on Security and Privacy in Public Cloud Computing [53] names many key cloud security and privacy issues and the corresponding precaution recommendations for organizations to follow when planning or initiating a public cloud service outsourcing arrangement. Ward and Sipior [71] focus on jurisdiction issues. They recommend 5 strategies for cloud customers to deal with jurisdiction problems.

Stage 3: Action

After evaluating the new computing model, the organization will be able to determine whether to adopt the service or not. If the answer is yes, it needs to draw up an implementation plan. This paper proposes a 5-step plan as follows.

Step 1: Determine the Cloud Service and Deployment Model

As discussed above, cloud computing can refer to several different service types (SaaS, PaaS, and IaaS) and different deployment models (private, public, community, and hybrid cloud). Each service type or deployment model has its own benefits and risks [55]. Therefore, the key considerations in contracting for different types of services or deployment models should be different.

Step 2: Compare Different Cloud Providers

Choosing a proper cloud provider is the most important part of the implementation plan. Different providers may offer different service models, pricing schemes, audit procedures, and privacy and security policies. The organization has to compare different offerings. Also, it needs to evaluate the provider's reputation and performance before it signs a contract.

Step 3: Obtain Assurance From Selected Cloud Provider

The organization needs assurances that the selected provider will provide quality of service and follow sound privacy, security, and legal practices and regulations. The quality-of-service assurances include on-demand access, pay-per-use, rapid-elasticity, on-time troubleshooting support, and operational transparency [54]. The privacy and security assurances cover data confidentiality, integrity, availability, authenticity, authorization, and nonrepudiation. Also, the provider must assure that the data, including all of its backups, are stored only in geographic locations permitted by contract, service level agreement, and regulation.

Step 4: Consider Future Data Migration

The organization may have to move data and services to another provider or back to an in-house IT environment because the provider ceases business or service operations (eg, the recent discontinuation of Google Health [2]), has an unacceptable decrease in service quality, or has a contract dispute. Data portability must be considered up front as part of the plan [84].

Step 5: Start a Pilot Implementation

Many previous strategic planning methods suggest that an organization with no previous cloud experience start with a pilot implementation [75,77]. The pilot should be suitable for providing proof of the advantages of cloud computing for the organization.

Stage 4: Follow-up

The last stage is to deploy the cloud computing infrastructure and develop a follow-up plan. The plan indicates when to measure and how to measure the service improvements. Reasonable targets are established beforehand, and the results of the new services are measured against the specified targets or performance indicators to assess the magnitude of the improvement [80]. If the new service condition is not satisfied, the health organization needs to review what facts influence the objective achievement. If the main cause of unsatisfied service condition is from the cloud provider, the organization will consult and discuss with the provider to improve service or may consider moving data and services to another provider or back to its in-house IT environment.

Discussion and Conclusion

Cloud computing is a new model of computing that promises to provide more flexibility, less expense, and more efficiency in IT services to end users. It offers potential opportunities for improving EHR adoption, health care services, and research. However, as discussed above, there are still many challenges to fostering the new model in health care. Perhaps the strongest resistance to the adoption of cloud computing in health IT centers concerns data security and legal issues. Fortunately, many main providers (eg, Microsoft, Google, Amazon) have commitments to develop best policies and practices to secure customer's data and privacy [59-61]. Some not-for-profit organizations, such as the Cloud Security Alliance and the

Trusted Computing Group, have developed comprehensive guidelines, and hardware and software technologies to enable the construction of trustworthy cloud applications. Governments also foster regulations (eg, HIPAA [63] and PIPEDA [64]) to protect cloud users' data security and privacy. In addition, most legal issues involved in cloud computing usually can be resolved through contract evaluation or negotiations [10,54].

When a health organization considers moving its service into the cloud, it needs strategic planning to examine environmental factors such as staffing, budget, technologies, organizational culture, and government regulations that may affect it, assess its capabilities to achieve the goal, and identify strategies designed to move forward. This paper provides useful strategic planning references for potential users to start cloud projects. Also a new model called HC^2SP is proposed that could be applied by a health organization to determine its direction, strategy, and resource allocation to move to the cloud paradigm. The model includes 4 stages: identification, evaluation, action, and follow-up. At the first stage, the organization analyzes the current status of the service process and identifies the fundamental service objective. Stage 2 is to evaluate the opportunities and challenges of adopting cloud computing. By using the SWOT analysis, the organization can determine the internal strength and weakness factors as well as the external opportunity and threat factors of adopting the new model. Some potential solutions to handle cloud issues have been also provided. Then, in stage 3, the organization draws up a cloud computing implementation plan. The author suggests that this should include at least the following: determine the cloud service and deployment model, compare different cloud providers, obtain assurance from the selected cloud provider, consider future data migration, and start a pilot implementation. The last stage is to deploy the cloud computing infrastructure and develop a follow-up plan to measure the health care service improvements.

As the Chief Executive Officer of a cloud IT company commented [85]:

> *If you woke up this morning and read in The Wall Street Journal that, say, Overstock.com has stopped using UPS and FedEx and the U.S. mail, and had bought fleets of trucks and started leasing airport hubs and delivering products themselves, you would say they were out of their minds. Why is that much more insane than a health care company spending $2 billion a year on (traditional) information technology?*

Cloud computing presents a compelling opportunity for consumers of IT and producers of information services [86,87]. Gartner Research also found that cloud computing ranked as the top technical priority of chief information officers in 2011 [88]. However, adopting cloud computing is a complex process involving many factors. It needs rigorous evaluation before introducing the new computing model to an organization. This paper focuses on 4 aspects of evaluation and strategic planning, which will assist health organizations in determining whether (or how) to migrate from traditional to cloud-based health services.

Acknowledgments

The paper has benefited greatly from the reviewers' insightful comments and suggestions. The author would like to thank Professor Francis Lau and Mr George Fraser for proofreading the manuscript.

Conflicts of Interest

None declared

References

1. Mell P, Grance T. The NIST definition of cloud computing. Commun ACM 2010;53(6):50.
2. Brown A, Weihl B. Official Google Blog. 2011 Jun 24. An Update on Google Health and Google PowerMeter URL: http://googleblog.blogspot.com/2011/06/update-on-google-health-and-google.html [accessed 2011-08-05] [WebCite Cache ID 60gjQjyd9]
3. Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, et al. A view of cloud computing. Commun ACM 2010;53(4):50-58. [doi: 10.1145/1721654.1721672]
4. Technology firms and health care: heads in the cloud: digitising America's health records could be a huge business. Will it? The Economist (US) 2011;399(8727):63.
5. Li ZJ, Chen C, Wang K. Cloud computing for agent-based urban transportation systems. IEEE Intell Syst 2011;26(1):73-79.
6. Behrend TS, Wiebe EN, London JE, Johnson EC. Cloud computing adoption and usage in community colleges. Behav Inf Technol 2011;30(2):231-240. [doi: 10.1080/0144929X.2010.489118]
7. DarkGovernment. 2009 Jul 23. NSA Embraces Cloud Computing URL: http://www.darkgovernment.com/news/nsa-embraces-cloud-computing [accessed 2011-07-11] [WebCite Cache ID 60kNQBw5S]
8. Chatman C. How cloud computing is changing the face of health care information technology. J Health Care Compliance 2010 Jun;12(3):37-70.
9. Dudley JT, Pouliot Y, Chen R, Morgan AA, Butte AJ. Translational bioinformatics in the cloud: an affordable alternative. Genome Med 2010;2(8):51 [FREE Full text] [doi: 10.1186/gm172] [Medline: 20691073]

10. Schweitzer EJ. Reconciliation of the cloud computing model with US federal electronic health record regulations. J Am Med Inform Assoc 2011 Jul 4. [doi: 10.1136/amiajnl-2011-000162] [Medline: 21727204]
11. Haughton J. Year of the underdog: Cloud-based EHRs. Health Manag Technol 2011;32(1):9.
12. Kabachinski J. What's the forecast for cloud computing in healthcare? Biomed Instrum Technol 2011;45(2):146-150. [doi: 10.2345/0899-8205-45.2.146] [Medline: 21466336]
13. Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: a new business paradigm for biomedical information sharing. J Biomed Inform 2010 Apr;43(2):342-353. [doi: 10.1016/j.jbi.2009.08.014] [Medline: 19715773]
14. Anderson NR, Lee ES, Brockenbrough JS, Minie ME, Fuller S, Brinkley J, et al. Issues in biomedical research data management and analysis: needs and barriers. J Am Med Inform Assoc 2007;14(4):478-488 [FREE Full text] [doi: 10.1197/jamia.M2114] [Medline: 17460139]
15. Dudley JT, Butte AJ. In silico research in the era of cloud computing. Nat Biotechnol 2010 Nov;28(11):1181-1185. [doi: 10.1038/nbt1110-1181] [Medline: 21057489]
16. Wall DP, Kudtarkar P, Fusaro VA, Pivovarov R, Patil P, Tonellato PJ. Cloud computing for comparative genomics. BMC Bioinformatics 2010;11:259 [FREE Full text] [doi: 10.1186/1471-2105-11-259] [Medline: 20482786]
17. Schatz MC, Langmead B, Salzberg SL. Cloud computing and the DNA data race. Nat Biotechnol 2010 Jul;28(7):691-693. [doi: 10.1038/nbt0710-691] [Medline: 20622843]
18. Avila-Garcia MS, Trefethen AE, Brady M, Gleeson F, Goodman D. Lowering the barriers to cancer imaging. In: eScience 2008: IEEE 4th International Conference on eScience. New York, NY: IEEE; 2008 Presented at: The 4th IEEE International Conference on eScience; December 8-12, 2008; Indiana, USA. [doi: 10.1109/eScience.2008.33]
19. Bateman A, Wood M. Cloud computing. Bioinformatics 2009 Jun 15;25(12):1475 [FREE Full text] [doi: 10.1093/bioinformatics/btp274] [Medline: 19435745]
20. Kudtarkar P, Deluca TF, Fusaro VA, Tonellato PJ, Wall DP. Cost-effective cloud computing: a case study using the comparative genomics tool, roundup. Evol Bioinform Online 2010;6:197-203 [FREE Full text] [doi: 10.4137/EBO.S6259] [Medline: 21258651]
21. Memon FN, Owen AM, Sanchez-Graillet O, Upton GJ, Harrison AP. Identifying the impact of G-quadruplexes on Affymetrix 3' arrays using cloud computing. J Integr Bioinform 2010;7(2):111. [doi: 10.2390/biecoll-jib-2010-111] [Medline: 20134078]
22. Vaquero LM, Rodero-Merino L, Caceres J, Lindner M. A break in the clouds: towards a cloud definition. ACM SIGCOMM Comput Commun Rev 2008 Jan;39(1):50-55. [doi: 10.1145/1496091.1496100]
23. Iyer B, Henderson JC. Preparing for the future: understanding the seven capabilities of cloud computing. MIS Q Exec 2010;9(2):117-131.
24. Vouk MA. Cloud computing: issues, research and implementations. J Comput Inf Technol 2008;16(4):235-246. [doi: 10.2498/cit.1001391]
25. Han Y. On the clouds: a new way of computing. Inf Technol Libr 2010 June; 87-92 2010 Jun 1;29(2).
26. Cervone HF. An overview of virtual and cloud computing. OCLC Syst Serv 2010;26(3):162-165. [doi: 10.1108/10650751011073607]
27. IBM and Juniper Networks Solutions Brief. IBM Global Services. 2009. IBM and Juniper Networks: Delivering Solutions That Transform Your Networking Infrastructure URL: ftp://public.dhe.ibm.com/common/ssi/ecm/en/jns03002usen/JNS03002USEN.PDF [accessed 2011-07-25] [WebCite Cache ID 60ko8hVTK]
28. Sittig DF, Singh H. Eight rights of safe electronic health record use. JAMA 2009 Sep 9;302(10):1111-1113. [doi: 10.1001/jama.2009.1311] [Medline: 19738098]
29. Wang X, Tan Y. Application of cloud computing in the health information system. In: Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM). New York, NY: IEEE; 2010 Presented at: International Conference on Computer Application and System Modeling; October 22-24, 2010; Taiyuan, China. [doi: 10.1109/ICCASM.2010.5619051]
30. He C, Jin X, Zhao Z, Xiang T. A cloud computing solution for hospital information system. In: Proceedings of the 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS). New York, NY: IEEE; 2010 Presented at: IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS 2010); October 29-31, 2010; Xiamen, China. [doi: 10.1109/ICICISYS.2010.5658278]
31. Botts N, Thoms B, Noamani A, Horan TA. Cloud computing architectures for the underserved: public health cyberinfrastructures through a network of healthATMs. In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences (HICSS). New York, NY: IEEE; 2010 Presented at: The 43rd Hawaii International Conference on System Sciences; January 5-8 , 2010; Hawaii, USA. [doi: 10.1109/HICSS.2010.107]
32. Yang CT, Chen LT, Chou WL, Wang KC. Implementation of a medical image file accessing system on cloud computing. In: Proceedings of the 2010 IEEE 13th International Conference on Computational Science and Engineering (CSE). New York, NY: IEEE; 2010 Presented at: The 13th IEEE International Conference on Computational Science and Engineering; December 11-13, 2010; Hong Kong, China. [doi: 10.1109/CSE.2010.48]
33. Hoang DB, Chen L. Mobile cloud for assistive healthcare (MoCAsH). In: Proceedings of the. Asia-Pacific: :; 2010 Presented at: the IEEE Asia-Pacific Services Computing Conference; December 6-10 , 2010; Hangzhou, China. [doi: 10.1109/APSCC.2010.102]

34. Guo L, Chen F, Chen L, Tang X. The building of cloud computing environment for e-health. In: Proceedings of the 2010 International Conference on E-Health Networking, Digital Ecosystems and Technologies (EDT). New York, NY: IEEE; 2010 Presented at: The IEEE International Conference on E-Health Networking; July 1-3, 2010; Lyon, France. [doi: 10.1109/EDT.2010.5496512]
35. Alagoz F, Valdez AC, Wilkowska W, Ziefle M, Dorner S, Holzinger A. From cloud computing to mobile Internet, from user focus to culture and hedonism: the crucible of mobile health care and wellness applications. In: Proceedings of the 2010 5th International Conference on Pervasive Computing and Applications (ICPCA). New York, NY: IEEE; 2010 Presented at: The 5th International Conference on pervasive Computing and Applications (ICPCA); December 1-3, 2010; Maribor, Slovenia. [doi: 10.1109/ICPCA.2010.5704072]
36. Rolim CO, Koch FL, Westphall CB, Werner J, Fracalossi A, Salvador GS. A cloud computing solution for patient's data collection in health care institutions. 2010 Feb 10 Presented at: In: Proceedings of the 2nd International Conference on eHealth, Telemedicine, and Social Medicine; February 10-16, 2010; New York, NY: IEEE. [doi: 10.1109/eTELEMED.2010.19]
37. Nkosi MT, Mekuria F. Cloud computing for enhanced mobile health applications. In: Proceedings of the 2010 IEEE 2nd International Conference on Cloud Computing Technology and Science (CloudCom). New York, NY: IEEE; 2010 Presented at: The 2nd IEEE International Conference on Cloud Computing Technology and Science; Nov 30- Dec 3, 2010; Indianapolis, USA. [doi: 10.1109/CloudCom.2010.31]
38. Rao GSVRK, Sundararaman K, Parthasarathi J. Dhatri: a pervasive cloud initiative for primary healthcare services. In: Proceedings of the 2010 14th International Conference on Intelligence in Next Generation Networks (ICIN). New York, NY: IEEE; 2010 Presented at: The 14th IEEE International Conference on Intelligence in Next Generation Networks (ICIN); October 11-14, 2010; Berlin, Germany. [doi: 10.1109/ICIN.2010.5640918]
39. Koufi V, Malamateniou F, Vassilacopoulos G. Ubiquitous access to cloud emergency medical services. In: Proceedings of the 2010 10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB). New York, NY: IEEE; 2010 Presented at: The 10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB); November 3-5, 2010; Corfu, Greece. [doi: 10.1109/ITAB.2010.5687702]
40. Arrais JP, Oliveira JL. On the exploitation of cloud computing in bioinformatics. In: Proceedings of the 2010 10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB). New York, NY: IEEE; 2010 Presented at: The IEEE 10th International Conference on Information Technology and Applications in Biomedicine (ITAB); November 3-5, 2010; Corfu, Greece. [doi: 10.1109/ITAB.2010.5687653]
41. Amazon Web Services. 2011. AWS Case Study: Harvard Medical School URL: http://aws.amazon.com/solutions/case-studies/harvard/ [accessed 2011-07-20] [WebCite Cache ID 60kNYfCi1]
42. Business Wire. The Free Library. 2008. DiskAgent Launches New Remote Backup and Loss Protection Software as a Service Offering URL: http://www.thefreelibrary.com/DiskAgent(TM)+Launches+New+Remote+Backup+and+Loss+Protection+Software...-a0182194404 [accessed 2011-07-25] [WebCite Cache ID 1312946350828095]
43. Strukhoff R, O'Gara M, Moon N, Romanski P, White E. SYS-CON Media, Inc. 2009 Mar 20. Cloud Expo: Healthcare Clients Adopt Electronic Health Records with Cloud-Based Services URL: http://cloudcomputing.sys-con.com/node/886530 [accessed 2011-07-18] [WebCite Cache ID 60kNgwnG7]
44. Editorial Staff. HealthImaging.com. 2010 Feb 16. Acumen Nabs ONC Cloud Computing Contract URL: http://www.healthimaging.com/index.php?option=com_articles&view=article&id=20648:acumen-nabs-onc-cloud-computing-contract&division=hiit [accessed 2011-07-19] [WebCite Cache ID 60kNoT21d]
45. Korea IT Times. IT Times. 2010 Jul 20. Telstra Plans Launch of E-Health Cloud Services, Tip of the Iceberg for Opportunity URL: http://www.koreaittimes.com/story/9826/telstra-plans-launch-e-health-cloud-services-tip-iceberg-opportunity [accessed 2011-08-05] [WebCite Cache ID 60kNrsOz0]
46. IBM Press Room. IBM. 2010 Nov 22. European Union Consortium Launches Advanced Cloud Computing Project With Hospital and Smart Power Grid Provider URL: http://www-03.ibm.com/press/us/en/pressrelease/33067.wss [accessed 2011-08-05] [WebCite Cache ID 60kNuNDjt]
47. Danek J. Public Works Government Services Canada. 2009 Oct 6. Cloud Computing and the Canadian Environment URL: http://www.scribd.com/doc/20818613/Cloud-Computing-and-the-Canadian-Environment [accessed 2011-08-05] [WebCite Cache ID 60kNvZJD2]
48. Avery P. IT Business Edge. 2009 Aug 26. Research Indicates Increase in Cloud Computing URL: http://www.itbusinessedge.com/cm/community/kn/blog/research-indicates-increase-in-cloud-computing/?cs=35256 [accessed 2011-08-05] [WebCite Cache ID 60kNxLIhD]
49. Cherry S. Forecast for cloud computing: up, up, and away. IEEE Spectrum 2009 Oct;46(10):68.
50. Bannerman PL. Cloud Computing Adoption Risks: State of Play. In: Proceedings of the 17th Asia Pacific Software Engineering Conference Cloud Workshop. New York, NY: IEEE; 2010:10-16.

51. Armbrust M, Fox A, Griffith R, Joseph AD, Katz RH, Konwinski A, et al. EECS Department, UC Berkeley. 2009. Above the Clouds: A Berkeley View of Cloud Computing. Technical Report URL: http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf [accessed 2011-09-08] [WebCite Cache ID 61YCjBAXv]
52. Everett C. Cloud computing: a question of trust. Comput Fraud Secur 2009 Jun 10(6):5-7. [doi: 10.1016/S1361-3723(09)70071-5]
53. Jansen W, Grance T. National Institute of Standards and Technology, US Department of Commerce. 2011 Jan. Guidelines on Security and Privacy in Public Cloud Computing URL: http://csrc.nist.gov/publications/drafts/800-144/Draft-SP-800-144_cloud-computing.pdf [accessed 2011-09-08] [WebCite Cache ID 61YDDUYzT]
54. European Network and Information Security Agency. ENISA. 2009. Cloud Computing: Benefits, Risks and Recommendations for Information Security URL: http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment [accessed 2011-09-08] [WebCite Cache ID 61YDDp324]
55. Zhang R, Liu L. Security models and requirements for healthcare application clouds. In: Proceedings of the 2010 IEEE 3rd International Conference on Cloud Computing (CLOUD). New York, NY: IEEE; 2010 Presented at: The 3rd IEEE International Conference on Cloud; July 5-10, 2010; Miami, FL, USA. [doi: 10.1109/CLOUD.2010.62]
56. Baliga J, Ayre RWA, Hinton K, Tucker RS. Green cloud computing: balancing energy in processing, storage, and transport. Proc IEEE 2011;99(1):149-167. [doi: 10.1109/JPROC.2010.2060451]
57. Durkee D. Why cloud computing will never be free. Commun ACM 2010;53(5):70-69. [doi: 10.1145/1735223.1735243]
58. European Network and Information Security Agency. ENISA. 2009 Nov. An SME Perspective on Cloud Computing: Survey URL: http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-sme-survey/ [accessed 2011-07-23] [WebCite Cache ID 60kO1c8l6]
59. Microscoft Corp. 2010 Nov. Privacy in the Cloud: A Microsoft Perspective URL: http://www.microsoft.com/privacy/cloudcomputing.aspx [accessed 2011-09-07] [WebCite Cache ID 61WdQxiyd]
60. Google Privacy Center. Google. 2010 Oct 3. Privacy Policy URL: http://www.google.com/google-d-s/intl/en/privacy.html [accessed 2011-08-06] [WebCite Cache ID 60kO3Cyfl]
61. Amazon Web Services. 2008 Oct 01. AWS Privacy Notice URL: http://aws.amazon.com/privacy/ [accessed 2011-08-05] [WebCite Cache ID 60kO4Fxhu]
62. Cloud Security Alliance. 2009 Dec. Security Guidance for Critical Areas of Focus in Cloud Computing V2.1 URL: http://www.cloudsecurityalliance.org/csaguide.pdf [accessed 2011-07-25] [WebCite Cache ID 60kO5MSs7]
63. US Department of Health & Human Services. 1996. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy and Security Rules URL: http://www.hhs.gov/ocr/privacy/ [accessed 2011-07-26] [WebCite Cache ID 60kO6ZCde]
64. Minister of Justice, Canada. 2011 Jan 1. Personal Information Protection and Electronic Documents Act (PIPEDA) URL: http://laws.justice.gc.ca/PDF/Readability/P-8.6.pdf [accessed 2011-08-05] [WebCite Cache ID 60kO7iHR0]
65. European Commission. EuroPriSe: The European Privacy Seal for IT Products and IT-Based Services URL: https://www.european-privacy-seal.eu/ [accessed 2011-08-06] [WebCite Cache ID 60kO8f1ro]
66. United Nations. United Nations Commission on International Trade Law. 2010. UNCITRAL Legislative Guide on Secured Transactions URL: http://www.uncitral.org/pdf/english/texts/security-lg/e/09-82670_Ebook-Guide_09-04-10English.pdf [accessed 2011-08-05] [WebCite Cache ID 60lQfpIr4]
67. Pearson S. Taking account of privacy when designing cloud computing services. In: Proceedings of the ICSE Workshop on Software Engineering Challenges of Cloud Computing (CLOUD'09). New York, NY: IEEE; 2009 Presented at: the IEEE First international workshop on software engineering challenges for Cloud Computing (ICSE); May 16-24, 2009; Vancouver, BC, Canada. [doi: 10.1109/CLOUD.2009.5071532]
68. Svantesson D, Clarke R. Privacy and consumer risks in cloud computing. Comput Law Secur Rev 2010;26(4):391-397. [doi: 10.1016/j.clsr.2010.05.005]
69. Mather T, Kumaraswamy S, Latif S. Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance (Theory in Practice). Sebastopol, CA: O'Reilly Media, Inc.; 2009.
70. Kuner C. Data protection law and international jurisdiction on the Internet (part 1). Int J Law Inf Technol 2010;18(2):176-201. [doi: 10.1093/ijlit/eaq002]
71. Ward BT, Sipior JC. The Internet jurisdiction risk of cloud computing. Inf Syst Manag 2010;27(4):334-339. [doi: 10.1080/10580530.2010.514248]
72. Financial Crimes Enforcement Network, US Department of Treasury. FinCEN. Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act of 2001. no date URL: http://www.fincen.gov/statutes_regs/patriot/index.html [accessed 2011-07-13] [WebCite Cache ID 60kQGBI3G]
73. Cavoukian A. Information and Privacy Commissioner, Ontario, Canada. 2009 Nov. A Discussion Paper on Privacy Externalities, Security Breach Notification and the Role of Independent Oversight URL: http://www.ipc.on.ca/images/Resources/privacy_externalities.pdf [accessed 2011-07-13] [WebCite Cache ID 60kQ6Ist2]
74. Javelin Strategy & Research. 2011. Data Breach Notifications: Victims Face Four Times Higher Risk of Fraud URL: https://www.javelinstrategy.com/brochure-158 [accessed 2011-07-23] [WebCite Cache ID 60kOBDUIf]
75. Marks EA, Lozano B. Executive's Guide to Cloud Computing. Hoboken, NJ: Wiley; 2010.

76. White Paper. Project Management Institute (PMI). 2011. Cloud Computing: The New Strategic Weapon URL: http://www.pmi.org/~/media/PDF/Home/CloudComputing_FINAL.ashx [accessed 2011-07-23] [WebCite Cache ID 60kPwduxn]
77. Stanoevska-Slabeva K, Wozniak T, Hoyer V. Practical guidelines for evolving IT infrastructure towards grids and clouds. In: Stanoevska-Slabeva K, Wozniak T, Ristol S, editors. Stanoevska- Slabeva K, Wozniak T, Ristol S. editors. Grid and Cloud Computing: A Business Perspective on Technology and Applications. Berlin: Springer; 2010:225-243.
78. US Department of Health & Human Services, Office of the National Coordinator for Health Information Technology. The ONC-Coordinated Federal Health IT Strategic Plan: 2008-2012. Washington, DC: ONC-HIT; 2008.
79. Kuo AM, Borycki E, Kushniruk A, Lee TS. A healthcare Lean Six Sigma System for postanesthesia care unit workflow improvement. Qual Manag Health Care 2011;20(1):4-14. [doi: 10.1097/QMH.0b013e3182033791] [Medline: 21192203]
80. Lee TS, Kuo MH. Toyota A3 report: a tool for process improvement in healthcare. Stud Health Technol Inform 2009;143:235-240. [Medline: 19380942]
81. Marston S, Li Z, Bandyopadhyay S, Zhang J, Ghalsasi A. Cloud computing: the business perspective. Decis Support Syst 2011;51(1):176-189. [doi: 10.1016/j.dss.2010.12.006]
82. Buyya R, Ranjan R. Special section: Federated resource management in grid and cloud computing systems. Future Generation Comput Syst 2010;26(8):1189-1191. [doi: 10.1016/j.future.2010.06.003]
83. Kuo MH, Kushniruk AW, Borycki EM. Design and implementation of a health data interoperability mediator. Stud Health Technol Inform 2010;155:101-107. [Medline: 20543316]
84. Gagliardi F, Muscella S. Cloud computing: data confidentiality and interoperability challenges. In: Antonopoulos N, Gillam L, editors. Antonopoulos N, Gillam L. editors. Cloud Computing: Principles, Systems and Applications (Computer Communications and Networks). London: Springer; 2010:257-270.
85. Knowledge@Wharton. Wharton Business School, University of Pennsylvania. 2009 Apr 1. No Man Is an Island: The Promise of Cloud Computing URL: http://knowledge.wharton.upenn.edu/article.cfm?articleid=2190 [accessed 2011-07-15] [WebCite Cache ID 60kOCuNNL]
86. Creeger M. CTO roundtable: cloud computing. Commun ACM 2009;52(8):50-56. [doi: 10.1145/1536616.1536633]
87. Fox A. Computer science. Cloud computing: what's in it for me as a scientist? Science 2011 Jan 28;331(6016):406-407. [doi: 10.1126/science.1198981] [Medline: 21273473]
88. Gartner Newsroom. Gartner, Inc. 2011 Jan 21. Gartner Executive Programs Worldwide Survey of More Than 2,000 CIOs Identifies Cloud Computing as Top Technology Priority for CIOs in 2011 URL: http://www.gartner.com/it/page.jsp?id=1526414 [accessed 2011-07-14] [WebCite Cache ID 60kOEKBmh]

Abbreviations

Amazon S3: Amazon Simple Storage Service
AWS: Amazon Web Services
CPU: central processing unit
EC2: Elastic Compute Cloud
EHR: electronic health record
ENISA: European Network and Information Security Agency
HC2SP: health care cloud computing strategic planning
HIPAA: Health Insurance Portability and Accountability Act
IaaS: infrastructure as a service
IT: information technology
NIST: National Institute of Standards and Technology
PaaS: platform as a service
PATRIOT: Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism
PIPEDA: Personal Information Protection and Electronic Documents Act
SaaS: software as a service
SSAHA: sequence search and alignment by hashing
SWOT: strengths, weaknesses, opportunities, and threats

Edited by G Eysenbach; submitted 10.06.11; peer-reviewed by P Mork, E Kimura, C Reynolds, F Lai; comments to author 02.07.11; revised version received 10.08.11; accepted 25.08.11; published 21.09.11.

Please cite as:
Kuo AMH
Opportunities and Challenges of Cloud Computing to Improve Health Care Services
J Med Internet Res 2011;13(3):e67
URL: http://www.jmir.org/2011/3/e67/
doi:10.2196/jmir.1867
PMID:21937354

Viewpoint

Institutionalizing Telemedicine Applications: The Challenge of Legitimizing Decision-Making

Paolo Zanaboni[1,2], PhD; Emanuele Lettieri[2], PhD

[1]Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway
[2]Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Milano, Italy

Corresponding Author:
Paolo Zanaboni, PhD
Norwegian Centre for Integrated Care and Telemedicine
University Hospital of North Norway
P.O. Box 6060
Tromsø, 9038
Norway
Phone: 47 07766
Fax: 47 77754098
Email: paolo.zanaboni@telemed.no

Abstract

During the last decades a variety of telemedicine applications have been trialed worldwide. However, telemedicine is still an example of major potential benefits that have not been fully attained. Health care regulators are still debating why institutionalizing telemedicine applications on a large scale has been so difficult and why health care professionals are often averse or indifferent to telemedicine applications, thus preventing them from becoming part of everyday clinical routines. We believe that the lack of consolidated procedures for supporting decision making by health care regulators is a major weakness. We aim to further the current debate on how to legitimize decision making about the institutionalization of telemedicine applications on a large scale. We discuss (1) three main requirements— rationality, fairness, and efficiency—that should underpin decision making so that the relevant stakeholders perceive them as being legitimate, and (2) the domains and criteria for comparing and assessing telemedicine applications—benefits and sustainability. According to these requirements and criteria, we illustrate a possible reference process for legitimate decision making about which telemedicine applications to implement on a large scale. This process adopts the health care regulators' perspective and is made up of 2 subsequent stages, in which a preliminary proposal and then a full proposal are reviewed.

(J Med Internet Res 2011;13(3):e72) doi:10.2196/jmir.1669

KEYWORDS

Telemedicine; decision making; institutionalization; technology assessment; implementation; program sustainability

Introduction

Telemedicine is the use of communications networks for delivering health care services and medical education from one geographical location to another [1]. In this regard, telemedicine is generally acknowledged as a subset of eHealth applications, which are health care services and information delivered or enhanced through the Internet and related technologies [2].

Telemedicine is not new. Although a variety of telemedicine applications have been trialed worldwide over the last decades [3,4], telemedicine is still an example of major potential benefits that have not been fully attained [5]. Many applications that were adopted with great expectations quickly became failures and were subsequently abandoned [6]. Health care regulators are still debating why institutionalizing telemedicine applications on a large scale has been so difficult [5,7] and why health care professionals are often averse or indifferent to telemedicine applications, thus preventing them from becoming part of everyday clinical routines [8,9].

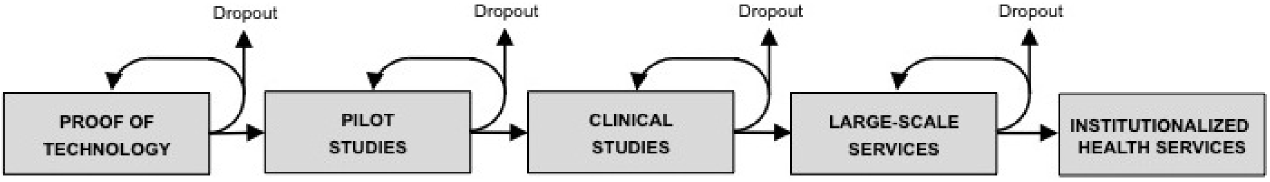
We believe that the reasons are many, but among others the lack of consolidated procedures for supporting decision making is a major weakness [10-13]. Recently the discipline of health technology assessment (HTA) has been acknowledged as the gold standard for the assessment of health care technology [14]. However, the major emphasis is on drugs, equipment, and medical devices; thus, telemedicine applications are far from having a reference assessment procedure [6]. Scholars of technology assessment are still debating whether telemedicine applications can be included within the scope of HTA [7]. Other studies investigated the problems [10] and conditions for successful implementation of telemedicine applications, identifying the most critical organizational [6,15] and contextual factors [16]. Within this debate, the normalization process theory developed by May et al [17] offers a theoretical lens through which to understand the implementation, embedding, and integration of complex health care interventions into everyday practice. Normalization process theory is today a well-established theory applied to several applications, including telemedicine [7]. However, this theory is more descriptive than prescriptive, and thus more pragmatic procedures for supporting decision making with respect to large-scale institutionalization of telemedicine applications are still missing [18]. Institutionalization is the last phase within the reference pathway that telemedicine applications follow to enter into clinical routine (Figure 1). Decisions about the institutionalization of telemedicine applications are usually taken by health care regulators, governments, and authorities in countries with a national health care service, such as the United Kingdom, France, Italy, and Norway.

In this Viewpoint, we consider this perspective, since we believe that in these cases the institutional actors (ie, health care regulators, governments, and authorities) are required by relevant stakeholders (eg, health care providers, patient groups, scientific societies, insurance companies, and technology suppliers) to legitimate their decision making about whether to institutionalize new health care technologies, such as telemedicine.

In particular, decision making about which telemedicine applications should be prioritized for large-scale implementation is tremendously complex [19]. In this context, we believe that a major issue for both academicians and practitioners should be to discuss how this peculiar decision-making exercise should be organized and which criteria should be adopted to support the selection of the most promising telemedicine applications.

We aim to further the debate on how to legitimize decision making about telemedicine applications by adding new elements to this debate. In particular, we discuss (1) the main requirements that should underpin this exercise for being perceived as legitimate by relevant stakeholders, and (2) the domains and criteria that should be taken into account for comparing and assessing different applications. Finally, we illustrate a possible reference decision-making process for telemedicine applications that might be taken as basis for both practical applications and academic discussions, and possibly extended to other eHealth applications.

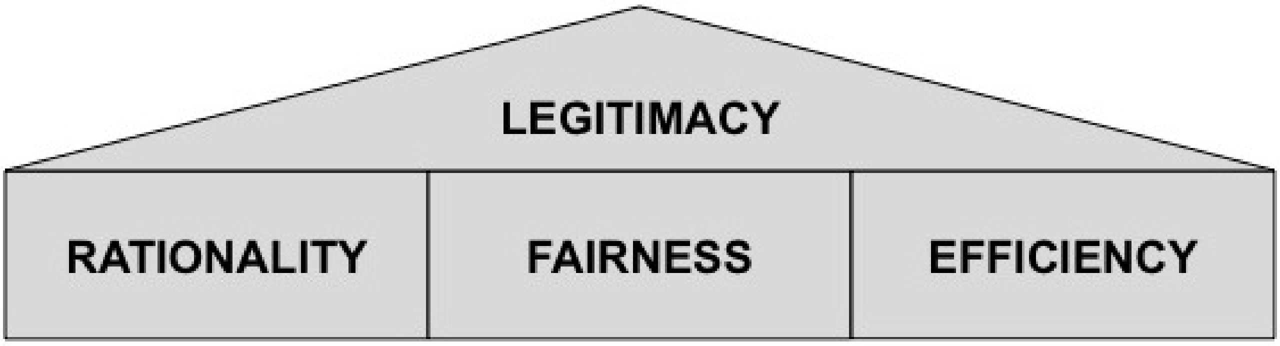
Figure 1. Implementation process of telemedicine applications.



Legitimizing Decision Making

Health care is a complex ecosystem of stakeholders [20]. Decision making about which telemedicine applications should be implemented into a large scale affects a variety of stakeholders and whether their goals are met. In this view, it is critical that they perceive this exercise as legitimate [7]. Legitimacy refers to “a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions” [21]. As a consequence, any organizational process should take into account the embracement of socially accepted techniques and procedures for being legitimate [22]. With respect to the peculiar case of the assessment of telemedicine applications, any recommendation should take into account a multiple-stakeholder perspective. Legitimization by the key stakeholders (ie, health care providers, patient groups, and technology suppliers) is gained when the decision-making exercise is respectful of their views and conforms to their perceptions of what is appropriate. Evidence from previous research suggests that 3 requirements are salient to gain stakeholders’ approval (Figure 2). First, decision making should be rational [23]. Second, decision making should be fair [24]. Third, decision making should be efficient [20]. In the following, we briefly discuss these requirements.

Figure 2. Requirements for legitimate decision making.



Rationality

Rationality of decision making is generally addressed through the adoption of multicriteria approaches. Decision making that covers a variety of criteria takes explicit account of the multifaceted nature of decisions under discussion and thus supports decision makers in exploring what really matters to key stakeholders [25]. All relevant goals and assessment criteria should be adequately captured and made accountable for decision makers [26], since shortcomings might arise if these goals and criteria have not been clarified [27]. Preferences between options should refer to an explicit set of objectives [28]. In this view, the selection of those telemedicine applications that appear to be the most promising should explicitly take into account the relevant and often contrasting goals. The peculiar nature of telemedicine applications and their expected performance should be assessed against multicriteria frameworks that facilitate a synthesis—and thus a conciliation—of different domains and perspectives. Previous contributions suggested a variety of criteria [11,13,29,30]; we believe that these criteria should be grouped into 2 main domains: (1) benefits-related criteria; and (2) sustainability-related criteria [31]. The last group is particularly critical for us. With “sustainability” we refer to the capability of use of a health care service to become routine and deliver high-quality and efficient care over time. A list of criteria that may be adopted for both the domains is discussed below.

Fairness

The second requirement for legitimacy is fairness. Decision making should be fair, accountable, and transparent [32]. This is particularly relevant when decision makers have contrasting views and goals [33]. The accountability for reasonableness (A4R) framework [32] offers a pragmatic solution to this issue, since it facilitates basing decisions on only those reasons that everybody will agree on and support [34]. The framework identifies 4 conditions that should be met by “fair-minded” decision makers: (1) relevance, (2) publicity, (3) revisability, and (4) enforcement [26]. Despite interest in the A4R framework, we must acknowledge that the literature offers little guidance on what a fair and legitimate process might look like in practice, and the process itself is still a “black box” [34]. The 4 conditions posited by the A4R framework are significant for the assessment of telemedicine applications. Relevance—required also for rationality—is necessary to address competing rationales and promote relevant stakeholders’ agreement. Publicity is critical for continuous learning. Proponents of telemedicine innovations must access, not just by demand, previous assessment exercises to learn about which criteria had been selected to compare benefits and sustainability in order to improve their further proposals. In particular, shortcomings of ongoing or concluded studies should be acknowledged and diffused to improve the reliability of further proposals [35]. In this view, adopting an appraisal-based process is also relevant for creating a trustworthy setting and facilitating learning and continuous improvement over time. Finally, decision makers must ensure that the above 3 conditions are met in order to win stakeholders’ legitimization.

Efficiency

We believe that efficiency is another key requirement for legitimate decision making [20], which should be timely and cost contained. The literature does not completely clarify how an efficient decision-making exercise should be organized with respect to technology assessment. Despite this limitation, observation of current practices by highly regarded agencies seems to acknowledge that a 2-stage decision-making exercise should be used to pursue timely and cost-contained decisions. A first example is the call for proposals within the Seventh Framework Program (2007–2013) of the European Commission. Applicants are required first to submit a limited outline proposal. All successful proposals are then invited to submit a full proposal [36]. Narrowing the view to health care, the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom adopted a similar process [37]. The process [38] is based on selection of new interventions first, followed by a structured appraisal of clinical effectiveness, cost effectiveness, and wider implications for the National Health Service as a whole. Emerging technologies that might require a NICE evaluation are identified by the National Horizon Scanning Centre at the University of Birmingham, England. First, a panel of experts prioritizes technologies for assessment. Second, the selected technologies are assessed and briefings are delivered to decision makers. Narrowing the view to telemedicine applications, the Scottish Telemedicine Initiative [39] invites organizations associated with the NHS in Scotland to submit assessment proposals to promote the implementation of telemedicine applications in Scotland. The procedure is also based on a 2-stage review process.

Based on these examples, we argue that decision-making exercise should possibly be organized through 2 sequential gates, with the first round aimed at selecting the most promising telemedicine applications and the second round aimed at selecting applications for large-scale implementation.

Criteria for Decision Making

As mentioned, rational decision making requires a multicriteria assessment. A variety of prescriptive multicriteria frameworks [11,13,29,30] converge on the identification of a list of assessment dimensions for telemedicine applications that should be reviewed by decision makers. Key dimensions include technical feasibility, legal and ethical issues, clinical effectiveness, economics, equity of access, acceptance by providers and patients, and organizational impacts. As previously mentioned, we think that decision makers would receive better support if the list of criteria covering these dimensions were grouped into 2 main domains: benefits and sustainability [31].

In this regard, we analyzed the literature to collect past contributions about benefits and sustainability measures. Our analysis was not intended to provide a systematic review of assessment criteria. Rather, it aimed at offering a first draft of how these 2 domains might be assessed through specific criteria. In the peculiar case of telemedicine applications, 3 main benefits-related criteria (Table 1) may be recognized: (1) clinical outcomes, (2) cost containment, and (3) access.

Table 1. Benefits-related criteria

Dimension	Measures	References
Clinical outcomes	Diagnostic accuracy	13,40–42
	Appropriateness of clinical decisions	42–45
	Therapeutic efficacy/effectiveness	13,41,46
	Timeliness of care	13,30,47,48
	Mortality	13,42,49–51
	Morbidity	13,50,52
	Disease-related measures	11,13,29,30,42
	Quality of life	11,13,42
	Hospitalizations	42,49,53,54
	Length of hospital stay	42,51,53
	In-clinic visits	42,45,49
	Emergency department accesses	45,53,55
Cost containment	Cost of the service	29,30,42,56–58
	Relative efficiency	11,29,30,58–60
Access	Geographic accessibility	29,30,42,45,61
	Availability	13,45,61,62
	Waiting lists	13,63
	Affordability	45,64

Clinical outcomes [30] consist of a wide set of measures that determine the effects of the implementation of telemedicine applications on patients' health status. They are diagnostic accuracy [13,40-42], appropriateness of clinical decisions [42-45], therapeutic efficacy/effectiveness [13,41,46], timeliness of care [13,30,47,48], mortality [13,42,49-51], morbidity [13,50,52], other efficacy/effectiveness or disease-related measures [11,13,29,30,42], quality of life [11,13,42], hospitalizations [42,49,53,54], length of hospital stay [42,51,53], in-clinic visits [42,45,49], and emergency department accesses [45,53,55].

Cost containment is intended to measure the value of resource use associated with an intervention [29], thus allowing an understanding of whether a telemedicine application is cost saving or cost effective. First, the cost of a telemedicine application has to be measured and compared with the alternative clinical practice [29,30,42,56-58]. Second, relative efficiency with other alternative health care strategies is critical for decision making when resources are limited [11,29,30,58-60].

Telemedicine applications are claimed to improve access to health care services, especially for people living in rural or remote areas where medical professionals and facilities are scarce or absent [29]. Access includes a set of specific dimensions describing the fit between the patient and the system, including geographical accessibility of patients [29,30,42,45,61], availability of health care resources and professionals [13,45,61,62], waiting lists for secondary care [13,63], and affordability of health care services [45,64].

A growing body of evidence suggests that decision makers should also take into account sustainability-related criteria when assessing health care technologies [26,31]. As mentioned, we refer to sustainability as the capability of a telemedicine application to deliver high-quality and efficient care over time. In this respect, sustainability is not a single concept, but is an umbrella of different dimensions. Five main dimensions (Table 2) may be recognized. They are (1) technological sustainability, (2) clinical sustainability, (3) organizational sustainability, (4) economic sustainability, and (5) contextual sustainability.

Table 2. Sustainability-related criteria

Dimension	Measures	References
Technological sustainability	Integration and interoperability	11,65,66
	Use over an extended period of time (proof of time)	30,58
Clinical sustainability	Malpractice, adverse events, and uncertainty in clinical practice	29,67,68
Organizational sustainability	Acceptance and satisfaction	13,29,30,42,69
	Staff, skills, learning, and training	30,70–73
	Leadership, communication, roles, and responsibilities	71–73
	Changes in organizational structure and work processes	42,70,72,73
	Collaboration, cooperation, partnership, and networks	71–73
Economic sustainability	Cost structure	11,13,42,56–58,60
	Total investment	13,57,58
	Level of use	13,56–58
	Costs for patients and caregivers	13,42,57–59
Contextual sustainability	Respect of ethical requirements	67,68
	Respect of legal requirements	30,67,68,74

The technological sustainability of a telemedicine application specifically addresses its integration and interoperability [11,65,66] with the existing technologies in a health care organization. Additionally, the telemedicine application has to be used over an extended period of time [30,58], thus attesting to the so-called *proof of time*.

Clinical sustainability addresses whether the clinical benefits will be maintained over time. Malpractice, including adverse events and uncertainty in clinical practice, should therefore be carefully considered [29,67,68].

Organizational sustainability refers to the readiness of organizations where the telemedicine application is implemented and potential resistance to change. The successful adoption of telemedicine applications depends on the acceptance and satisfaction of health care professionals, patients, caregivers, and other users [13,29,30,42,69]. Internal resistance to change can be also driven by issues related to staff, skills, and learning and training needs [30,70-73]. Leadership, roles, and responsibilities [71-73] have to be carefully defined and communicated among the health care professionals while introducing a new technology. Additionally, telemedicine applications imply the introduction of changes within the organizational structure and work processes [42,70,72,73]. Collaboration, cooperation, partnership, and networks [71-73] are also required beyond the borders of the organization, in relation to other health care providers and stakeholders.

Economic sustainability refers to the economic value generation derived from the use of telemedicine applications and the capability to maintain it over time. Economic sustainability is related to the cost structure [11,13,42,56-58,60] and to the total amount of capital investments [13,57,58]. Moreover, the economic impact depends on the level of use of the telemedicine application [13,56-58] and on the costs sustained by patients and caregivers [13,42,57-59].

Finally, contextual sustainability is necessary for the long-term use of a service. Contextual sustainability of telemedicine applications refers to the respect of ethical [67,68] and legal requirements [30,67,68,74].

We believe that clustering criteria in terms of benefits and sustainability offers better support to decision makers, since this approach makes clear to all key stakeholders both the potential value of a telemedicine application and the difficulties in pursuing high performance over time. Rationality and accountability of decision making will benefit from the transparent disclosure of decision makers' expectations in terms of benefits and sustainability, and this will facilitates stakeholders' acceptance of the final decision.

A Possible Reference Process for Decision Making

In this section, we illustrate a possible reference process for legitimate decision making about which telemedicine applications to implement on a large scale. As mentioned, we adopt the health care regulators' perspective and we address the need for key stakeholders, such as health care providers, patient groups, and technology suppliers, to legitimize this process. Our proposal of a reference process is based on the previous discussion. In this regard, we argue for a process that is made up of 2 subsequent stages, in which a preliminary proposal and then a full proposal are reviewed: this structure improves efficiency and leverages on the real examples that we have previously illustrated.

Stage 1: Assessment of Preliminary Proposals

The process would start with a request by an accredited subject (eg, hospitals, research centers) to decision makers to implement a large-scale telemedicine application. We assume a bottom-up process (and not top-down), since this perspective coheres to what happens in those countries that have a national health care system. The request derives from previous studies that enrolled a limited number of patients and providers, whose results justify the request to move to a larger scale. Subjects that were involved in these studies would submit a preliminary proposal.

The preliminary proposal is a document that provides decision makers with summary information about the telemedicine application under discussion. At this stage, all the preliminary proposals should be prioritized. Only those proposals with a satisfactory relevance score would be asked to submit the full proposal. This prioritization exercise would be conducted by an institutional committee (eg, members of governments or authorities) on the basis of the documentation provided. The possibility to establish an advisory panel as suggested in previous research [26] should be discussed. Preliminary proposals structure information in 2 main sections: (1) a description of the context into which the telemedicine application should be introduced, and (2) a description of the proposed telemedicine application and its benefits.

The description of the context (Table 3) provides information necessary to understand the nature of the problem that the telemedicine application aims to solve. For instance, a telemedicine application might address a clinical problem (eg, patients' health conditions affected by a chronic disease), an economic problem (eg, the containment of present expenditure for a specific health need), or a problem of accessibility to health care services (eg, the provision of secondary care to rural areas). Published data about the problem under consideration should support decision makers in understanding its relevance at both the national and the local level in order to contextualize the assessment exercise. Moreover, a telemedicine application is often introduced as an alternative to current clinical practice, commonly face to face [75]. In these cases, decision makers expect that the telemedicine application will produce at least the same clinical outcomes as the conventional practice before being adopted into routine [11]. Telemedicine applications could be either supplementary [45], offering additional activities, or alternative, changing the way of providing a service [76]. More rarely, they address needs still uncovered by traditional practice, thus allowing access to services not otherwise available [77].

Table 3. Preliminary proposal: description of the context

Dimension	Information
Nature of the problem	Description of the problem (clinical, economic, access) the telemedicine application aims to solve
	Description of data that support the relevance of the problem at the national or regional level
Current clinical practice	Description of current clinical practice (if existing) to which the telemedicine application is presented as an alternative practice (integrative or substitutive)

The description of the proposed telemedicine application (Table 4) aims to provide that information necessary to prioritize the implementation on a large scale. First, the proposal must clarify the type of service (eg, doctor-to-doctor consultation via telemedicine systems [78]), the description of the target patients (eg, patients affected by a specific chronic disease), the technology used for providing the service (eg, Internet), and the subjects who will be involved or interested in the service provision. Second, the proposal must clarify the benefits expected from the implementation of a large-scale telemedicine application. These can be measured through clinical outcomes, costs, and access.

Table 4. Preliminary proposal: description of the proposed telemedicine application

Dimension	Information
Description of the service	Characterization of the type of telemedicine application
	Characterization of patients
	Description of technologies used to provide the service
	Description of subjects involved in the service and their roles
Expected benefits	Impact on clinical outcomes
	Impact on costs
	Impact on access

The institutional committee, after having examined all the information provided, would make a decision based on the content of the proposal and its relevance to the health care system. Three different replies might arise: (1) the preliminary proposal is accepted due to positive feedback, (2) the preliminary proposal is accepted with reservation due to satisfactory feedback with some weaknesses to be addressed, (3) the preliminary proposal is rejected due to unsatisfactory feedback. Proposals in the first or second scenarios would be asked to be completed as a full proposal.

Stage 2: Assessment of Full Proposals

The full proposal provides detailed information about the telemedicine application. Full proposals would be assessed mainly by a board of experts, selected from a list of experts from different scientific fields (eg, medicine, sociology, health economics) who would be required to disclose their possible conflicts of interest. This board would produce recommendations for the institutional committee, responsible for the final selection.

The information collected into the full proposals is structured in 4 parts: (1) a description of the context in which the telemedicine application should be introduced, (2) a description of the proposed telemedicine application and its benefits, (3) documentation supporting the assessment, and (4) an assessment of the scientific basis and available evidence.

The context and the proposed telemedicine application in the full proposal contain the same information as in the preliminary proposal, but in more detail. Additional fields should require the specification of inclusion and exclusion criteria of enrollment, clinical protocol, technical requirements, expected and proper level of utilization, data and information exchanged, and methods for data collection and assessment.

The third part collects evidence that supports previous claims and figures. Studies, papers, and reports about the specific telemedicine application under consideration should be provided. The board of experts should verify the reliability of this evidence in terms of levels of evidence and strength of the recommendations [79].

The process would end with an assessment of the scientific basis. Both benefits-related criteria and sustainability-related criteria should be addressed as previously described in Table 1 and Table 2. Where information is available, the proponent should clarify the metrics used to measure the specific criteria, results, and references. The benefits-related criteria should be measured preferably with respect to a comparator (eg, control group of patients). Referring to the sustainability-related criteria, results and references should be classified with respect to the main stakeholders: patients and caregivers, health care professionals, hospitals, vendors and service centers, or the health care system.

At the end of the process, the board of expert should examine the context and the proposed telemedicine application and assess its benefits and sustainability with respect to a large-scale implementation. The institutional committee would then make the final decision. The implementation of a large-scale telemedicine application should finally require periodic monitoring to verify the correct implementation and level of use by health care providers, and the achievement of the expected benefits and its sustainability over time, and eventually to modify the clinical protocol or to interrupt the service.

Conclusions

Final Remarks

This Viewpoint provides new insights into the ongoing debate about the institutionalization of telemedicine applications. In particular, our discussion originates from the acknowledgement that a major weakness inhibiting the large-scale implementation of telemedicine applications is the lack of a legitimate reference process for decision making by health care regulators. Moving from this, we shed some light on (1) 3 main requirements that should characterize this process—that is, rationality, fairness, and efficiency, (2) some criteria for measuring the benefits and sustainability of a telemedicine application, and (3) what a legitimate process might look like in practice. We believe that these contributions will promote further discussion among practitioners and academicians about supporting the institutionalization of telemedicine applications.

Practical Implications

Two examples from real-world practice may illustrate some implications of this discussion. First, health care regulators complain that HTA reports often do not provide them with all the information they require for decision making, since salient information such as organizational impacts and resistance to change are not fully disclosed [80]. This limitation is particularly true for telemedicine applications [81], since contributions in the literature are more interested in discussing clinical outcomes (in particular they aim to verify whether telemedicine services deliver at least the same effectiveness as current face-to-face ones) or technical feasibility (to persuade professionals that telemedicine services are safe and robust to possible technical failures). Our view is that decision making should be rational and accountable, and thus our examples of a preliminary proposal and a full proposal help to frame this requirement in 2 groups of criteria. By knowing which information is salient for supporting decision making, both the agencies that deliver HTA reports and scholars who publish the main results of their studies about new telemedicine applications will be able to increase the impact of their contributions. In fact, they will know in advance which information will be searched through their contributions.

Second, professionals (physicians in particular) are largely involved in clinical studies aimed at understanding the feasibility and the performance of telemedicine applications. This Viewpoint paper offers them some insights into how health care regulators should assess the possibility of implementing their innovations on a large scale. These studies often collect only a small part of the information that decision makers would require to make a decision. This would be a lost opportunity for both patients and professionals, and missing information will be collected through further studies. Consider a hospital involved by a technology supplier in a randomized controlled trial on telemonitoring of patients with chronic disease (eg, chronic heart failure, chronic obstructive pulmonary disease). By knowing in advance which information is salient for decision makers to implement this application on a larger scale, both physicians and technicians will be able to define the most appropriate study design and to collect all relevant data, such as professionals' perception of utility and ease of use as predictor of potential resistance to changing current practice.

Limitations

There are 2 main limitations to our proposed reference process for decision making. These limitations suggest avenues for future research and debates. First, we did not systematically review the literature to identify which criteria should be included in the proposed framework. As mentioned, our literature analysis was intended to provide a first example without attempting to be definitive.

Second, we adopted the health care regulators' perspective for the proposed 2-stage process. As a consequence, it has been designed to support primarily governments and authorities in prioritizing which telemedicine applications should be institutionalized, and therefore funded, at a large scale. Although we believe that this perspective is the most relevant in countries that have a national health care system, such as the United Kingdom, France, Italy, and Norway, we acknowledge that other relevant insights could be gathered by adopting other perspectives, such as that of technology suppliers or insurance companies.

Acknowledgments

This work was undertaken in the framework of a project funded by the Italian Ministry of Health. The authors would like to thank the Lombardy Region, Health Care Directorate, Italy, and all the participants involved in the project. The authors are responsible for the views expressed in this paper.

Conflicts of Interest

None declared

References

1. Sood S, Mbarika V, Jugoo S, Dookhy R, Doarn CR, Prakash N, et al. What is telemedicine? A collection of 104 peer-reviewed perspectives and theoretical underpinnings. Telemed J E Health 2007 Oct;13(5):573-590. [doi: 10.1089/tmj.2006.0073] [Medline: 17999619]
2. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
3. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
4. Whitten P, Holtz B, Nguyen L. Keys to a successful and sustainable telemedicine program. Int J Technol Assess Health Care 2010 Apr;26(2):211-216. [doi: 10.1017/S026646231000005X] [Medline: 20392326]
5. MacFarlane A, Murphy AW, Clerkin P. Telemedicine services in the Republic of Ireland: an evolving policy context. Health Policy 2006 May;76(3):245-258. [doi: 10.1016/j.healthpol.2005.06.006] [Medline: 16026889]
6. Obstfelder A, Engeseth KH, Wynn R. Characteristics of successfully implemented telemedical applications. Implement Sci 2007;2:25 [FREE Full text] [doi: 10.1186/1748-5908-2-25] [Medline: 17662134]
7. May C, Harrison R, Finch T, MacFarlane A, Mair F, Wallace P, Telemedicine Adoption Study Group. Understanding the normalization of telemedicine services through qualitative evaluation. J Am Med Inform Assoc 2003 Dec;10(6):596-604. [doi: 10.1197/jamia.M1145] [Medline: 12925553]
8. Gagnon MP, Légaré F, Fortin JP, Lamothe L, Labrecque M, Duplantie J. An integrated strategy of knowledge application for optimal e-health implementation: a multi-method study protocol. BMC Med Inform Decis Mak 2008;8:17 [FREE Full text] [doi: 10.1186/1472-6947-8-17] [Medline: 18435853]
9. de Bont A, Bal R. Telemedicine in interdisciplinary work practices: on an IT system that met the criteria for success set out by its sponsors, yet failed to become part of every-day clinical routines. BMC Med Inform Decis Mak 2008;8:47 [FREE Full text] [doi: 10.1186/1472-6947-8-47] [Medline: 18954428]
10. Ammenwerth E, Gräber S, Herrmann G, Bürkle T, König J. Evaluation of health information systems-problems and challenges. Int J Med Inform 2003 Sep;71(2-3):125-135. [Medline: 14519405]
11. Ohinmaa A, Hailey D, Roine R. Elements for assessment of telemedicine applications. Int J Technol Assess Health Care 2001;17(2):190-202. [Medline: 11446131]
12. Lettieri E, Masella C. Measuring the value and sustainability of internet-based ICTs in healthcare organisations. Int J Healthcare Technol Manag 2006;7(3/4):319-332. [doi: 10.1504/IJHTM.2006.008439]
13. Masella C, Zanaboni P. Assessment models for telemedicine services in National Health Systems. Int J Healthc Technol Manag 2008;9(5/6):446-472. [doi: 10.1504/IJHTM.2008.020198]
14. Battista RN. Expanding the scientific basis of health technology assessment: a research agenda for the next decade. Int J Technol Assess Health Care 2006;22(3):275-80; discussion 280. [Medline: 16984053]
15. Gagnon MP, Duplantie J, Fortin JP, Landry R. Implementing telehealth to support medical practice in rural/remote regions: what are the conditions for success? Implement Sci 2006;1:18 [FREE Full text] [doi: 10.1186/1748-5908-1-18] [Medline: 16930484]
16. Boddy D, King G, Clark JS, Heaney D, Mair F. The influence of context and process when implementing e-health. BMC Med Inform Decis Mak 2009;9:9 [FREE Full text] [doi: 10.1186/1472-6947-9-9] [Medline: 19183479]
17. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]

18. Miller EA. Solving the disjuncture between research and practice: telehealth trends in the 21st century. Health Policy 2007 Jul;82(2):133-141. [doi: 10.1016/j.healthpol.2006.09.011] [Medline: 17046097]
19. Norum J, Pedersen S, Størmer J, Rumpsfeld M, Stormo A, Jamissen N, et al. Prioritisation of telemedicine services for large scale implementation in Norway. J Telemed Telecare 2007;13(4):185-192. [doi: 10.1258/135763307780908076] [Medline: 17565774]
20. Sibbald SL, Singer PA, Upshur R, Martin DK. Priority setting: what constitutes success? A conceptual framework for successful priority setting. BMC Health Serv Res 2009;9:43 [FREE Full text] [doi: 10.1186/1472-6963-9-43] [Medline: 19265518]
21. Suchman MC. Managing legitimacy: strategic and institutional approaches. Acad Manag Rev 1995;20(3):571-610. [doi: 10.2307/258788]
22. Scott WR. Organizations: Rational, Natural, and Open Systems. Englewood Cliffs, NJ: Prentice-Hall; 1987.
23. Menon D, Stafinski T, Martin D. Priority-setting for healthcare: who, how, and is it fair? Health Policy 2007 Dec;84(2-3):220-233. [doi: 10.1016/j.healthpol.2007.05.009] [Medline: 17628202]
24. Milewa T. Representation and legitimacy in health policy formulation at a national level: perspectives from a study of health technology eligibility procedures in the United Kingdom. Health Policy 2008 Mar;85(3):356-362. [doi: 10.1016/j.healthpol.2007.09.001] [Medline: 17936401]
25. Belton V, Stewart TJ. Multiple Criteria Decision Analysis: An Integrated Approach. Boston, MA: Kluwer Academic Publishers; 2002.
26. Peacock S, Mitton C, Bate A, McCoy B, Donaldson C. Overcoming barriers to priority setting using interdisciplinary methods. Health Policy 2009 Oct;92(2-3):124-132. [doi: 10.1016/j.healthpol.2009.02.006] [Medline: 19346024]
27. Peacock S, Ruta D, Mitton C, Donaldson C, Bate A, Murtagh M. Using economics to set pragmatic and ethical priorities. BMJ 2006 Feb 25;332(7539):482-485. [doi: 10.1136/bmj.332.7539.482] [Medline: 16497768]
28. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc 2006;4:14 [FREE Full text] [doi: 10.1186/1478-7547-4-14] [Medline: 16923181]
29. Field ME, Institute of Medicine. Telemedicine: A Guide to Assessing Telecommunications in Health Care. Washington, DC: National Academy Press; 1996.
30. Hailey D, Jacobs P, Simpson J, Doze S. An assessment framework for telemedicine applications. J Telemed Telecare 1999;5(3):162-170. [Medline: 10628031]
31. Lettieri E, Masella C. Priority setting for technology adoption at a hospital level: relevant issues from the literature. Health Policy 2009 Apr;90(1):81-88. [doi: 10.1016/j.healthpol.2008.07.007] [Medline: 18783844]
32. Daniels N, Sabin J. The ethics of accountability in managed care reform. Health Aff (Millwood) 1998 Oct;17(5):50-64 [FREE Full text] [Medline: 9769571]
33. Daniels N. Accountability for reasonableness. BMJ 2000 Nov 25;321(7272):1300-1301 [FREE Full text] [Medline: 11090498]
34. Hasman A, Holm S. Accountability for reasonableness: opening the black box of process. Health Care Anal 2005 Dec;13(4):261-273. [doi: 10.1007/s10728-005-8124-2] [Medline: 16435464]
35. Madden S, Martin DK, Downey S, Singer PA. Hospital priority setting with an appeals process: a qualitative case study and evaluation. Health Policy 2005 Jul;73(1):10-20. [doi: 10.1016/j.healthpol.2004.11.002] [Medline: 15911053]
36. CORDIS. Rules for Submission of Proposals, and the Related Evaluation, Selection and Award Procedures, Version 3, 21 August 2008: COM(2008)4617. Luxembourg: European Commission; 2008. URL: ftp://ftp.cordis.europa.eu/pub/fp7/docs/fp7-evrules_en.pdf [accessed 2011-08-01] [WebCite Cache ID 60cCBMEOQ]
37. Sculpher M, Drummond M, O'Brien B. Effectiveness, efficiency, and NICE. BMJ 2001 Apr 21;322(7292):943-944 [FREE Full text] [Medline: 11312215]
38. Department of Health (UK). NHS Executive. 1999. "Faster Access to Modern Treatment": How NICE Appraisal Will Work URL: http://www.nice.org.uk/aboutnice/whatwedo/niceandthenhs/fasteraccesstomoderntreatment/faster_access_to_modern_treatment.jsp [accessed 2011-08-01] [WebCite Cache ID 60cCWVNvD]
39. Scottish Telemedicine Initiative. Scottish Telemedicine Action Forum. 2000. Procedure for Submission and Evaluation of Proposals URL: http://www.telemedicine.scot.nhs.uk/Programme/SubmissionProposals.pdf [accessed 2011-05-04] [WebCite Cache ID 5yQicQCiJ]
40. Krupinski EA, LeSueur B, Ellsworth L, Levine N, Hansen R, Silvis N, et al. Diagnostic accuracy and image quality using a digital camera for teledermatology. Telemed J 1999;5(3):257-263. [doi: 10.1089/107830299312005] [Medline: 10908439]
41. Harno K, Arajärvi E, Paavola T, Carlson C, Viikinkoski P. Clinical effectiveness and cost analysis of patient referral by videoconferencing in orthopaedics. J Telemed Telecare 2001;7(4):219-225. [Medline: 11506757]
42. Martínez A, Everss E, Rojo-Alvarez JL, Figal DP, García-Alberola A. A systematic review of the literature on home monitoring for patients with heart failure. J Telemed Telecare 2006;12(5):234-241. [doi: 10.1258/135763306777889109] [Medline: 16848935]
43. Mair F, Ferguson J. More patients with minor injuries could be seen by telemedicine. J Telemed Telecare 2008;14(3):132-134. [doi: 10.1258/jtt.2008.003009] [Medline: 18430279]

44. Al-Abdullah T, Plint AC, Shaw A, Correll R, Gaboury I, Pitters C, et al. The appropriateness of referrals to a pediatric emergency department via a telephone health line. CJEM 2009 Mar;11(2):139-148 [FREE Full text] [Medline: 19272215]
45. Zanaboni P, Scalvini S, Bernocchi P, Borghi G, Tridico C, Masella C. Teleconsultation service to improve healthcare in rural areas: acceptance, organizational impact and appropriateness. BMC Health Serv Res 2009;9:238 [FREE Full text] [doi: 10.1186/1472-6963-9-238] [Medline: 20021651]
46. Scalvini S, Zanelli E, Volterrani M, Martinelli G, Baratti D, Buscaya O, et al. A pilot study of nurse-led, home-based telecardiology for patients with chronic heart failure. J Telemed Telecare 2004;10(2):113-117. [doi: 10.1258/135763304773391576] [Medline: 15068649]
47. Whited JD, Hall RP, Foy ME, Marbrey LE, Grambow SC, Dudley TK, et al. Patient and clinician satisfaction with a store-and-forward teledermatology consult system. Telemed J E Health 2004;10(4):422-431. [doi: 10.1089/tmj.2004.10.422] [Medline: 15689645]
48. Fox KC, Somes GW, Waters TM. Timeliness and access to healthcare services via telemedicine for adolescents in state correctional facilities. J Adolesc Health 2007 Aug;41(2):161-167. [doi: 10.1016/j.jadohealth.2007.05.001] [Medline: 17659220]
49. Goldberg LR, Piette JD, Walsh MN, Frank TA, Jaski BE, Smith AL, WHARF Investigators. Randomized trial of a daily electronic home monitoring system in patients with advanced heart failure: the Weight Monitoring in Heart Failure (WHARF) trial. Am Heart J 2003 Oct;146(4):705-712. [doi: 10.1016/S0002-8703(03)00393-4] [Medline: 14564327]
50. Waldmann A, Katalinic A, Schwaab B, Richardt G, Sheikhzadeh A, Raspe H. The TeleGuard trial of additional telemedicine care in CAD patients. 2 Morbidity and mortality after 12 months. J Telemed Telecare 2008;14(1):22-26. [doi: 10.1258/jtt.2007.070512] [Medline: 18318925]
51. Thomas EJ, Lucke JF, Wueste L, Weavind L, Patel B. Association of telemedicine for remote monitoring of intensive care patients with mortality, complications, and length of stay. JAMA 2009 Dec 23;302(24):2671-2678 [FREE Full text] [doi: 10.1001/jama.2009.1902] [Medline: 20040555]
52. Finkelstein SM, Speedie SM, Potthoff S. Home telehealth improves clinical outcomes at lower cost for home healthcare. Telemed J E Health 2006 Apr;12(2):128-136. [doi: 10.1089/tmj.2006.12.128] [Medline: 16620167]
53. Rees RS, Bashshur N. The effects of TeleWound management on use of service and financial outcomes. Telemed J E Health 2007 Dec;13(6):663-674. [doi: 10.1089/tmj.2007.9971] [Medline: 18069917]
54. Giordano A, Scalvini S, Zanelli E, Corrà U, Longobardi GL, Ricci VA, et al. Multicenter randomised trial on home-based telemanagement to prevent hospital readmission of patients with chronic heart failure. Int J Cardiol 2009 Jan 9;131(2):192-199. [doi: 10.1016/j.ijcard.2007.10.027] [Medline: 18222552]
55. McConnochie KM, Wood NE, Herendeen NE, Ng PK, Noyes K, Wang H, et al. Acute illness care patterns change with use of telemedicine. Pediatrics 2009 Jun;123(6):e989-e995 [FREE Full text] [doi: 10.1542/peds.2008-2698] [Medline: 19482750]
56. Suri JS, Dowling A, Laxminarayan S, Singh S. Economic impact of telemedicine: a survey. Stud Health Technol Inform 2005;114:140-156. [Medline: 15923771]
57. Norum J, Bergmo TS, Holdø B, Johansen MV, Vold IN, Sjaaeng EE, et al. A tele-obstetric broadband service including ultrasound, videoconferencing and cardiotocogram. A high cost and a low volume of patients. J Telemed Telecare 2007;13(4):180-184. [doi: 10.1258/135763307780908085] [Medline: 17565773]
58. Le Goff-Pronost M, Sicotte C. The added value of thorough economic evaluation of telemedicine networks. Eur J Health Econ 2010 Feb;11(1):45-55. [doi: 10.1007/s10198-009-0162-5] [Medline: 19562393]
59. Dávalos ME, French MT, Burdick AE, Simmons SC. Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis. Telemed J E Health 2009 Dec;15(10):933-948. [doi: 10.1089/tmj.2009.0067] [Medline: 19954346]
60. Bergmo TS. Can economic evaluation in telemedicine be trusted? A systematic review of the literature. Cost Eff Resour Alloc 2009;7:18 [FREE Full text] [doi: 10.1186/1478-7547-7-18] [Medline: 19852828]
61. Groves RH, Holcomb BW, Smith ML. Intensive care telemedicine: evaluating a model for proactive remote monitoring and intervention in the critical care setting. Stud Health Technol Inform 2008;131:131-146. [Medline: 18305328]
62. Hicks LL, Boles KE, Hudson ST, Madsen RW, Kling B, Tracy J, et al. Using telemedicine to avoid transfer of rural emergency department patients. J Rural Health 2001;17(3):220-228. [Medline: 11765886]
63. Stalker HJ, Wilson R, McCune H, Gonzalez J, Moffett M, Zori RT. Telegenetic medicine: improved access to services in an underserved area. J Telemed Telecare 2006;12(4):182-185. [doi: 10.1258/135763306777488762] [Medline: 16774698]
64. Galli R, Keith JC, McKenzie K, Hall GS, Henderson K. TelEmergency: a novel system for delivering emergency care to rural hospitals. Ann Emerg Med 2008 Mar;51(3):275-284. [doi: 10.1016/j.annemergmed.2007.04.025] [Medline: 17764784]
65. Tobman M, Nätscher C, Sussmann H, Horsch A. A new approach for integration of telemedicine applications into existing information systems in healthcare. Stud Health Technol Inform 2002;90:152-155. [Medline: 15460679]
66. Craft RL. Toward technical interoperability in telemedicine. Telemed J E Health 2005 Jun;11(3):384-404. [doi: 10.1089/tmj.2005.11.384] [Medline: 16035933]
67. Stanberry B. The legal and ethical aspects of telemedicine. 3: Telemedicine and malpractice. J Telemed Telecare 1998;4(2):72-79. [Medline: 9744161]

68. Silverman RD. Current legal and ethical concerns in telemedicine and e-medicine. J Telemed Telecare 2003;9 Suppl 1:S67-S69. [doi: 10.1258/135763303322196402] [Medline: 12952729]
69. Mair F, Whitten P. Systematic review of studies of patient satisfaction with telemedicine. BMJ 2000 Jun 3;320(7248):1517-1520 [FREE Full text] [Medline: 10834899]
70. Jennett P, Yeo M, Pauls M, Graham J. Organizational readiness for telemedicine: implications for success and failure. J Telemed Telecare 2003;9 Suppl 2:S27-S30. [doi: 10.1258/135763303322596183] [Medline: 14728753]
71. Robinson DF, Savage GT, Campbell KS. Organizational learning, diffusion of innovation, and international collaboration in telemedicine. Health Care Manage Rev 2003 Mar;28(1):68-78. [Medline: 12638374]
72. Monrad Aas IH. The Organizational Challenge for Health Care From Telemedicine and e-Health. Oslo: Work Research Institute; 2007.
73. Bartoli L, Zanaboni P, Masella C, Ursini N. Systematic review of telemedicine services for patients affected by chronic obstructive pulmonary disease (COPD). Telemed J E Health 2009 Nov;15(9):877-883. [doi: 10.1089/tmj.2009.0044] [Medline: 19919194]
74. Callens S, Cierkens K. Legal aspects of E-HEALTH. Stud Health Technol Inform 2008;141:47-56. [Medline: 18953124]
75. Currell R, Urquhart C, Wainwright P, Lewis R. Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2000(2):CD002098. [doi: 10.1002/14651858.CD002098] [Medline: 10796678]
76. Wootton R, Bloomer SE, Corbett R, Eedy DJ, Hicks N, Lotery HE, et al. Multicentre randomised control trial comparing real time teledermatology with conventional outpatient dermatological care: societal cost-benefit analysis. BMJ 2000 May 6;320(7244):1252-1256 [FREE Full text] [Medline: 10797038]
77. Wootton R, Craig J, Patterson V. In: Wootton RWF, Craig J, Patterson V, editors. Introduction to Telemedicine. London, UK: Rittenhouse Book Distributors; 2006.
78. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth (3): a systematic review of published definitions. J Med Internet Res 2005;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]
79. Hailey D, Ohinmaa A, Roine R. Study quality and evidence of benefit in recent assessments of telemedicine. J Telemed Telecare 2004;10(6):318-324. [doi: 10.1258/1357633042602053] [Medline: 15603628]
80. Milne R, Clegg A, Stevens A. HTA responses and the classic HTA report. J Public Health Med 2003 Jun;25(2):102-106 [FREE Full text] [Medline: 12848397]
81. Coughlan J, Eatock J, Eldabi T. Evaluating telemedicine: a focus on patient pathways. Int J Technol Assess Health Care 2006;22(1):136-142. [Medline: 16673690]

Abbreviations

A4R: accountability for reasonableness
HTA: health technology assessment
NICE: National Institute for Health and Clinical Excellence

Edited by G Eysenbach; submitted 18.10.10; peer-reviewed by MP Gagnon; comments to author 01.04.11; revised version received 19.05.11; accepted 18.05.11; published 28.09.11.

Please cite as:
Zanaboni P, Lettieri E
Institutionalizing Telemedicine Applications: The Challenge of Legitimizing Decision-Making
J Med Internet Res 2011;13(3):e72
URL: http://www.jmir.org/2011/3/e72/
doi:10.2196/jmir.1669
PMID:21955510

Original Paper

Using the Internet to Understand Smokers' Treatment Preferences: Informing Strategies to Increase Demand

J Lee Westmaas[1,2], PhD; Lorien Abroms[3], ScD; Jeuneviette Bontemps-Jones[1], MPH, CHES; Joseph E Bauer[4], PhD; Jeanine Bade[5], MA

[1]Behavioral Research Center, American Cancer Society, Atlanta, GA, United States

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

[3]Department of Prevention and Community Health, School of Public Health and Health Services, George Washington University, Washington, DC, United States

[4]Statistics and Evaluation Center, American Cancer Society, Atlanta, GA, United States

[5]Department of Health Promotions, American Cancer Society, Atlanta, GA, United States

Corresponding Author:
J Lee Westmaas, PhD
Behavioral Research Center
American Cancer Society
250 Williams St.
Atlanta, GA, 30303
United States
Phone: 1 404 329 7730 ext 7730
Fax: 1 404 929 6832
Email: lee.westmaas@cancer.org

Abstract

Background: Most smokers attempt to quit on their own even though cessation aids can substantially increase their chances of success. Millions of smokers seek cessation advice on the Internet, so using it to promote cessation products and services is one strategy for increasing demand for treatments. Little is known, however, about what cessation aids these smokers would find most appealing or what predicts their preferences (eg, age, level of dependence, or timing of quit date).

Objective: The objective of our study was to gain insight into how Internet seekers of cessation information make judgments about their preferences for treatments, and to identify sociodemographic and other predictors of preferences.

Methods: An online survey assessing interest in 9 evidence-based cessation products and services was voluntarily completed by 1196 smokers who visited the American Cancer Society's Great American Smokeout (GASO) webpage. Cluster analysis was conducted on ratings of interest.

Results: In total, 48% (572/1196) of respondents were "quite a bit" or "very much" interested in nicotine replacement therapy (NRT), 45% (534/1196) in a website that provides customized quitting advice, and 37% (447/1196) in prescription medications. Only 11.5% (138/1196) indicated similar interest in quitlines, and 17% (208/1196) in receiving customized text messages. Hierarchical agglomerative cluster analysis revealed that interest in treatments formed 3 clusters: *interpersonal–supportive* methods (eg, telephone counseling, Web-based peer support, and in-person group programs), *nonsocial–informational* methods (eg, Internet programs, tailored emails, and informational booklets), and *pharmacotherapy* (NRT, bupropion, and varenicline). Only 5% (60/1196) of smokers were "quite a bit" or "very much" interested in interpersonal–supportive methods compared with 25% (298/1196) for nonsocial–informational methods and 33% (399/1196) for pharmacotherapy. Multivariate analyses and follow-up comparisons indicated that level of interest in pharmacotherapy ("quite a bit or "very much" vs. "not at all") varied as a function of education (n = 575, χ^2_3 =16.6, *P* = .001), age (n = 528, χ^2_3 = 8.2, *P* = .04), smoking level (n = 514, χ^2_3 = 9.5, *P* = .02), and when smokers were planning to quit (n = 607, χ^2_4 = 34.0, $P < .001$). Surprisingly, greater age was associated with stronger interest in nonsocial–informational methods (n = 367, χ^2_3 = 10.8, *P* = .01). Interest in interpersonal–supportive methods was greater if smokers had used a quitline before (n = 259, χ^2_1 = 18.3, $P < .001$), or were planning to quit earlier rather than later (n = 148, χ^2_1 = 4.9, *P* = .03).

Conclusions: Smokers accessing the Internet for information on quitting appear to differentiate cessation treatments by how much interpersonal interaction or support the treatment entails. Quitting date, smoking level, and sociodemographic variables can identify smokers with varying levels of interest in the 3 classes of cessation methods identified. These results can potentially be used to more effectively target and increase demand for these treatments among smokers searching the Internet for cessation information.

(J Med Internet Res 2011;13(3):e58) doi:10.2196/jmir.1666

KEYWORDS

Consumer demand; pharmacotherapy; quitline counseling; smoking cessation; social support

Introduction

Several effective tobacco-cessation products and services are available to help smokers who want to quit. These can double or triple the rate of cessation compared with quitting without help [1]. In spite of their availability, however, the use of these products and services is low, as most smokers opt to quit on their own [2]. For example, the North American Quitline Consortium estimated that the reach of quitlines, calculated as the proportion of all smokers in a US state who complete a program of phone counseling, was between only 1% and 2% [3].

There has been little research on why smokers are not using evidenced-based treatments to quit. One study found that, although smokers planning to quit expressed several barriers around quitline use, none of the self-reported barriers predicted actual calls made to a quitline [4]. Another study of 636 German smokers who had never used cessation aids when trying to quit found that the most endorsed barrier to not using cessation aids was belief in one's own ability to quit [5]. A quarter of smokers also believed that cessation aids would not be helpful, and cited their cost as a reason for not using them. Lack of awareness of cessation aids also appears to be a barrier. In a study of smokers in the United Kingdom, only 30% of current and former smokers were aware of various cessation services provided by the National Health Service [6]. This is consistent with findings from a study of Canadian smokers, who demonstrated low recall of cessation methods [7].

To increase utilization of cessation products and services, the National Tobacco Cessation Collaborative, an American and Canadian consortium of leading nonprofit and government agencies dedicated to reducing the burden of tobacco use, delineated several core strategies to increase demand for available treatments. One of these strategies includes the recommendation to understand what smokers need and want, instead of viewing them as "passive treatment beneficiaries rather than treatment consumers" (p. S308) [8]. The millions of smokers who seek cessation information on the Internet [9] are a population that can be targeted to increase demand for treatments. The Internet can also be used to understand their cessation needs and wants, and to easily provide access to the treatments they prefer. However, little is known about preferences for cessation methods in this population. To that end we asked smokers who were seeking cessation information on the Internet to rate their interest in various evidenced-based cessation products and services. We subjected these ratings to a cluster analysis, an analytic technique used extensively by market researchers to evaluate brand or product preferences [10]. We used this approach to determine whether particular groups of products and services were preferred over others so that we could infer underlying motivations for preferences. We also examined whether sociodemographic and smoking behavior variables were associated with preferences for particular treatments. Knowledge of who is interested in what kinds of treatments can be used to promote cessation treatments on the Internet in a more targeted manner, and could potentially increase demand for them.

Using the Internet to Inform Smokers About Cessation Treatments

Many smokers thinking of quitting access the Internet for general information on smoking cessation. The Internet may therefore be an effective medium for promoting evidenced-based cessation aids [11]. According to a random-digit-dial survey conducted in the United States in 2006, an estimated 9% of Internet users (approximately 12.7 million individuals) searched the Internet for information on "how to quit smoking" [9]. Moreover, compared with traditional media such as radio or television advertising, online advertising has been shown to be a more cost-efficient mode of recruiting smokers to Internet- and telephone-based cessation treatments [12]. It is now also possible for marketers to construct a sociodemographic profile for a computer user by gathering information on different websites visited. This information is then used to market products to groups of people who would be receptive to messages about the products. This strategy, termed "behavioral targeting," could be employed to promote specific cessation products or services to smokers most likely to use them [13]. In addition, some products and services (eg, Internet programs for cessation) can be made immediately available—that is, at the time smokers are searching the Internet for help on quitting. Overall, these developments speak to the tremendous potential of the Internet as an ideal medium through which large numbers of smokers can be reached and provided with information on, or access to, various evidence-based cessation treatments.

Identifying and targeting potential quitters by providing them information on or immediate access to treatments would be maximally effective if smokers' preferences for treatments and predictors of those preferences were known. Only 1 study to our knowledge, however, asked smokers seeking cessation information on the Internet to rate their perceptions of various treatments [11]. Results indicated that smokers perceived that telephone counseling or receiving support through Internet chat, forums, or email would be least helpful. In contrast, smokers perceived that information about withdrawal or individually tailored information would be most helpful [11].

Research on what methods are used most frequently by smokers when attempting to quit could provide some indication of what cessation products and services would be favored. Low use or nonuse of a particular product or service, however, may be due to a smoker simply not knowing it exists. Some products may also be used only because they were recommended or were available at the time. Nevertheless, data on actual use of evidence-based cessation treatments can be used to form hypotheses about what types of treatments might be preferred, or how preferences might vary as a function of sociodemographic characteristics. For example, in the 2005 National Health Interview Survey (NHIS) [14], pharmacotherapy was used more frequently than behavioral methods, which was used infrequently (<5%). Results also indicated that older smokers (aged ≥25 years) were more likely than younger smokers to use pharmacotherapy. In addition, for younger smokers, but not for older smokers, greater educational attainment was associated with having used pharmacotherapy. Based on these results we hypothesized that pharmacotherapy and behavioral methods would form separate clusters of preferences, and that interest in pharmacotherapy would be greater than interest in behavioral methods. We also hypothesized that, after controlling for age, greater educational attainment would be associated with greater interest in pharmacotherapy (as more educated smokers are likely to be more knowledgeable about nicotine replacement therapy [NRT] and less susceptible to myths about it).

We also expected preferences to vary by gender. In the 2-year longitudinal National Youth Smoking Cessation Survey of smokers aged 16–24 years [15], among those who had tried to quit smoking at least once, seeking help from a professional was more common among females than among males. The proportion of female smokers who contact quitlines is also greater [16]. This may be a result of the male stereotype emphasizing independence and avoidance of emotional disclosure, or the greater acceptability for females to seek others' assistance [17]. Based on these results we hypothesized that female smokers would be more interested than male smokers in methods involving counseling or seeking advice from others. In sum, our hypotheses, based on the findings described earlier, were that (1) pharmacotherapy and behavioral methods would form separate clusters, and interest in pharmacotherapy would be greater than interest in behavioral methods, (2) more educated smokers would be more interested than less educated smokers in using pharmacotherapy, and (3) compared with male smokers, female smokers would be more interested in methods involving counseling such as quitlines or group cessation programs. Because of younger smokers' greater use of the Internet [18] we also wanted to explore the following hypothesis: younger smokers seeking cessation information on the Internet would be more interested than older smokers in using Internet programs for cessation, or other technologically involved treatments (eg, text messages).

Methods

Participants

Participants' mean age was 38.4 years (SD 9.1) and they smoked an average of 15.9 cigarettes per day (SD 9.1). The majority of the sample was female (840, 74.3%). A college degree or higher was reported by 34.2% (388), with 39.6% (449) reporting some college, 21.1% (239) completing high school or its equivalent, and 5.1% (58) completing grade 11 or lower. A minority (43, 3.8%) were Latino/Hispanic, whereas the majority were not (1068, 94.8%), or indicated “Don't know” (15, 1.4%). The majority of participants were white (984, 87%). Black or African American smokers constituted 5% (57) and Asian smokers, 1.6% (18) of the sample. The remaining racial groups were collapsed into 1 category that comprised 6.4% (75) of the sample (ie, Pacific Islander, American Indian, Alaska Native, Other, and Don't know) (see Table 1).

Table 1. Descriptive statistics of the study population

Characteristic	n	%
Smoking rate (cigarettes per day)		
<10	230	22.1
10–19	379	36.4
20–29	329	31.6
≥30	103	9.9
Age group (years)		
≤25	164	15.7
26–40	438	41.9
41–55	358	34.2
≥56	86	8.2
Gender		
Female	840	74.3
Male	291	25.7
Education level		
Grade ≤11	58	5.1
High school graduate or GED[a]	239	21.1
Some college	449	39.6
College graduate or higher	388	34.2
Race		
White	984	87.0
Black/African American	57	5.0
Asian	18	1.6
Pacific Islander/American Indian/Alaskan Native/Other/Don’t know	72	6.4
Quit in past year		
Yes	730	63.2
No	425	36.8
Quit date		
In the next 24 hours	266	22.2
In next week or two	362	30.3
In next month	209	17.5
In next 6 months	114	9.5
In future/undecided	245	20.5
Free quitline help available?		
Yes	256	21.4
No	35	3.0
Don’t know	861	74.7
Ever used a quitline		
Yes	112	9.7
No	884	76.6
N/A[b]—never tried to quit before	158	13.7

[a] General equivalency diploma.

[b] Not applicable.

Procedure

A 12-item online questionnaire was posted for 11 months on the American Cancer Society's (ACS) Great American Smokeout (GASO) webpage. The survey was posted 1 week prior to the 2008 GASO event (November 20). The GASO webpage is the ACS's online portal for information about quitting and received 92,946 unique views during the study period. An introductory paragraph on the GASO webpage explained that the ACS was interested in learning about how smokers quit, and that if they were interested to click on the link appearing below. The link led respondents to a page in which a consent section appeared above the survey. The survey was completed by 1594 current smokers over the entire study period. Approximately half of responses (845, 53%) were collected by 6 days after GASO, and the other half during the remaining months of the study period. We excluded 90 participants due to missing data on sociodemographic variables. Analyses reported below were restricted to participants who provided responses on sociodemographic variables and who did not choose the "don't know" option on the items assessing interest in cessation products and services (N = 1196). The voluntary and anonymous survey, which was approved by the Emory University institutional review board, did not provide incentives, financial or otherwise, for completion.

The current study also subjected ratings of interest in cessation products and services to a cluster analysis. Cluster analysis is an assumption-free classification technique that is commonly used in market research to understand consumer behavior [10]. It simultaneously minimizes within-group variance and maximizes between-groups variance, so that members or variables of the same group are more similar to each other than to those of other groups [19]. It can thus be used to infer the underlying dimensions that form the basis of smokers' preferences among treatments. For the present report we used SPSS 18 (IBM Corporation, Somers, NY, USA) to perform hierarchical agglomerative clustering, with between-groups linkage and squared euclidean distance as the similarity metric. To determine the reliability of clusters obtained we first conducted a cluster analysis on a random sample of 50% of cases (598) and a second one on the remaining cases. For both samples membership in clusters at each level of agglomeration was identical. The composition of clusters was also the same when analyses were conducted on the first 50% and the latter 50% of cases. The final cluster analysis presented thus used the full sample.

Predictors of Cessation Methods

To examine the relationship between interest in cessation products and services, and sociodemographic and smoking behavior variables, a multivariate analysis of variance (MANOVA) was conducted. Cessation methods that formed clusters were averaged to obtain summary variables that were the set of dependent variables in the MANOVA. Independent variables (categorical) were gender, age, race, education, when smokers were planning to quit, prior use of a quitline, smoking rate, knowledge of quitlines' free availability, and whether an attempt to quit was made in the past year. A second MANOVA was conducted that excluded independent variables with nonsignificant multivariate main effects (gender, quitline knowledge, and past-year attempt). Multivariate results from this final MANOVA are presented. Analyses that controlled for the number of days between when the survey was completed and the start date of the survey were also conducted.

Follow-up Tukey pairwise comparisons that controlled for the familywise error rate were conducted. These examined differences on dependent variables among levels of the significant independent variables (ie, significant according to the univariate results). For comparisons that the Tukey tests indicated were statistically significant (ie, $P < .05$) we also conducted chi-square tests to demonstrate the association between the independent and dependent variables. For these chi-square analyses we first created 2 contrasting groups for each dependent variable: smokers who were "not at all" interested versus smokers who were "quite a bit" or "very much" interested. Smokers who indicate they are "quite a bit" or "very much" interested in a cessation method are likely to be most receptive to trying a cessation method if it is available to them, at least much more so than smokers who report that they are "not at all" interested. These latter smokers are likely to be more difficult to reach with marketing efforts aimed at encouraging use of a particular cessation method. The chi-square follow-up analyses thus aided interpretation by illustrating which groups might be fairly easy or more challenging to encourage to use particular cessation methods. We followed a conservative approach of reporting only the chi-square associations that were also significant (in addition to the Tukey tests described above). We also provide an effect size measure, Cramer's V, that ranges from 0 to 1 [20,21].

Measures

Sociodemographic and Smoking Behavior Variables

Participants were asked to indicate when they planned to quit (in the next 24 hours, in the next week or two, in the next month, in the next 6 months, sometime in the future but haven't decided when, not applicable (N/A)—already quit, other), whether they knew if "free help from a counselor at a quitline" was available to all smokers in their state (yes, no, don't know), whether they had ever called a quitline to help them quit smoking (yes, no, N/A—never tried to quit before), the number of cigarettes smoked per day, whether they tried to seriously quit in the past year (yes, no), their gender, and their age.

Interest in Cessation Products and Services

Smokers were asked to indicate how interested they would be in using the following cessation products and services on a scale from 1 (not at all) to 5 (very much). Specific items, as written, were (1) "Using a telephone quitline (a quitline has trained counselors help you over the phone with your quit attempt," (2) "Using a website that gives professional advice about quitting smoking that is customized for you," (3) "Using the Internet to chat with other smokers who are trying to quit," (4) "Receiving emails timed around your quit date that contain professional advice about quitting that is customized for you," (5) "Attending a program led by a professional and that involves a few meetings with other smokers trying to quit," (6) "Using nicotine replacement therapy (eg, the patch) which is available without a prescription," (7) "Using booklets or other printed materials that give professional advice on how to quit," (8) "Receiving text messages on your cell phone timed around your quit date that contain professional advice about quitting that is customized for you," and (9) "Using prescription medications for quitting such as Zyban (bupropion) or Chantix (varenicline)."

Results

Descriptive Statistics

The majority of respondents had tried to quit in the past year (730, 63.2%) (see Table 1). In addition, 22.2% (266) of smokers planned to quit immediately (ie, in the next 24 hours), 30.3% (362) in the next week or two, 17.5% (209) in the next month, 9.5% (114) in 6 months, and 20.5% (245) at some undecided time in the future.

A large majority of smokers (861, 74.7%) did not know whether free help from a quitline counselor was available to all smokers in their state; 3.0% (35) indicated that such help was *not* available and 21.4% (256) indicated that it was. Not surprisingly, a similar majority, 76.6% (884), reported not ever having used a quitline, 9.7% (112) reported previously using a quitline, and 13.7% (158) indicated this question was not applicable because they had never before tried to quit.

The cessation method that received the greatest proportion of respondents who indicated being "quite a bit" or "very much" interested was NRT (572/1196, 47.8%), followed by a website that provides customized quitting advice (534/1196, 44.6%) and prescription medications (447/1196, 37.4%). Only 11.5% (138/1196) of respondents indicated being "quite a bit" or "very much" interested in using quitlines, and only 17.4% (208/1196) reported similar interest in receiving customized text messages. The proportion of respondents in each sociodemographic category who were "quite a bit" or "very much" interested in a particular cessation method are presented in Table 2. For example, among respondents 25 years or younger (n = 164), only 8.5% (14) were "quite a bit" or "very much" interested in quitlines whereas 39.6% (65) had a similarly strong interest in NRT.

Table 2. Number and percentage of respondents in each sociodemographic category who were “quite a bit” or “very much” interested in each cessation method

	Quitline	Website	Web peer support	Emails	Group programs	NRT[a]	Prescription medications	Cessation booklets	Text messages	n
Age group (years)										
≤25	14 (8.5%)	64 (39.0%)	33 (20.1%)	47 (28.7%)	30 (18.3%)	65 (39.6%)	51 (31.1%)	34 (20.7%)	45 (27.4%)	164
26–40	44 (10.0%)	94 (43.2%)	94 (21.5%)	162 (37.0%)	85 (19.4%)	217 (49.5%)	187 (42.7%)	146 (33.3%)	79 (18.0%)	438
41–55	54 (15.1%)	87 (50.8%)	87 (24.3%)	143 (39.9%)	94 (26.3%)	185 (51.7%)	138 (38.5%)	121 (33.8%)	52 (14.5%)	358
≥56	9 (10.5%)	21 (50.0%)	21 (24.4%)	35 (40.7%)	24 (27.9%)	47 (54.7%)	30 (34.9%)	28 (32.6%)	10 (11.6%)	86
Gender										
Male	24 (8.2%)	120 (41.2%)	50 (17.2%)	90 (30.9%)	56 (19.2%)	142 (48.8%)	93 (32.0%)	71 (24.4%)	44 (15.1%)	291
Female	106 (12.6%)	391 (46.5%)	200 (23.8%)	324 (38.6%)	193 (23.0%)	411 (48.9%)	338 (40.2%)	284 (33.8%)	154 (18.3%)	840
Education										
Grade ≤11	10 (17.2%)	23 (39.7%)	15 (25.9%)	21 (36.2%)	9 (15.5%)	17 (29.3%)	17 (29.3%)	17 (29.3%)	9 (15.5%)	58
High school graduate/GED[b]	24 (10.0%)	103 (43.1)	53 (22.2%)	82 (34.3%)	47 (19.7%)	122 (51.0%)	87 (36.4%)	65 (27.2%)	30 (12.6%)	239
Some college	46 (10.2%)	196 (43.7%)	99 (22.0%)	157 (35.0%)	99 (22.0%)	226 (50.3%)	184 (41.0%)	154 (34.3%)	79 (17.6%)	449
College graduate or higher	51 (13.1%)	190 (49.0%)	84 (21.6%)	155 (39.9%)	95 (24.5%)	189 (48.7%)	144 (37.1%)	120 (30.9%)	81 (20.9%)	388
Race										
White	105 (10.7%)	452 (45.9%)	210 (21.3%)	359 (36.5%)	206 (20.9%)	490 (49.8%)	376 (38.2%)	303 (30.8%)	164 (16.7%)	984
Black/African American	15 (26.3%)	28 (49.1%)	18 (31.6%)	28 (49.1%)	22 (38.6%)	26 (45.6%)	20 (35.1%)	24 (42.1%)	18 (31.6%)	57
Asian	2 (11.1%)	8 (44.4%)	5 (27.8%)	5 (27.8%)	4 (22.2%)	8 (44.4%)	8 (44.4%)	8 (44.4%)	4 (22.2%)	18
Other	9 (12.5%)	23 (31.9%)	16 (22.2%)	23 (31.9%)	17 (23.6%)	28 (38.9%)	27 (37.5%)	21 (29.2%)	13 (18.1%)	72

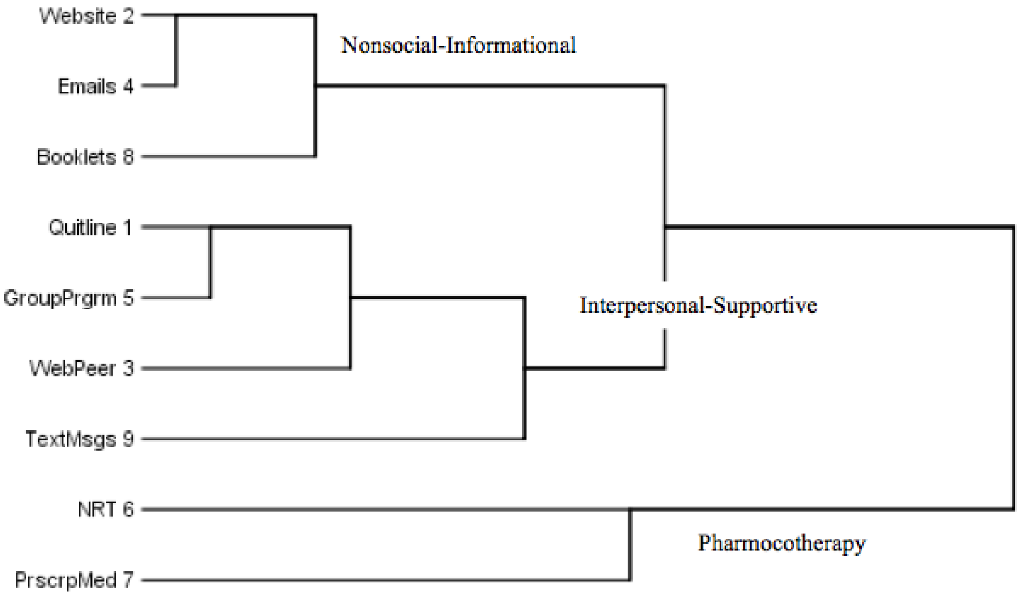
[a] Nicotine replacement therapy.

[b] General equivalency diploma.

Cluster Analysis

Inspection of the dendrogram from the cluster analysis and membership at each cluster stage suggested that a 3-cluster solution was appropriate (Figure 1). Of the 3 clusters, one consisted of methods that all involved interpersonal interaction or support, specifically using a telephone quitline, attending a group-based cessation program, using the Internet to chat with other smokers who are trying to quit, and receiving text messages by cell phone. This group of cessation products and services were labeled *interpersonal–supportive* methods. The appearance of text messages in this cluster is likely attributable to the fact that text messages are typically used for interpersonal reasons—for example, to communicate among friends.

Figure 1. Dendrogram indicating clusters from cluster analysis of interest in cessation products and services (Msgs = messages, NRT = nicotine replacement therapy, Prgrm = program, PrscrpMed = prescription medication)



A second cluster consisted of using a customized website, receiving timed emails, and using printed materials that give professional advice on how to quit. We labeled this group of treatments *nonsocial–informational* methods. The third and final cluster consisted of using NRT and prescription medications. We labeled this cluster *pharmacotherapy*.

Composite variables to represent interest in each of the 3 types of cessation methods were computed as the average of the variables in each cluster. The means and standard deviations for each cluster were as follows: nonsocial–informational methods (mean 2.9, SD 1.2); interpersonal–supportive methods (mean 2.1, SD 1.0); and pharmacotherapy (mean 3.0, SD 1.3). Paired samples *t* tests indicated that interest in interpersonal–supportive methods was significantly lower than interest in pharmacotherapy ($t_{1195} = –21.64$, $P < .001$) or interest in nonsocial–informational methods ($t_{1195} = 30.03$, $P < .001$). Interest in nonsocial–informational methods and in pharmacotherapy were only marginally different ($t_{1195} = –1.72$, $P = .09$). Only 5% (60) of smokers were “quite a bit” or “very much” interested in interpersonal–supportive methods compared to 25% (298) for nonsocial–informational methods and 33% (399) for pharmacotherapy.

Results from multivariate analyses are reported in Table 3. Analyses that controlled for the number of days between when the survey was completed and the start date of the survey revealed the same pattern of significance for multivariate and univariate associations. Univariate analyses indicated that all independent variables except race and education were associated with interest in pharmacotherapy. Also, all independent variables except age were associated with interest in interpersonal–supportive methods. Only quit date was significantly associated with interest in nonsocial–informational methods ($F_{4,938} = 8.10$, $P < .001$). Age group showed a marginal association ($F_{3,938} = 2.50$, $P = .06$).

Table 3. Results of multivariate analysis of variance indicating main effects of independent variables on smokers' interest in cessation treatments

Multivariate			Univariate			
Independent variable	Pillai trace	*P* [a]	Dependent variable	*F*	df	*P* [b]
Quit date	4.21	<.001	Pharmacotherapy	4.11	4, 938	.003
			Interpersonal–supportive	4.04	4, 938	.003
			Nonsocial–informational	8.10	4, 938	<.001
Used a quitline before	3.79	<.001	Pharmacotherapy	4.18	2, 938	.016
			Interpersonal–supportive	8.69	2, 938	<.001
			Nonsocial–informational	1.26	2, 938	.28
Smoking rate	1.95	.04	Pharmacotherapy	3.02	3, 938	.03
			Interpersonal–supportive	2.74	3, 938	.04
			Nonsocial–informational	1.48	3, 938	.22
Age group	2.48	.008	Pharmacotherapy	3.30	3, 938	.02
			Interpersonal–supportive	0.31	3, 938	.82
			Nonsocial–informational	2.50	3, 938	.06
Race	2.59	.006	Pharmacotherapy	0.31	3, 938	.82
			Interpersonal–supportive	6.71	3, 938	<.001
			Nonsocial–informational	1.64	3, 938	.18
Education level	2.16	.02	Pharmacotherapy	2.15	3, 938	.09
			Interpersonal–supportive	3.37	3, 938	.02
			Nonsocial–informational	0.84	3, 938	.47

[a] Indicates significance of multivariate relationship between the independent variable and the set of 3 dependent variables.

[b] Indicates significance of univariate relationship between the independent variable and each dependent variable.

Quit date

The later smokers planned to quit, the greater their interest in pharmacotherapy. While 48% (69/143) of smokers planning to quit within a day were "quite a bit" or "very much" interested in pharmacotherapy, a significantly greater proportion (≥75%); planning to quit later were similarly interested, either "in the next week or two" (132/175, 75%; n = 318, $\chi^2_1 = 25.0$, *P* < .001, *V* = .28), or "in the next month" (76/97, 78%; n = 240, $\chi^2_1 = 21.9$, *P* < .001, *V* = .30). In contrast, interest in interpersonal–supportive methods was greater the earlier smokers planned to quit; over twice as many smokers planning to quit "in the next week or two" (22/82, 27%) were "quite a bit" or "very much" interested in interpersonal–supportive methods as smokers who planned to quit at some undecided time in the future (8/66, 12%; n = 148, $\chi^2_1 = 4.9$, *P* < .02, *V* = .18).

Interest in nonsocial–informational methods was greater among smokers planning to quit in the next week or two (121/144, 84%) compared with those quitting in the next 24 hours (72/107, 67%; n = 251, $\chi^2_1 = 9.7$, *P* < .002, *V* = .20). On the other hand, the proportion of smokers planning to quit at some undecided time in the future who were "quite a bit" or "very much" interested was significantly lower (28/71, 39%) than the 67% (72/107) of smokers planning to quit in the next 24 hours (n = 178, $\chi^2_1 = 13.4$, *P* < .001, *V* = .28), the 84% (121/144) of smokers planning to quit in the next week or two (n = 215, $\chi^2_1 = 44.4$, *P* < .0001, *V* = .46), the 77% (53/69) planning to quit in the next month (n = 140, $\chi^2_1 = 20.0$, *P* < .0001, *V* = .38), or the 75% (24/32) planning to quit in the next 6 months (n = 103, $\chi^2_1 = 11.2$, *P* < .001, *V* = .33).

Previous Use of a Quitline

A greater proportion of smokers who had used a quitline before were "quite a bit" or "very much" interested in pharmacotherapy (49/63, 78%), than the proportion who had never before used a quitline (284/439, 65%; n = 502, $\chi^2_1 = 4.2$, *P* = .03, *V* = .09). In addition, smokers who had used a quitline before were much more likely to be "quite a bit" or "very much" interested in interpersonal–supportive methods (13/25, 52%) than smokers who had never before used a quitline (38/234, 16%; n = 259, $\chi^2_1 = 18.3$, *P* < .001, *V* = .27) or smokers for whom the question was not applicable (because they had never before tried to quit) (9/33, 27%; n = 58, $\chi^2_1 = 3.7$, *P* = .05, *V* = .25).

Cigarettes Per Day

Interest in pharmacotherapy was significantly associated with smoking level. Among lighter smokers (<10 cigarettes per day), 57% (63/111) were "quite a bit" or "very much" interested in pharmacotherapy, whereas a greater proportion (132/188, 70%)

of smokers of between 10 and 19 cigarettes per day (n = 299, $\chi^2_1 = 5.6$, $P = .01$, $V = .14$) or between 20 and 29 cigarettes per day (118/164, 72%) were similarly highly interested (n = 275, $\chi^2_1 = 6.7$, $P = .007$, $V = .16$). The very heaviest smokers, however, were somewhat less interested in pharmacotherapy than moderate smokers, although the differences were marginally significant. Specifically, 59% (30/51) of participants who smoked 30 or more cigarettes per day were “quite a bit” or “very much” interested in pharmacotherapy, compared with the 70% (132/188) of those smoking between 10 and 19 cigarettes per day (n = 239, $\chi^2_1 = 2.4$, $P = .09$, $V = .10$), or the 72% (118/164) of those smoking between 20 and 29 cigarettes per day (n = 215, $\chi^2_1 = 3.1$, $P = .06$, $V = .12$).

Although the MANOVA and univariate tests indicated a significant relationship between cigarettes per day and interest in interpersonal–supportive methods, none of the pairwise Tukey test comparisons were significant (all $P > .05$). These results are thus not reported.

Age Group

A significantly smaller proportion of younger smokers (≤25 years) were “quite a bit” or “very much” interested in pharmacotherapy (51/88, 58%) than the proportion (163/220, 74%) of smokers in the 26 to 40 age group (n = 308, $\chi^2_1 = 7.2$, $P = .005$, $V = .16$) who were similarly interested.

The univariate F test for the relationship between age group and interest in nonsocial–informational methods was marginally significant ($P < .06$) but Tukey pairwise comparisons indicated significant differences between age groups (all $P < .05$) that were corroborated by chi-square analyses. Fewer younger smokers (≤25 years) were “quite a bit” or “very much” interested in nonsocial–informational methods (32/56, 57%) than all other age groups: 74% (111/151) of smokers 26–40 years (n = 207, $\chi^2_1 = 5.1$, $P = .02$, $V = .16$), 79% (102/130) of smokers 41–55 years (n = 186, $\chi^2_1 = 8.8$, $P = .003$, $V = .22$), and 83% (25/30) of smokers aged ≥56 years (n = 86, $\chi^2_1 = 6.0$, $P = .01$, $V = .26$).

Race

White and black/African American smokers differed in their interest in interpersonal–supportive methods. While only 18% (44/242) of white smokers were “quite a bit” or “very much” interested in interpersonal–supportive methods, a much greater percentage (7/15, 47%) of black/African American smokers were similarly interested (n = 257, $\chi^2_1 = 7.2$, $P = .01$, $V = .17$). These results should be interpreted cautiously, however, due to the small sample of African American smokers contributing to these analyses.

Education

Although the univariate F test for the relationship between education level and interest in pharmacotherapy was marginally significant ($P < .10$), Tukey pairwise comparisons indicated significant differences between education levels that were corroborated by chi-square analyses. These indicated that a smaller proportion of smokers who had completed grade 11 or less were “quite a bit” or “very much” interested in pharmacotherapy (12/34, 35%) than the proportion of smokers who had achieved higher levels of education; specifically high school or its equivalent (87/122, 71%; n = 156, $\chi^2_1 = 14.9$, $P < .001$, $V = .31$), some college (161/237, 68%; n = 271, $\chi^2_1 = 13.7$, $P < .001$, $V = .23$), or graduation from college or higher (124/182, 68%; n = 216, $\chi^2_1 = 13.2$, $P < .001$, $V = .25$).

Discussion

Increasing the demand for cessation products and services will lead to more quit attempts, higher cessation rates, and greater reductions in smoking prevalence [22]. Smokers who seek cessation information on the Internet are a large group of smokers [11] who can be targeted and exposed to online messages that could potentially increase demand for evidence-based cessation products. Little is known, however, about what these smokers need or want to help them quit, or the reasons for or predictors of their preferences for various treatments. The current study asked smokers who visited the website of the ACS’s GASO for cessation information to rate their level of interest in evidence-based cessation products and services. Our results for predictors of interest in cessation methods were obtained controlling for all other variables including gender, knowledge of availability of free quitlines, and past-year attempts.

Cluster analysis of smokers’ ratings suggested that smokers’ interest in behavioral treatments centered on the degree of social or interpersonal involvement or social support the treatment would provide. Specifically, one cluster we obtained comprised products and services that involve high levels of interpersonal interaction and/or support. These included group-support cessation programs, telephone counseling, and using the Internet to chat with other smokers trying to quit. A contrasting cluster consisted of cessation methods that would provide tailored or individualized information on cessation, but that would *not* involve interpersonal interaction. The latter comprised tailored emails timed around a quit date, a website providing tailored cessation information, or booklets with information on quitting. Among the current sample of smokers, these nonsocial–informational programs were preferred to a significantly greater degree than methods requiring interpersonal interaction or support.

A separate cluster was also obtained that, as hypothesized, consisted of pharmacotherapies for cessation. This result suggests that motivation to use pharmacotherapy extends to both nicotinic and non-nicotinic medications. We had also hypothesized that interest in pharmacotherapy would be greater than interest in behavioral methods. We found, however, that smokers’ interest in medications as a whole was comparable with their interest in nonsocial–informational methods. Interest in interpersonal–supportive methods received lower ratings than pharmacotherapy or nonsocial–informational methods. These results are consistent with Cobb and Graham’s [11] finding that Internet seekers of cessation information believed that tailored information would be more helpful than telephone counseling or support received through Internet chat or forums. The present study builds on their findings by (1) suggesting that the level of interpersonal interaction and social support involved in these treatments makes them less appealing to Internet seekers of cessation information, and (2) examining sociodemographic and other factors associated with interest in interpersonal–supportive, nonsocial–informational, and pharmacological methods for cessation.

As noted earlier, compared with traditional media such as radio or television advertising, online advertising has been shown to be a more cost-efficient mode of recruiting smokers to Internet and telephone-based cessation treatments [12]. Moreover, behavioral targeting can be conducted whereby marketers target online consumers who fit particular sociodemographic criteria based on histories of websites visited and other publicly available information. These strategies can be used to promote cessation products and services to smokers seeking assistance via the Internet. Our results suggest, however, that this is more likely to be successful if the focus is on nonsocial–informational methods and/or pharmacotherapy. The low interest in interpersonal–supportive methods among online seekers of cessation information, however, suggests that methods involving interpersonal interaction and/or support would need to be carefully marketed. They should also consider the various smoker characteristics associated with greater interest in these methods.

Pharmacotherapy

Our results indicated that more smokers who were quitting the next day were not at all interested in pharmacotherapy compared with smokers quitting later on. Smokers planning to quit the next day most likely have made up their minds about quitting right away. They therefore may not want to spend the time to learn about and choose a medication, go to the drugstore, or ask a doctor to write a prescription. However, given the efficacy of pharmacotherapy, efforts might be aimed at these smokers to set a quit date with enough time to consider pharmacotherapy.

As hypothesized, results indicated that a greater proportion of older smokers (26–40 years) than younger smokers (≤25 years) were strongly interested in pharmacotherapy. This suggests that older smokers would be more receptive than younger smokers to efforts that encourage pharmacotherapy use. It is not immediately clear why a smaller proportion of younger smokers were interested, but one possibility is that more of them may believe they cannot afford medications. Alternatively, younger smokers may hold myths about medications that older smokers know are not true. Another possibility is that older smokers are more accustomed to taking medications for various health conditions. These possibilities can be examined in future research addressing preferences for cessation methods. The affordability of medications, or myths about them, may also explain why a much smaller proportion of less educated smokers were interested in pharmacotherapy than more educated smokers. Understanding the main reasons for less educated smokers' reluctance to use pharmacotherapy through further research would be important in devising strategies to increase demand for its use in this population.

Strategies to increase demand for pharmacotherapy should also consider smoking level, as the number of cigarettes smoked per day was associated with interest in pharmacotherapy. Fewer of the lightest smokers (defined as smoking <10 cigarettes per day) and fewer of the heaviest smokers (≥30 cigarettes per day) were interested in pharmacotherapy than were those who smoked between 10 and 29 cigarettes per day. A possible explanation is that lighter smokers may be more likely to believe that they are not so addicted that they need pharmacotherapy to help them quit. In contrast, the heaviest smokers may be more likely to believe that they are so addicted that even medications cannot help them quit. If evidence is obtained supporting these reasons, they can be addressed in online messages to increase demand for pharmacotherapy such as NRT.

Based on the results obtained, behavioral marketing could potentially be used to increase demand for NRT by targeting smokers 26–40 years old who have at least a high school education (as these individuals appear to be most interested in pharmacotherapy). Messages for these individuals could emphasize the ease of purchasing NRT, compare its cost relative to continued smoking, clarify concerns about using nicotine for treatment, and note its effectiveness relative to no medications. Focus-group research would be helpful in determining the precise content of messages that would resonate most for this demographic segment.

Interpersonal–Supportive Methods

In general, smokers were less interested in interpersonal–supportive cessation methods than in pharmacotherapy or nonsocial–informational methods. Moreover, interest in interpersonal–supportive methods could not be explained by whether smokers knew that free quitlines were available (as results were no different when this variable was controlled). In spite of the generally lower interest in interpersonal–supportive methods, however, results indicated that there were differences in interest as a function of quit date and whether smokers had previously used a quitline. Specifically, more smokers planning to quit "in the next week or two" were interested in interpersonal–supportive methods than were smokers who planned to quit at some undecided time in the future. This could be interpreted as a function of the greater seriousness about quitting among smokers planning to quit in the next week or two than among smokers who simply report planning to quit at some undecided time in the future.

Smokers who had previously used a quitline were also much more likely to be interested in interpersonal–supportive methods than smokers who had never before called a quitline. Smokers who had previously used a quitline may have had a positive enough experience to consider using an interpersonal–supportive method again. Alternatively, these methods may appeal to these smokers because of preexisting personal characteristics. For example, smokers higher on the personality trait of extraversion may find social interactional methods more appealing.

We had hypothesized that male smokers would be less interested than female smokers in interpersonal–supportive methods. Gender, however, was unrelated to interest in these treatments or, for that matter, in the two other types of treatments. Previous research has found gender differences among young adult smokers in the seeking of help from a professional [15]. Female smokers are also more frequent users of quitline services [14].

The lack of gender differences in *interest* in interpersonal–supportive methods in this study suggests that there may be other reasons why men have not actually used these methods as frequently as women. One possibility is that men may not be as knowledgeable about their existence. In this study, however, men and women did not differ in their knowledge about the availability of free quitlines. Alternatively, perhaps men are intrinsically interested in interpersonal–supportive methods but choose to not use them because of concerns about conforming to gender stereotypes of masculinity and help-seeking. Future research testing this hypothesis would be useful. Support for this hypothesis would suggest that online marketing messages directed to male smokers that also address the issue of masculinity might be able to encourage their greater use of interpersonal–supportive methods.

Overall, these results indicate that it may take a greater amount of effort to persuade smokers using the Internet for cessation advice to use interpersonal–supportive methods such as quitlines, group counseling, or Internet forums for peer support. However, given the demonstrated efficacy of quitlines, efforts should be made to encourage these options, at least among those smokers who would be more receptive (ie, smokers planning to quit earlier and who may have used a quitline before).

Nonsocial–Informational Methods

Nonsocial–informational methods include Internet-based treatments for tobacco use, which reviews and meta-analyses have concluded are effective compared with minimal or no treatments [23,24]. These and other Internet-based methods are typically free. Moreover, they can affect prevalence at the population level because large numbers of smokers can be reached for a very low cost.

Only age and when smokers were planning to quit were significantly related to interest in nonsocial–informational methods. While among all age groups most were "quite a bit" or "very much" interested in these methods, a greater majority of *older* smokers expressed strong interest. This counters the perception that older individuals are not receptive to newer technologies such as Web-delivered treatments or tailored email-based cessation programs. At the time that smokers are looking for cessation information on the Internet it seems logical to promote and provide access to these methods. Such a strategy may be most effective if targeted toward older smokers and to smokers who are a week or two away from their planned quit date.

Limitations

The sample of individuals in this study may not be representative of the population of smokers who use the Internet for cessation advice. Nevertheless, some of our results for interest in cessation methods parallel actual use found by studies using different recruitment methods. For example, adult smokers in the 2005 NHIS reported a low rate of having used behavioral treatments (which included in-person or telephone counseling or group cessation programs) compared with pharmacotherapy. This is consistent with the current study's finding of greater interest in pharmacotherapy than in behavioral methods. The NHIS also found that young adult smokers were less likely than older smokers to use pharmacotherapy [14]. Consistent with this we found that a significantly greater proportion of older smokers than younger smokers were "quite a bit" or "very much" interested in using pharmacotherapy. In addition, our results are similar to Cobb and Graham's finding that a greater proportion of Internet seekers of cessation information were women [11]. That study also found that telephone counseling and support received through Internet chat or forums were perceived as treatments that would not be helpful. In addition, although Cobb and Graham did not assess education levels, the higher levels of education noted in the current sample is consistent with that of users of the Internet in general [25], the population of interest in the current study. Nevertheless, replication of our findings would add confidence to our results for smokers' preferences and factors associated with them. In addition, our finding suggesting that black smokers were more interested than white smokers in interpersonal–supportive methods for quitting should be further investigated. Replication of this result would justify devoting extra resources to encourage the use of quitlines and other interpersonal–supportive methods among smokers from this community.

It is also not certain whether interest in a type of cessation method would translate into actual use of that method if it were made available. The aim of the study, however, was not to predict actual use but rather to first understand the basis of smokers' preferences. This is information that can then be used to increase adoption of evidenced-based treatments. Self-report is a reasonable approach to understanding these preferences and is typically a first step in market research. Future research may then focus on how to increase demand based on knowledge of these preferences by developing appropriate messaging for particular segments of consumers. Making preferred treatments appealing and available at the time smokers are seeking cessation advice on the Internet would help translate preferences to actual behavior, especially given the convenience of the Internet for purchasing products and services. There is also no reason to believe that ratings of preferences for particular cessation methods would be subject to social desirability biases. In the case of tobacco use behavior, self-reports have been shown to be reliable and valid [26], including from online surveys [27]. The validity of self-reports of tobacco behavior in online surveys is attributed to the lack of incentive to present oneself in a favorable light. In contrast, in a clinical setting where smokers are face to face with researchers, social desirability biases are greater [26]. In the current study, smokers planning to quit completed an anonymous survey with no incentives.

Due to concerns for brevity, the study was unable to test specific underlying reasons for associations between predictors and interest in treatments. For example, why a greater proportion of older smokers were strongly interested in nonsocial–informational cessation methods than were younger smokers could not be answered by the present study. Understanding smokers' reasons behind their interest (or noninterest) in cessation methods would be a fruitful topic for further research. Results of that research would help in refining strategies to promote use of cessation products and services for all sociodemographic groups. Evaluating smokers' interest in cessation aids with no evidence of effectiveness (eg, hypnosis, acupuncture) would also be informative.

Conclusions

Smokers who want to quit have available to them several effective cessation products and services. Demand for these is relatively low, however. Fortunately, many smokers access the Internet to help them quit, so targeting these smokers to promote cessation aids is a potentially effective way of increasing demand. Results of the current study indicated that they have relatively greater interest in pharmacotherapy and in cessation methods that provide tailored information or advice, but that do *not* involve interpersonal or socially supportive interactions. Moreover, the study indicated that smokers' level of cigarette consumption, when they were planning to quit, and sociodemographic variables were all associated with level of interest in using these treatments. Future research investigating the causes of interest in evidence-based treatments and whether targeted messages can encourage use among different groups of smokers will be an important step in understanding how to increase demand for treatments among smokers who are Internet users.

Conflicts of Interest

None declared

References

1. Fiore MC, Jaen CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ. Treating Tobacco Use and Dependence: 2008 Update. Clinical Practice Guideline. Rockville, MD: Agency for Health Care Policy and Research (US); 2008.
2. Cokkinides VE, Ward E, Jemal A, Thun MJ. Under-use of smoking-cessation treatments: results from the National Health Interview Survey, 2000. Am J Prev Med 2005 Jan;28(1):119-122. [doi: 10.1016/j.amepre.2004.09.007] [Medline: 15626567]
3. North American Quitline Consortium. NAQC. 2009. Measuring Reach of Quitline Programs. Quality Improvement Initiative URL: http://www.naquitline.org/resource/resmgr/docs/naqc_issuepaper_measuringrea.pdf [accessed 2010-06-12] [WebCite Cache ID 5rOJw40MM]
4. Solomon LJ, Hughes JR, Livingston A, Naud S, Callas PW, Peters EN, et al. Cognitive barriers to calling a smoking quitline. Nicotine Tob Res 2009 Nov;11(11):1339-1346. [doi: 10.1093/ntr/ntp143] [Medline: 19793785]
5. Gross B, Brose L, Schumann A, Ulbricht S, Meyer C, Völzke H, et al. Reasons for not using smoking cessation aids. BMC Public Health 2008;8:129 [FREE Full text] [doi: 10.1186/1471-2458-8-129] [Medline: 18430206]
6. Edwards R, McElduff P, Jenner D, Heller RF, Langley J. Smoking, smoking cessation, and use of smoking cessation aids and support services in South Derbyshire, England. Public Health 2007 May;121(5):321-332. [doi: 10.1016/j.puhe.2006.11.002] [Medline: 17339045]
7. Hammond D, McDonald PW, Fong GT, Borland R. Do smokers know how to quit? Knowledge and perceived effectiveness of cessation assistance as predictors of cessation behaviour. Addiction 2004 Aug;99(8):1042-1048. [doi: 10.1111/j.1360-0443.2004.00754.x] [Medline: 15265101]
8. Backinger CL, Thornton-Bullock A, Miner C, Orleans CT, Siener K, DiClemente CC, et al. Building consumer demand for tobacco-cessation products and services: the national tobacco cessation collaborative's consumer demand roundtable. Am J Prev Med 2010 Mar;38(3 Suppl):S307-S311. [doi: 10.1016/j.amepre.2009.12.002] [Medline: 20176301]
9. Fox S. Pew Internet. Washington, DC: Pew Internet & American Life Project; 2006 Oct 29. Online health search 2006 URL: http://www.pewinternet.org/~/media/Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf [accessed 2010-06-25] [WebCite Cache ID 5qlBlCMNh]
10. Punj G, Stewart DW. Cluster analysis in marketing research: review and suggestions for application. J Marketing Res 1983;20(May):134-148 [FREE Full text]
11. Cobb NK, Graham AL. Characterizing Internet searchers of smoking cessation information. J Med Internet Res 2006;8(3):e17 [FREE Full text] [doi: 10.2196/jmir.8.3.e17] [Medline: 17032633]
12. Graham AL, Milner P, Saul JE, Pfaff L. Online advertising as a public health and recruitment tool: comparison of different media campaigns to increase demand for smoking cessation interventions. J Med Internet Res 2008;10(5):e50 [FREE Full text] [doi: 10.2196/jmir.1001] [Medline: 19073542]
13. Atlanta Journal-Constitution. AJC. 2010. Digital targeting URL: http://ajcmediakit.com/digital/digital-targeting.html [accessed 2010-07-07] [WebCite Cache ID 5r30qll3Y]
14. Curry SJ, Sporer AK, Pugach O, Campbell RT, Emery S. Use of tobacco cessation treatments among young adult smokers: 2005 National Health Interview Survey. Am J Public Health 2007 Aug;97(8):1464-1469 [FREE Full text] [doi: 10.2105/AJPH.2006.103788] [Medline: 17600243]
15. Centers for Disease Control and Prevention (CDC). Use of cessation methods among smokers aged 16-24 years--United States, 2003. MMWR Morb Mortal Wkly Rep 2006 Dec 22;55(50):1351-1354 [FREE Full text] [Medline: 17183227]
16. North American Quitline Consortium. NAQC. 2009. Tobacco Cessation Quitlines: A Good Investment to Save Lives, Decrease Direct Medical Costs and Increase Productivity URL: http://www.naquitline.org/resource/resmgr/docs/naqc_issuepaper_tobaccocessa.pdf [accessed 2011-06-03] [WebCite Cache ID 5zAgEIklh]

17. Pederson EL, Vogel DL. Male gender role conflict and willingness to seek counseling: Testing a mediation model on college-aged men. J Couns Psychol 2007;54(4):373-384. [doi: 10.1037/0022-0167.54.4.373]
18. Jones S, Fox S. Pew Internet Project Data Memo. Washington, DC: Pew Internet & American Life Project; 2009 Jan 28. Generations online in 2009 URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Generations_2009.pdf [accessed 2011-07-29] [WebCite Cache ID 60XomJ6j3]
19. Aldenderfer MS, Blashfield RK. Cluster Analysis. Beverly Hills, CA: Sage Publications; 1984.
20. Agresti A. Categorical Data Analysis, 2nd edition. New York, NY: Wiley-Interscience; 2002.
21. Cramér H. Mathematical Methods of Statistics. Princeton, NJ: Princeton University Press; 1999.
22. Levy DT, Mabry PL, Graham AL, Orleans CT, Abrams DB. Reaching Healthy People 2010 by 2013: A SimSmoke simulation. Am J Prev Med 2010 Mar;38(3 Suppl):S373-S381. [doi: 10.1016/j.amepre.2009.11.018] [Medline: 20176310]
23. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937 [FREE Full text] [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
24. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]
25. Pew Internet. Washington, DC: Pew Internet & American Life Project; 2011. Demographics of Internet users URL: http://www.pewinternet.org/Static-Pages/Trend-Data/Whos-Online.aspx [accessed 2011-06-03] [WebCite Cache ID 5zAhWCjub]
26. Velicer WF, Prochaska JO, Rossi JS, Snow MG. Assessing outcome in smoking cessation studies. Psychol Bull 1992 Jan;111(1):23-41. [Medline: 1539088]
27. Ramo DE, Hall SM, Prochaska JJ. Reliability and validity of self-reported smoking in an anonymous online survey with young adults. Health Psychol 2011 May 16. [doi: 10.1037/a0023443] [Medline: 21574709]

Abbreviations

ACS: American Cancer Society
GASO: Great American Smokeout
MANOVA: multivariate analysis of variance
NHIS: National Health Interview Survey
NRT: nicotine replacement therapy

Edited by G Eysenbach; submitted 28.10.10; peer-reviewed by A Graham, B Bock; comments to author 20.04.11; revised version received 05.07.11; accepted 06.07.11; published 26.08.11.

Please cite as:
Westmaas JL, Abroms L, Bontemps-Jones J, Bauer JE, Bade J
Using the Internet to Understand Smokers' Treatment Preferences: Informing Strategies to Increase Demand
J Med Internet Res 2011;13(3):e58
URL: http://www.jmir.org/2011/3/e58/
doi:10.2196/jmir.1666
PMID:21873150

Original Paper

A Systematic Review of Information and Communication Technology–Based Interventions for Promoting Physical Activity Behavior Change in Children and Adolescents

Patrick WC Lau[1], PhD; Erica Y Lau[1], Mphil; Del P Wong[2], PhD; Lynda Ransdell[3], PhD

[1]Department of Physical Education, Hong Kong Baptist University, Hong Kong, China
[2]Department of Health and Physical Education, Hong Kong Institute of Education, Hong Kong, China
[3]Department of Kinesiology, College of Education, Boise State University, Boise, United States

Corresponding Author:
Patrick WC Lau, PhD
Department of Physical Education
Hong Kong Baptist University
DLB 643, David C. Lam Bldg,
Shaw Campus, HKBU
Hong Kong, 0000
China
Phone: 852 34115634
Fax: 852 34115757
Email: wclau@hkbu.edu.hk

Abstract

Background: A growing body of research has employed information and communication technologies (ICTs) such as the Internet and mobile phones for disseminating physical activity (PA) interventions with young populations. Although several systematic reviews have documented the effects of ICT-based interventions on PA behavior, very few have focused on children and adolescents specifically.

Objectives: The present review aimed to systematically evaluate the efficacy and methodological quality of ICT-based PA interventions for children and adolescents based on evidence from randomized controlled trials.

Methods: Electronic databases Medline, PsycInfo, CINAHL, and Web of Science were searched to retrieve English language articles published in international academic peer-reviewed journals from January 1, 1997, through December 31, 2009. Included were articles that provided descriptions of interventions designed to improve PA-related cognitive, psychosocial, and behavioral outcomes and that used randomized controlled trial design, included only children (6-12 years old) and adolescents (13-18 years old) in both intervention and control groups, and employed Internet, email, and/or short message services (SMS, also known as *text messaging*) as one or more major or assistive modes to deliver the intervention.

Results: In total, 9 studies were analyzed in the present review. All studies were published after 2000 and conducted in Western countries. Of the 9 studies, 7 demonstrated positive and significant within-group differences in at least one psychosocial or behavioral PA outcome. In all, 3 studies reported positive and significant between-group differences favoring the ICT group. When between-group differences were compared across studies, effect sizes were small in 6 studies and large in 3 studies. With respect to methodological quality, 7 of the 9 studies had good methodological quality. Failure to report allocation concealment, blinding to outcome assessment, and lack of long-term follow-up were the criteria met by the fewest studies. In addition, 5 studies measured the intervention exposure rate and only 1 study employed objective measures to record data.

Conclusion: The present review provides evidence supporting the positive effects of ICTs in PA interventions for children and adolescents, especially when used with other delivery approaches (ie, face-to-face). Because ICT delivery approaches are often mixed with other approaches and these studies sometimes lack a comparable control group, additional research is needed to establish the true independent effects of ICT as an intervention delivery mode. Although two-thirds of the studies demonstrated satisfactory methodological quality, several quality criteria should be considered in future studies: clear descriptions of allocation concealment and blinding of outcome assessment, extension of intervention duration, and employment of objective measures in intervention exposure rate. Due to the small number of studies that met inclusion criteria and the lack of consistent evidence, researchers should be cautious when interpreting the findings of the present review.

(J Med Internet Res 2011;13(3):e48) doi:10.2196/jmir.1533

KEYWORDS

Internet; email; text messages

Introduction

Regular physical activity (PA) is associated with reduced risk of breast cancer, hypertension, coronary heart disease, type 2 diabetes mellitus, obesity, and osteoporosis in children and adolescents [1-3]. However, the majority of our young population is not engaging in sufficient PA to achieve these health benefits [4-7]. Developing effective interventions to promote active lifestyles among children and adolescents is one way to address the lack of PA in this population.

In the past decade, numerous PA interventions have been developed and implemented [8,9]. The typical modes of delivery for these interventions have been face-to-face and mass media [10,11]. Studies [12-14] have indicated that interventions delivered using a face-to-face approach (ie, structured PA programs and individual counseling) have been effective for PA behavior changes, but effects have been small. Due to time schedules, high running costs, and geographic restrictions, these interventions could not reach and be accessed by a large population [15-18]. With the combination of small effects and limited reach, the impact of face-to-face PA interventions on public health has been modest [19]. On the other hand, PA interventions disseminated through mass media approaches (ie, TV, radio, and pamphlets) have the potential to reach large numbers of individuals. One limitation of mass media–based PA interventions is that they mainly contain generic content and feedback that is less relevant to individuals who may need different strategies to change PA behavior [10]. The aforementioned limitation probably explains why the majority of these mass media–based interventions (except the VERB campaign [20]) were only successful in raising awareness and increasing knowledge and not in improving PA behavioral outcomes (ie, increasing PA levels) [21]. In sum, innovative approaches that can reach large groups of people while at the same time enhancing accessibility and personal relevance are needed. Interestingly, studies have shown that this lack of personalization, mentioned as a limitation of media-based PA interventions, could be addressed with the use of advanced information and communication technology (ICT) such as Internet, personal digital assistants, computer kiosks, and mobile phones [22,23].

Advantages of ICT-based interventions

The proliferation of the Internet and mobile phones has provided a powerful channel to widen the reach of PA interventions in children and adolescents [23]. In developed countries, over 90% of children and adolescents were found to have had access to the Internet at school and/or home [14]. More importantly, they perceived the Internet as their primary resource for seeking health information [24,25]. Additionally, 45% to 99% of children and adolescents have been found to own a personal mobile phone, and half of them use short message services (SMS, also known as *text messaging*) [10]. In addition to this broadened reach, advanced Web technologies can enhance the personal relevance of an intervention's contents. Researchers can now tailor PA interventions based on a variety of factors that influence PA behavior change in children and adolescents (eg, gender, ethnicity, weight status, stage of change, PA self-efficacy, and PA barriers). In addition, ICT interventions can present materials in various forms (ie, text, sound, video, and animation) to satisfy children’s and adolescents’ preferences [22]. Moreover, email and SMS have provided a means for researchers to deliver individualized feedback, automatic reminders, and social support. These elements could enhance children’s and adolescents’ attention toward and understanding of the materials [26-28], which could lead to subsequent improvements in PA behavior.

In recent years, the evidence base of ICT-based health behavior change interventions has been growing [29-31]. Several systematic reviews [26,28,32-36] have evaluated the efficacy of these interventions; however, very few studies have focused on PA behavior. Even though some studies have concentrated specifically on PA behavior [32,33,35,36], none has focused on children and adolescents. To our knowledge, the systematic review conducted by Norman et al [33] was the only paper published in the previous 5 years that has documented the efficacy of ICT-based PA interventions in children and adolescents. These authors conducted an electronic database search through the year 2005, and they identified 33 studies with PA as an outcome. As ICT-based PA interventions were still in a development stage at that time, studies that focused on children and adolescents were scant. These authors located 5 studies that focused on children and adolescents, and only 3 of them were randomized controlled trials (RCT). Although Norman and colleagues concluded that ICT-based interventions were effective for changing PA behavior, this conclusion was mainly drawn using evidence collected with adult subjects. Whether their conclusions can be generalized to children and adolescents needs further investigation. Since 2005, several additional RCTs have been published. It is, therefore, timely to conduct a new systematic review to evaluate the efficacy and methodological quality of ICT-based interventions relative to promoting PA behavior change in children and adolescents.

Although there are various ICTs, the present review focused on the Internet and SMS only. The reason for limiting the scope of ICT to Internet, email, and SMS is that these modes are most frequently used among children and adolescents. Other ICTs such as interactive CD-ROMs and computer kiosks were excluded because they are less popular, and, in that format, the advantages of ICT-based interventions (eg, free of time and geographic restrictions) are not fully utilized. Moreover, ICT-based PA interventions for children and adolescents are still in a developing stage; a systematic review focused on both Internet and SMS should demonstrate their usefulness in various research designs. This could inform the choice of ICT in future studies.

The purpose of the present review was to systematically evaluate the efficacy and methodological quality of ICT-based interventions that applied Internet and/or SMS as a delivery mode for PA behavior change in children and adolescents based on evidence from randomized controlled trials.

Definitions

In this review, an ICT-based intervention is defined as an intervention that employs Internet, email, and/or SMS as one of the intervention delivery modes. The following types of intervention are excluded from this definition: (1) interventions that only involved ICT for data collection (ie, online surveys and electronic medical records) and (2) interventions that used a computer to generate individually tailored printed materials and delivered those materials using a non-ICT mode. The aforementioned interventions were excluded because there was little or no information exchanged, and interactions between the ICTs and participants were minimal.

Methods

An electronic database search was conducted to retrieve English articles from CINAHL, Medline, PsycInfo, PubMed, and Web of Science. The search targeted articles published from January 1, 1997, through December 31, 2009, because ICT-based interventions began in the late 1990s [11]. For CINAHL, Medline, PsycInfo, and Web of Science, we performed a keyword search using the following search strings: (child* OR adolescent* OR teenag* OR youth) AND (Internet OR Web-based OR Web-delivered messages OR email OR e-mail OR electronic mail OR mobile phone OR text messag* OR SMS) AND (daily physical activit* OR exercise OR walk* OR motor activ* OR leisure activit* OR physical fitness OR sport*) AND (health OR health behavior OR weight loss OR obesity OR overweight). In addition, we conducted a MeSH search in PubMed using the following search strings: ("Adolescent"[Mesh] OR "Child"[Mesh]) AND ("Internet"[Mesh] OR "Telecommunications"[Mesh]) AND ("Exercise"[Mesh] OR "Motor Activity"[Mesh] OR "Sports"[Mesh]) AND ("Health Behavior"[Mesh] OR "Obesity"[Mesh] OR "Weight Loss"[Mesh]).

Selection Criteria

To be included, articles had to (1) be published in international academic peer-reviewed journals (book chapters, abstracts of conference proceeding, and dissertations were excluded); (2) use a randomized controlled trials design; (3) evaluate an intervention that aimed to promote PA behavior; (4) include at least one PA behavior variable as the outcome (no restriction was defined regarding the types of PA behavior outcomes, which could be cognitive [ie, PA knowledge], psychosocial [eg, PA intention, PA self-efficacy, social support to PA, stage of change], or behavioral [ie, energy expenditure, step counts, or self-reported PA level]; (5) focus only on children (6-12 years old) and adolescents (13-18 years old) in both the intervention and control group; and (6) employ Internet, email, and/or SMS as one or more major or assistive modes to deliver the intervention. No further limits were set on the types and content of the control group. Control groups were non–ICT-based, no treatment, or different types of ICT-based interventions.

To attain additional eligible articles, the reference list of the located studies and relevant reviews were also checked. The selection of articles was independently performed by two investigators (authors PWCL and EYL).

Data Extraction

The present review provided a narrative evaluation of the selected articles because of the heterogeneity in study designs, measures, and outcomes across studies. Information about the selected articles was extracted into a structured summary table by one investigator (EYL) and checked by another investigator (PWCL). The following data were extracted: source (year of publication, country in which study was conducted); study characteristics (study design, setting that the information was delivered via ICT, and target behavior); participant characteristics (sample size, age, and group of participants); intervention descriptions, intervention characteristics (intervention duration, mode of delivery, contact frequency, theoretical basis, types and numbers of behavior change technique [BCT] used, and ICT initiation strategy). Use of BCT was coded according to the definition of the taxonomy of BCT developed by Abraham and Michie [37]. ICT initiation strategy was divided into participant-initiated (participants have to decide when, where, and what information to access or transmit by ICT) or investigator-initiated (an investigator delivered the information to participants via ICT at a fixed time, venue, and under specific conditions). For example, with a participant-initiated process, participants were told to access a PA website twice a week during their free time, but they were not told which day of the week or number of pages to read each time. For an investigator-initiated process, participants might receive a hyperlink via their personal email or individual feedback to their mobile phone.

The Efficacy of ICT-Based PA Interventions for Children and Adolescents

In addition, types of outcome measures and main findings were coded. The pre-post difference on PA behavior outcome in the intervention group was coded as "↑" for positive and significant change, "→" for no significant change and "↓" for significant negative change. The pre-post difference in PA behavior outcome between the intervention group and control group was coded as "+" (significant difference favoring the ICT intervention group), "O" (no significant difference between groups), and "—" (significant difference favoring the control group). To compare the potential effect of ICT interventions on children and adolescents across studies, an effect size (ES) was calculated when sufficient information was reported. An ES of less than 0.5 was interpreted as small, 0.5 to 0.8, as medium, and greater than 0.8 as large [38]. When a study measured outcomes across several time points, the longest follow-up was selected for effect size calculation. For example, if a study measured exercise behavior at 6, 12, 18, and 24 months the 24-month data were selected for comparison. For studies that employed more than one comparison group, following previous systematic reviews of ICT-based intervention [26,34], the control group with the least contact was selected for ease of interpretation. To ensure the accuracy of the data extraction, original authors of the included studies were contacted for further information and clarification when needed.

Assessment of Methodological Quality

Methodological quality was assessed using a 13-item scale developed in a previous review [35]. Studies were rated independently by one investigator (EYL) and checked by another investigator (PWCL). Disagreements were discussed until consensus was reached. Each item was rated as *yes*, *no*, or *unknown*. A total methodological quality score (ranging from 0-13) was calculated by summing up all *yes* items. Studies were rated as having good methodological quality if they met at least two-thirds of the criteria (ie, ≥ 9 items).

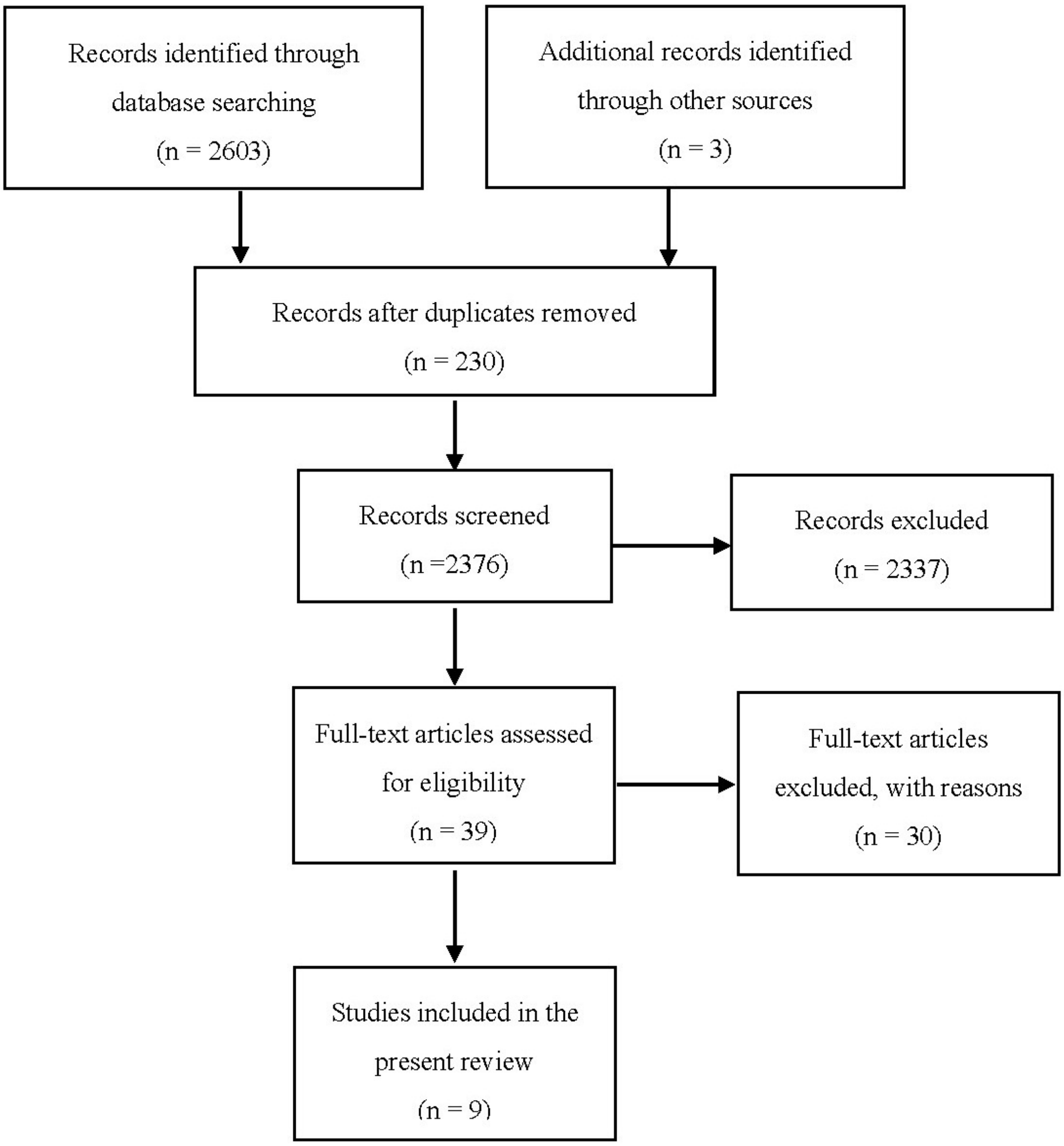
In addition, the intervention exposure rate of the included studies was also extracted because this was suggested as an important quality criterion of ICT-based interventions [39,40]. The present review also assessed whether the included studies measured exposure frequency (ie, how frequently the participants accessed materials via ICT) of the intervention. For those studies that involved an Internet program, exposure duration (ie, how much time did participants spend on reading materials via ICT each time?) was also extracted [40].

Results

Selection of Articles

The search and selection process for articles is illustrated in Figure 1. A total of 2606 articles were identified initially. After removing duplicates articles (n = 230) and irrelevant studies (n = 2337), 39 articles were retained for further considerations. Of these, 30 articles were excluded as they were descriptive or feasibility studies (n = 13), they were not targeted to the population of interest (n = 7), they did not use an RCT design (n = 6), they did not include any PA behavior outcome (n=3), or they did not use ICT as the mode of delivery (n = 1). In the end, 9 studies [41-49] were included in the present review.

Figure 1. Search and selection process for the articles



Data Extraction

Details regarding the 9 included studies are summarized in Table 1. All studies were published after 2000 and conducted in Western countries. Specifically, 6 studies were conducted in the United States [41,44,46-49], and 1 each was conducted in the United Kingdom [42], Australia [43], and New Zealand [45].

Study Characteristics

In all, 8 studies [41,43-49] randomized the participants on an individual basis, while one study [42] randomized at the school level. Another 7 studies [41-45,48,49] delivered the intervention through ICTs in a home setting, whereas the remaining studies disseminated the intervention at schools [47] or clinics [46]. In regard to the target behavior, 3 studies [42,44,45] focused on PA behavior alone; 6 studies [41,43,46-49] focused on PA together with other health behaviors (eg, sedentary behavior, dietary intake, and diabetes management) (See Table 1).

Participant Characteristics

The 9 included studies contained 1456 total participants and the sample sizes ranged from 57 to 473 in each study. Approximately 48.1% of the total sample participants were female (701/1456). In all, 6 studies included both genders [41,43,45-48], while 3 studies focused on a single gender: 2 studies [44,49] focused on females and 1 study focused on [42] males. In the majority of studies, the intervention was offered to healthy children and adolescents recruited from schools and communities (eg, a scout troop). Also, 3 studies focused on overweight [49] or diabetic patients [41,45] from pediatric clinics (See Table 1).

Intervention Characteristics

Intervention duration ranged from 2 weeks to 2 years. Of the 9 studies, 5 [42,44,45,47,48] focused on short-term results (≤ 3 months), 2 [43,46] focused on medium-term results (4-6 months), and 2 focused on long-term results (> 6 months) [41,49]. Regarding the mode of delivery, 8 studies [41-48] employed a single ICT mode (Internet = 4, email = 1, and SMS = 3) to deliver the intervention. The only exception was Williamson et al [49] who combined the use of Internet and email. The contacts made via the ICT mode varied from twice per day to once in 12 weeks. Also employed by 6 studies [41-43,45,46,48] were non-ICT modes (face-to-face, mail, and telephone) to contact the participants; the frequency of contact by non-ICT modes ranged from once per week to once every 3 months.

Of the 9 studies, 5 [41-43,46,47] had developed the intervention based on health behavior change theories: 3 [41-43] were guided by social cognitive theory (SCT) [50,51] and 2 [46,47] were developed based on SCT [50,51], transtheoretical model (TTM) [52], and relapse prevention model (RPM) [53]. Use of BCT in each study is summarized in Table 2. It was found that 3 types of BCT were used the most among the included studies: prompt specific goal setting, prompt self-monitoring of behavior, and provide feedback on performance. Overall, there was great variability in the number of BCTs used (the range was 3 to 9). In terms of the ICT initiation strategy, 5 studies [41,43,45-47] used investigator-initiated strategies and 4 studies [42,44,48,49] employed participant-initiated strategies.

Table 1. Overview of characteristics and main findings of included studies

Author, Year, and Country	Participant Characteristics	Study characteristics	Intervention Descriptions	Intervention Characteristics	Main Findings
Franklin et al, 2006, United States [41]	92 diabetic patients (8-18 years of age), 43 were female	Design: RCT Setting: home Focus: diabetes	ICT group 1: Sweet Talk SMS (automatic scheduled SMS reminder of the goal set, daily tips, and monthly text newsletter on diabetes issues. Participants could reply to the SMS and get an extra SMS for reply) plus intensive insulin treatment and standard treatment (clinic visit once per 3 to 4 months and access to emergency hotline). ICT group 2: Sweet Talk SMS in addition to conventional insulin treatment and standard treatment Control: standard treatment only	Duration: 1 year Mode and contact: 1 or 2 SMS per day[a]; 1 face-to-face every 3 to 4 months Theory: SCT Number of BCTs used: 8 Communication initiation: investigator	- Significantly greater increase in perceived social support to exercise in both ICT groups compared with control group - No significant difference between the 2 ICT groups
Jago et al, 2006, United Kingdom [42]	473 boy scouts (10-14 years), 0 were female	Design: cluster RCT Setting: home Focus: PA	ICT group 1: Internet-based PA program contained goal setting and a comic story on overcoming PA barriers plus face-to-face troop training started in the spring. ICT group 2: Internet-based PA program contained goal setting and a comic story on overcoming PA barriers plus face-to-face troop training started in the fall Control: dietary intervention plus face-to-face troop training	Duration: 9 weeks Mode and contact: 2 per week via Internet and 1 per week face-to-face Theory: SCT[a] Number of BCTs used: 8 Communication initiation: participant	- Significant increase in light PA in ICT group 1 - No significant between-group differences
Lubans et al, 2009, Australia [43]	124 school children (mean age 14.1), 71 were female	Design: RCT Setting: home Focus: PA and Diet	ICT: social support email from investigator and pedometer self-monitoring plus printed PA and nutrition handbook and printed monthly newsletter for parents plus face-to-face school-based sport education program (structured PA session focusing on lifetime activities) plus information session (weekly diet and PA messages with teacher-demonstrated related activities to reinforce the message and summary lectures) Control: face-to-face school-based sport education program plus exercise handbook	Duration: 6 months Mode and contact: 1 email per week[a], face-to-face contact once per week for the first 10 weeks Theory: SCT Number of BCTs used: 7 Communication initiation: investigator	- Significantly greater increase in step counts in ICT group compared with control for both boys and girls - Significantly greater increase in low active participants of ICT group compared with control group - No significant difference in step counts in active participants between ICT group and control group
Marks et al, 2006, United States [44]	319 girls (12-14 years of age), that is, all 319 female	Design: RCT Setting: home Focus: PA	ICT: Web-based PA program containing interactive games, quiz, downloadable charts to plan daily activity, and audio demonstration of PA activities Control: print-based PA program with the content identical to the Internet-based intervention	Duration: 2 weeks Mode and contacts: Internet 2 times per week Theory: none mentioned Number of BCTs used: 3 Communication initiation: participant	- Significant increase in PA self-efficacy and PA intention in both ICT and control groups - Significantly greater increase in PA intention in control group compared with ICT group but not in PA, self-efficacy and PA level

Author, Year, and Country	Participant Characteristics	Study characteristics	Intervention Descriptions	Intervention Characteristics	Main Findings
Newton et al, 2009, New Zealand [45]	78 diabetic patients (11-18 years), 42 were female	Design: RCT Setting: home Focus: PA	ICT: pedometer self-monitoring with a goal of at least 10,000 step/day plus motivational SMS reminder (participants could reply the SMS and get an extra SMS for reply[a]) plus standard diabetes treatment Control: standard diabetes treatment	Duration: 12 weeks Mode and contact: 1 or 2 SMS contacts per week Theory: none mentioned Number of BCTs used: 5 Communication initiation: investigator	• Both groups decreased in step counts, but this was not statistically significant. • No significant between-group differences
Patrick et al, 2001, United States [46]	117 adolescents (11-18 years, mean age 14.1), 43 were female	Design: RCT Setting: clinic Focus: PA plus diet	ICT: interactive computer program (assessed and compared participant's self-reported PA and diet behavior with recommendations, gave feedback, and instructed the participants to select two behaviors that they are most ready to change and construct an action plan) plus provider counseling (provider review and discussion of the action with the participants) plus extended follow-up by mail for group 1 and by infrequent telephone plus mail contact for group 2, and frequent telephone plus mail contact for group 3 Control: interactive computer program plus provider counseling but no further extended follow-up	Duration: 16 weeks Mode and contact: Internet, face-to-face and mail, each once during the 16 weeks plus telephone once every 2 weeks Theories: SCT, TTM, and RPM Number of BCTs used: 8 Communication initiation: investigator	• Significant improvement in moderate PA in all groups but no effect in vigorous PA • No significant between-group differences
Prochaska et al, 2004, United States [47]	138 school children (12-14 years), 90 were female	Design: RCT Setting: school Focus: PA plus diet	ICT group 1: one-session Internet-based PA assessment with tailored feedback ICT group 2: one-session Internet-based PA plus dietary assessment with tailored feedback Control: no treatment	Duration: 12 weeks Mode and contact: 1 Internet contact in the 12 week period Theories: SCT, TTM, and RPM Number of BCTs used: 9 Communication initiation: investigator	• Significantly greater increase in PA level in ICT groups compared with control groups for boys but not for girls

Author, Year, and Country	Participant Characteristics	Study characteristics	Intervention Descriptions	Intervention Characteristics	Main Findings
Shapiro et al, 2008, United States [48]	58 children (5-13 years), 36 were female	Design: RCT Setting: home Focus: PA plus diet plus sedentary activity	ICT: parent and child to report their sugar-sweetened beverages, screen time, and PA goals by SMS plus immediate and automatic SMS feedback plus 3 face-to-face psychologist-led educational sessions Control group 1: parent and child to report their sugar-sweetened beverages, screen time, and PA goals by using a paper diary plus 3 face-to-face psychologist-led educational sessions plus verbal feedback during the educational session Control group 2: 3 face-to-face psychologist-led educational sessions only	Duration: 8 weeks Mode and contact: 2 SMS per day, face-to-face once a week Theory: none mentioned Number of BCTs used: 7 Communication initiation: participants	• No significant difference in step counts in all groups
Williamson et al, 2006, United States [49]	57 overweight girls (11-15 years), that is all 57 were female	Design: RCT Setting: home Focus: weight loss	ICT: Internet-based PA and dietary program with tailored information and prescriptions and online counseling plus 4 face-to-face meetings at 1, 3, 6, and 12 weeks Control: Internet-based program with general PA and dietary information plus 4 face-to-face sessions at 1, 3, 6, and12 months	Duration: 2 years Mode and contact: Internet once per week, email once per week, and face-to-face 4 times in 12 weeks Theory: none mentioned Number of BCTs used: 7 Communication initiation: participant	• Significant improvement in self-reported exercise behavior in all groups • No significant between-group differences

[a] Information obtained from original author

Table 2. Use of behavior change techniques (indicated by Abraham and Michie [37])

Items	Franklin et al [41]	Jago et al [42]	Lubans et al [43]	Marks et al [44]	Newton et al [45]	Patrick et al [46]	Prochaska et al [47]	Shapiro et al [48]	Williamson et al [49]
Provide information about behavior-health link	X					X	X		X
Provide information on consequences				X		X	X		X
Provide information about other's approval									
Prompt intention formation						X	X		
Prompt barrier identification		X	X			X	X		X
Provide general encouragement								X	
Set graded tasks		X							
Provide instruction	X	X	X						
Model or demonstrate the behavior		X	X	X					
Prompt specific goal setting	X	X			X	X	X	X	X
Prompt review of behavioral goals		X			X	X	X	X	
Prompt self-monitoring of behavior	X	X	X	X				X	X
Provide feedback on performance			X		X	X	X	X	X
Provide contingent rewards	X	X			X[a]			X	
Teach to use prompts and cues									
Agree on behavioral contract	X		X						X
Prompt practice	X				X			X	
Use follow-up prompts									
Provide opportunities for social comparison									
Plan social support or social change	X		X			X	X		
Prompt identification as a role model.									
Prompt self-talk									
Relapse prevention									
Stress management									
Motivational interviewing									
Time management									

[a] Information obtained from original author

Intervention Efficacy

Table 3 illustrates the effects of ICT-based interventions on PA behavior outcomes. Changes in behavioral variables were reported in 7 studies [42-48], and 4 of these [42,43,46,47] demonstrated significant within-group differences. Changes in psychosocial variables were presented in 3 studies [41,44,49], and all demonstrated significant within-group differences.

In all, 7 studies [41-45,47,48] compared the effects between an ICT group and a non-ICT control group. Of these, 3 [41,43,47] reported a positive effect favoring the ICT group, and 1 study showed a positive effect favoring the non-ICT control group [44]. In addition, 2 studies [46,49] contrasted the effect between two ICT groups. They examined whether different tailoring levels and follow-up methods would affect intervention efficacy. There were no significant between-group differences. On average, ICT-based interventions had a small effect size (0.03 to 0.41) on PA behavior change when compared with the control group. Notable exceptions were studies by Franklin et al [41], Lubans et al [43], and Prochaska et al [47], who reported large effect sizes.

Assessment of Methodological Quality

In Table 4, the results of the methodological quality assessment are described. Of the 9 included studies, 7 [41-44,47-49] were rated as having good methodological quality. The low methodological quality scores were attributed to failure to report the concealment method for randomization, blinding of the assessors, and failure to follow-up long term. In all, 7 studies [41-44,46,47,49] measured intervention exposure frequency, while 5 studies [42,44,46,47,49] utilized an Internet program, and 3 of those studies [44,46,47] recorded Internet exposure duration.

Table 3. Effect of ICT-based intervention on PA outcomes

Sources		Effect[a]		
	Outcome Measure	Within-Group	Between-Group	Effect Size
Franklin et al [41]	Perceived social support to exercise	↑	+	0.76
Jago et al [42]	Light PA	↑	○	0.03
	Moderate PA	→	○	0.08
	Step counts	→	○	0.24
Lubans et al [43]	Step count, boys	↑	+	0.80
	Step count, girls	↑	+	1.2
Marks et al [44]	PA self-efficacy	↑	○	Not applicable
	PA intention	↑	—	0.41
	Self-reported PA	→	○	0.39
Newton et al [45]	Step counts	→	○	Not applicable
Patrick et al [46]	Moderate PA	↑	○	Not applicable
	Vigorous PA	→	○	Not applicable
Prochaska et al [47]	PA level, boys	↑	+	0.95
	PA level, girls	→	○	0.03
Shapiro et al. [48]	Self-reported PA	→	○	0.14
Williamson et al [49]	Self-reported exercise behavior	↑	○	Not applicable

[a] The pre-post difference in PA behavior outcome in the intervention group was indicated by: "↑" for positive and significant, "→" for no significant change and "↓" for significant negative change. The pre-post difference in PA behavior outcome between the intervention group and control group was coded as "+" (significant difference favoring the ICT intervention group), "O" (no significant difference between groups), and "—" (significant difference favoring the control group).

Table 4. Assessment of methodological quality and intervention exposures of the studies

Items	Franklin et al [41]	Jago et al [42]	Lubans et al [43]	Marks et al [44]	Newton et al [45]	Patrick et al [46]	Prochaska et al [47]	Shapiro et al [48]	Williamson et al [49]
Methodological quality									
Were the eligible criteria specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the method of randomization described?	Yes	No	Yes	Yes	No	No	No	Yes	Yes
Was the random allocation concealed? (ie, was the assignment generated by an independent person not responsible for determining eligibility of the participants)	Yes	Yes[a]	Yes	Unknown	Yes[a]	No[a]	Yes[a]	Yes[a]	No[a]
Were the groups similar at baseline regarding important prognostic indicators?	Yes	Yes[a]	Yes	Yes	No[a]	Yes	Yes	Yes	Yes
Were both the index and the control interventions explicitly described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the compliance or adherence with the interventions described?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Was the outcome assessor blinded to the interventions?	No[a]	No[a]	No[a]	Unknown	No	Yes[a]	Yes[a]	No[a]	No[a]
Was the dropout rate described, and were the characteristics of the dropouts compared with the completers of the study?	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Was a long-term follow-up measurement in both groups comparable?	Yes	No	Yes	No	No	No	No	No	Yes
Was the timing of the outcome measurements in both groups comparable?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the sample size for each group described by means of a power calculation?	Yes	Yes[a]	Yes	Yes	Yes	No[a]	Yes	No	Yes
Did the analysis include an intention-to-treat analysis?	Yes	No	No[a]	Yes	Yes	No	Yes	No	Yes
Were point estimates and measures of variability presented for the primary outcome measures?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intervention exposure									
Was the exposure frequency measured?	Yes[a]	Yes[a]	Yes	Yes	No	Yes	Yes	No	Yes
Was the exposure duration measured?	Na	No	Na	Yes	Na	Yes[a]	Yes	Na	No

[a] Information obtained from original author
Na, Not applicable

Discussion

Effects of ICT-based PA interventions for children and adolescents

The present review systematically evaluated the efficacy and methodological quality of ICT-based interventions that applied Internet and/or SMS as a delivery mode for PA behavior change in children and adolescents based on the evidence of randomized controlled trials during the past 12 years (1997-2009). As mentioned earlier, the review by Norman et al [33] only included 3 RCTs focused on children and adolescents. Although the present review located 6 more studies, the small number of included studies in both reviews indicates the needs for additional studies. The 3 RCTs reviewed by Norman et al were Internet-based interventions. In contrast, the present review illustrated that the proportion of Internet- and SMS-based interventions was almost equal. This finding suggests the emerging role of SMS in changing PA behaviors in a young population. However, there are no existing criteria to inform the choice of ICT for different research purposes. Future studies investigating this issue are suggested.

The present review demonstrates consistent evidence supporting the efficacy of improving psychosocial variables through ICT-based Interventions (eg, self-efficacy). For behavioral variables (ie, PA level), evidence was less consistent. Unfortunately, there is insufficient information explaining the underlying mechanisms for change because many of the included studies have an incomplete theoretical foundation.

Baranowski and Jago [54] indicated that a complete theoretical foundation of an intervention played an imperative role in explaining the effects of PA interventions. Their framework stated that a complete theoretical foundation not only includes employing theory and theory-based strategies to design the intervention, but also an evaluation of mediating variables. Further, changes as a result of an intervention should be associated with changes in outcome variables and potential confounders (eg, gender and ethnicity) should also be assessed for their role in influencing the relationship between the mediating variables and the target behavior.

In the present review, only half of the studies reviewed developed their interventions based on health behavior change theories. We found that 4 theory-based interventions measured a behavioral variable (PA level) as the outcome and that only 2 studies [43,46] explicitly described how the desired outcomes were manipulated by the intervention components. None have measured changes in the theoretical constructs as the outcome variables. Although a few studies [42,43,47] analyzed the confounders (ie, season, baseline PA level, and gender) that serve as moderators for intervention effects, it was still difficult to determine the underlying mechanism that drives an intervention's success and failure [55,56]. These findings reinforce the need to strengthen theoretical foundations in future studies.

In all, 7 studies [41-45,47,48] compared the effects of the ICT groups with either non-ICT or no treatment control groups. Also, 6 studies showed that ICT groups were either as effective as (n = 3) [42,45,48] or superior to (n = 3) [41,43,47] non-ICT groups. However, it is inconclusive whether ICT is equivalent or superior to other delivery approaches (ie, face-to-face). The problem with existing research is that the majority of the studies employed both ICT and face-to-face modes. In addition, these studies did not include a comparable control group. When analyzing the intervention characteristics of the ICT group and the non-ICT control group, the contact frequency of the two groups varied. For instance, Fanklin et al [41] and Lubans et al [43] employed ICT to provide tailored feedback to participants in the ICT group, but it was not offered to the non-ICT control group. Shapiro et al [48] gave tailored feedback to participants in both the ICT group and non-ICT control group, but the contact frequency in the ICT group (once per day) was far more frequent than in the non-ICT control group (once per week). Despite the fact that the impact of varied contact frequency on intervention efficacy was unclear, existing evidence [36] supported the notion that higher contact frequency was associated with enhanced efficacy. Although these studies reported significant between-group differences favoring the ICT group, it is difficult to determine whether the surplus effects in the ICT group were a result of the use of ICT or increased contact frequency. Nonetheless, the findings provide evidence supporting the effectiveness of ICT in PA interventions for children and adolescents, especially when used along with other delivery approaches.

In all, 3 studies [41,43,47] demonstrated significant between-group differences and large effect sizes, which are obviously larger than in the remaining studies. Attempts were made to examine whether any specific intervention characteristics contributed to larger effect sizes. We found that ICT-based interventions that were grounded in behavior change theory and utilized an investigator-initiation strategy were more likely to show more significant between-group differences and larger effect sizes than those did not (See Table 1).

There is extensive prior evidence suggesting that use of theory has a beneficial effect on interventions [26,55-57]. The mechanism that explained this evidence was that behavior change theories could inform researchers of the most influential mediating variables of the target behavior [58]. Through intervening on these influential mediating variables, people would be more likely to initiate behavior change. In this review, efforts were made to examine whether improved efficacy was associated with the use of a specific behavior change theory. Due to the small number of theory-based interventions and the heterogeneity of study designs, direct comparisons examining the effects across different behavior change theories on intervention efficacy could not be performed. Again, this finding indicates the importance of using a theoretical framework to facilitate an intervention's success and designing experimental studies that compare the effects of different behavior change theories on intervention efficacy.

It is important to note that ICT-based interventions that used behavior change theory along with the adoption of investigator-initiated strategies showed significant between-group difference and larger effect size when compared with a non-ICT control group. A possible reason for the improved efficacy is that the investigator-initiation strategy uses a "pushed" approach [59], where automatic and specific materials (ie, Web hyperlinks and personalized feedback) are directly addressed to participant's personal email or SMS. This can save participants the cognitive effort it takes to plan when, where, what, and how to prevent time conflicts with other daily tasks before using ICT to access the materials. All the participants have to do is to check their email and use their mobile phone in a typical fashion. This makes ICT-based interventions more compatible with a participant's existing practice and lifestyle. According to the diffusion of innovation theory [60], increased compatibility of an intervention could enhance the likelihood that children and adolescents would read the materials and adhere to the interventions. Consequently, the initiation of behavior change could be more likely to happen [61]. However, the present review could not confirm the effect of ICT initiation strategy on intervention exposure and adherence rate since these data were not available in most of the included studies. Clearly, more studies are needed to investigate the impact of ICT initiation strategy on intervention exposure rate, adherence rate, and efficacy.

With respect to methodological quality, two-thirds of the included studies were classified as having good methodological quality. Most of the studies failed to report the allocation concealment and blinding of outcome assessments. Appropriate allocation concealment is important to avoid selection and confounding bias [62,63] while blinding of outcome assessment can lower the risk of exaggerating treatment effectiveness [64]. Absence of the above information will prevent us from appraising the risk of bias. Without understanding the risk of bias, we should remain cautious about the positive effects of

ICT-based interventions [62,65]. Another shortcoming to this body of research is that very few studies (n = 2) conducted long-term follow up. As the development of a PA habit is a life-long task and maintenance of a new adopted behavior may require at least 6 months [66], interventions with a long-term follow-up period (> 6 months) may better allow us to assess the effects of ICT-based PA interventions.

There are two methodological issues related to the intervention exposure rate. First, there was improper reporting of the intervention exposure rate. In the present review, only 2 studies reported both exposure frequency and exposure duration. Second, there was an unclear description of the measurements used to assess intervention exposure rate. Williamson et al [49] was the only research group that employed objective measurements, and theirs was also the only study that defined intervention exposure rate and described how to measure it. Without an objective measure, the risk of response bias may increase [67]. These methodological issues have prevented us from estimating the extent to which the prescribed intervention dosage was received by the participants. If the dosage received by participants could not be estimated, it is hard to determine whether improvements in the measured outcomes were an effect of the interventions or other factors. Using Lubans et al [43] as an example, participant's exposure to the social support email was low, but they still observed significant and positive results. It is possible that the positive effects were influenced by the face-to-face sports program. In addition, Crutzen and colleagues [68] suggested that intervention exposure rate may reflect the salience of intention of behavior change, which varied during different time periods throughout the intervention. It is important for researchers to understand participants' behavior change patterns so necessary adjustments can be made. These findings indicate the need for adopting objective and valid instruments to measure intervention exposure rates. It is also important for future studies to report these data when presenting results.

Conclusion

The present review provides evidence supporting the positive effects of ICTs in PA interventions for children and adolescents, especially when used with other delivery approaches (ie, face-to-face). Because ICT delivery approaches are often mixed with other approaches and these studies sometimes lack a comparable control group, additional research is needed to establish the true independent effects of ICT as an intervention delivery mode. Nevertheless, this review has found that combining the use of behavior change theory and investigator-initiated strategies could be associated with enhanced intervention efficacy and larger effect sizes. However, more studies are needed prior to reaching solid conclusions. Although two-thirds of the studies demonstrated satisfactory methodological quality, several quality criteria still have room for improvement (eg, providing clear descriptions of allocation concealment, blinding of outcome assessment, and intervention exposure rate). Furthermore, researchers should also consider intervening for a longer duration and employing objective instruments for assessing intervention exposure rate. Due to the small number of studies that met the inclusion criteria, researchers should be cautious when interpreting the findings of the present review.

Acknowledgments

The authors wish to thanks Ms Dianne Cmor and Mr. Christopher Chan for their assistances throughout the literature search process.

Conflicts of Interest

None declared

References

1. Hallal PC, Victora CG, Azevedo MR, Wells JC. Adolescent physical activity and health: a systematic review. Sports Med 2006;36(12):1019-1030. [Medline: 17123326]
2. Boreham C, Riddoch C. The physical activity, fitness and health of children. J Sports Sci 2001 Dec;19(12):915-929. [Medline: 11820686]
3. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
4. Janssen I, Katzmarzyk PT, Boyce WF, Vereecken C, Mulvihill C, Roberts C, Health Behaviour in School-Aged Children Obesity Working Group. Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns. Obes Rev 2005 May;6(2):123-132. [doi: 10.1111/j.1467-789X.2005.00176.x] [Medline: 15836463]
5. Brodersen NH, Steptoe A, Boniface DR, Wardle J. Trends in physical activity and sedentary behaviour in adolescence: ethnic and socioeconomic differences. Br J Sports Med 2007 Mar;41(3):140-144. [doi: 10.1136/bjsm.2006.031138] [Medline: 17178773]
6. Koezuka N, Koo M, Allison KR, Adlaf EM, Dwyer JJ, Faulkner G, et al. The relationship between sedentary activities and physical inactivity among adolescents: results from the Canadian Community Health Survey. J Adolesc Health 2006 Oct;39(4):515-522. [doi: 10.1016/j.jadohealth.2006.02.005] [Medline: 16982386]
7. Hamar P, Biddle S, Soós I, Takács B, Huszár A. The prevalence of sedentary behaviours and physical activity in Hungarian youth. Eur J Public Health 2010 Feb;20(1):85-90 [FREE Full text] [doi: 10.1093/eurpub/ckp100] [Medline: 19587226]

8. Jago R, Baranowski T. Non-curricular approaches for increasing physical activity in youth: a review. Prev Med 2004 Jul;39(1):157-163. [doi: 10.1016/j.ypmed.2004.01.014] [Medline: 15207997]
9. 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]
10. Marcus BH, Williams DM, Dubbert PM, Sallis JF, King AC, Yancey AK, American Heart Association Council on Nutrition‚ Physical Activity‚ and Metabolism (Subcommittee on Physical Activity), American Heart Association Council on Cardiovascular Disease in the Young, Interdisciplinary Working Group on Quality of Care and Outcomes Research. Physical activity intervention studies: what we know and what we need to know: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity); Council on Cardiovascular Disease in the Young; and the Interdisciplinary Working Group on Quality of Care and Outcomes Research. Circulation 2006 Dec 12;114(24):2739-2752 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.106.179683] [Medline: 17145995]
11. Cassell MM, Jackson C, Cheuvront B. Health communication on the Internet: an effective channel for health behavior change? J Health Commun 1998;3(1):71-79. [Medline: 10947375]
12. van Sluijs EM, McMinn AM, Griffin SJ. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. BMJ 2007 Oct 6;335(7622):703 [FREE Full text] [doi: 10.1136/bmj.39320.843947.BE] [Medline: 17884863]
13. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007;29:144-159. [doi: 10.1093/epirev/mxm010] [Medline: 17556765]
14. Dobbins M, De Corby K, Robeson P, Husson H, Tirilis D. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6-18. Cochrane Database Syst Rev 2009(1):CD007651. [doi: 10.1002/14651858.CD007651] [Medline: 19160341]
15. Fotheringham MJ, Owen N. Introduction. Interactive health communication in preventive medicine. Am J Prev Med 2000 Aug;19(2):111-112. [Medline: 10913901]
16. Patrick K. Information technology and the future of preventive medicine: potential, pitfalls, and policy. Am J Prev Med 2000 Aug;19(2):132-135. [Medline: 10913905]
17. Wylie-Rosett J, Swencionis C, Ginsberg M, Cimino C, Wassertheil-Smoller S, Caban A, et al. Computerized weight loss intervention optimizes staff time: the clinical and cost results of a controlled clinical trial conducted in a managed care setting. J Am Diet Assoc 2001 Oct;101(10):1155-62; quiz 1163. [Medline: 11678486]
18. Sevick MA, Dunn AL, Morrow MS, Marcus BH, Chen GJ, Blair SN. Cost-effectiveness of lifestyle and structured exercise interventions in sedentary adults: results of project ACTIVE. Am J Prev Med 2000 Jul;19(1):1-8. [Medline: 10865157]
19. Abrams DB, Orleans CT, Niaura RS, Goldstein MG, Prochaska JO, Velicer W. 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. [doi: 10.1007/BF02895291] [Medline: 18425675]
20. Huhman M, Potter LD, Wong FL, Banspach SW, Duke JC, Heitzler CD. Effects of a mass media campaign to increase physical activity among children: year-1 results of the VERB campaign. Pediatrics 2005 Aug;116(2):e277-e284 [FREE Full text] [doi: 10.1542/peds.2005-0043] [Medline: 16061581]
21. 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]
22. Nigg CR. Technology's influence on physical activity and exercise science: the present and the future. Psychol Sport Exerc 2003;4:57-65.
23. 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]
24. Bastian H. Internet, Health Information on the. In: Heggenhougan K, Quah S, editors. International Encyclopedia of Public Health. London, England: Elsevier Inc; 2008:679-683.
25. Borzekowski DL, Rickert VI. Adolescent cybersurfing for health information: a new resource that crosses barriers. Arch Pediatr Adolesc Med 2001 Jul;155(7):813-817 [FREE Full text] [Medline: 11434849]
26. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
27. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [Medline: 14672383]
28. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
29. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
30. Atkinson NL, Gold RS. The promise and challenge of eHealth interventions. Am J Health Behav 2002;26(6):494-503. [Medline: 12437024]

31. Brug J, Oenema A, Campbell M. Past, present, and future of computer-tailored nutrition education. Am J Clin Nutr 2003 Apr;77(4 Suppl):1028S-1034S [FREE Full text] [Medline: 12663313]
32. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
33. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345. [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
34. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16. [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
35. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
36. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
37. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
38. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.
39. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Internet-delivered interventions aimed at adolescents: a Delphi study on dissemination and exposure. Health Educ Res 2008 Jun;23(3):427-439 [FREE Full text] [doi: 10.1093/her/cym094] [Medline: 18209115]
40. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. A conceptual framework for understanding and improving adolescents' exposure to Internet-delivered interventions. Health Promot Int 2009 Sep;24(3):277-284 [FREE Full text] [doi: 10.1093/heapro/dap018] [Medline: 19515716]
41. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
42. Jago R, Baranowski T, Baranowski JC, Thompson D, Cullen KW, Watson K, et al. Fit for Life Boy Scout badge: outcome evaluation of a troop and Internet intervention. Prev Med 2006 Mar;42(3):181-187. [doi: 10.1016/j.ypmed.2005.12.010] [Medline: 16458955]
43. Lubans DR, Morgan PJ, Callister R, Collins CE. Effects of integrating pedometers, parental materials, and E-mail support within an extracurricular school sport intervention. J Adolesc Health 2009 Feb;44(2):176-183. [doi: 10.1016/j.jadohealth.2008.06.020] [Medline: 19167667]
44. Marks JT, Campbell MK, Ward DS, Ribisl KM, Wildemuth BM, Symons MJ. A comparison of Web and print media for physical activity promotion among adolescent girls. J Adolesc Health 2006 Jul;39(1):96-104. [doi: 10.1016/j.jadohealth.2005.11.002] [Medline: 16781967]
45. Newton KH, Wiltshire EJ, Elley CR. Pedometers and text messaging to increase physical activity: randomized controlled trial of adolescents with type 1 diabetes. Diabetes Care 2009 May;32(5):813-815 [FREE Full text] [doi: 10.2337/dc08-1974] [Medline: 19228863]
46. Patrick K, Sallis JF, Prochaska JJ, Lydston DD, Calfas KJ, Zabinski MF, et al. A multicomponent program for nutrition and physical activity change in primary care: PACE+ for adolescents. Arch Pediatr Adolesc Med 2001 Aug;155(8):940-946 [FREE Full text] [Medline: 11483123]
47. Prochaska JJ, Sallis JF. A randomized controlled trial of single versus multiple health behavior change: promoting physical activity and nutrition among adolescents. Health Psychol 2004 May;23(3):314-318. [doi: 10.1037/0278-6133.23.3.314] [Medline: 15099173]
48. Shapiro JR, Bauer S, Hamer RM, Kordy H, Ward D, Bulik CM. Use of text messaging for monitoring sugar-sweetened beverages, physical activity, and screen time in children: a pilot study. J Nutr Educ Behav 2008;40(6):385-391. [doi: 10.1016/j.jneb.2007.09.014] [Medline: 18984496]
49. Williamson DA, Walden HM, White MA, York-Crowe E, Newton RL, Alfonso A, et al. Two-year internet-based randomized controlled trial for weight loss in African-American girls. Obesity (Silver Spring) 2006 Jul;14(7):1231-1243 [FREE Full text] [doi: 10.1038/oby.2006.140] [Medline: 16899804]
50. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
51. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
52. Prochaska JO, Marcus BH. The transtheoretical model: applications to exercise. In: Dishman RK, editor. Advances in Exercise Adherence. Champaign, IL: Human Kinetics; 1994:161-180.
53. Marlatt GA, George WH. Relapse prevention: introduction and overview of the model. Br J Addict 1984 Sep;79(3):261-273. [Medline: 6595020]

54. Baranowski T, Jago R. Understanding the mechanisms of change in children's physical activity programs. Exerc Sport Sci Rev 2005 Oct;33(4):163-168. [Medline: 16239832]
55. Rothman AJ. "Is there nothing more practical than a good theory?": Why innovations and advances in health behavior change will arise if interventions are used to test and refine theory. Int J Behav Nutr Phys Act 2004 Jul 27;1(1):11 [FREE Full text] [doi: 10.1186/1479-5868-1-11] [Medline: 15279674]
56. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010 Jan;29(1):1-8. [doi: 10.1037/a0016939] [Medline: 20063930]
57. Painter JE, Borba CP, Hynes M, Mays D, Glanz K. The use of theory in health behavior research from 2000 to 2005: a systematic review. Ann Behav Med 2008 Jun;35(3):358-362. [doi: 10.1007/s12160-008-9042-y] [Medline: 18633685]
58. Baranowski T, Cullen KW, Nicklas T, Thompson D, Baranowski J. Are current health behavioral change models helpful in guiding prevention of weight gain efforts? Obes Res 2003 Oct;11 Suppl:23S-43S. [doi: 10.1038/oby.2003.222] [Medline: 14569036]
59. Waller A, Franklin V, Pagliari C, Greene S. Participatory design of a text message scheduling system to support young people with diabetes. Health Informatics J 2006 Dec;12(4):304-318. [doi: 10.1177/1460458206070023] [Medline: 17093001]
60. Roger ME, Roger E. Diffusion of Innovations. 5th edition. New York, NY: Free Press; 2003.
61. Petty RE, Cacioppo JT. Communication and Persuasion: Central and Peripheral Routes to Attitude Change. New York, NY: Springer/Verlag; 1986.
62. Schulz KF. Assessing allocation concealment and blinding in randomized controlled trials: why bother? ACP J Club 2000;132(2):A11-A12. [Medline: 10750446]
63. Greenhalgh T. Assessing the methodological quality of published papers. BMJ 1997 Aug 2;315(7103):305-308 [FREE Full text] [Medline: 9274555]
64. Poolman RW, Struijs PA, Krips R, Sierevelt IN, Marti RK, Farrokhyar F, et al. Reporting of outcomes in orthopaedic randomized trials: does blinding of outcome assessors matter? J Bone Joint Surg Am 2007 Mar;89(3):550-558. [doi: 10.2106/JBJS.F.00683] [Medline: 17332104]
65. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. Lancet 2002 Feb 23;359(9307):696-700. [doi: 10.1016/S0140-6736(02)07816-9] [Medline: 11879884]
66. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
67. Biemer PP, Witt M. Repeated measures estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse. NIDA Res Monogr 1997;167:439-476. [Medline: 9243573]
68. Crutzen R, de Nooijer J, Candel MJ, de Vries NK. Adolescents who intend to change multiple health behaviours choose greater exposure to an internet-delivered intervention. J Health Psychol 2008 Oct;13(7):906-911. [doi: 10.1177/1359105308095064] [Medline: 18809641]

Abbreviations

BCT: behavior change technique
ICT: information and communication technology
PA: physical activity
RCT: randomized controlled trial
RPM: relapse prevention model
SCT: social cognitive theory
SMS: short message services
TTM: transtheoretical model

Edited by G Eysenbach; submitted 19.03.10; peer-reviewed by T Baranowski; comments to author 25.06.10; revised version received 17.02.11; accepted 24.03.11; published 13.07.11.

Please cite as:
Lau PWC, Lau EY, Wong DP, Ransdell L
A Systematic Review of Information and Communication Technology–Based Interventions for Promoting Physical Activity Behavior Change in Children and Adolescents
J Med Internet Res 2011;13(3):e48
URL: http://www.jmir.org/2011/3/e48/
doi:10.2196/jmir.1533
PMID:21749967

Original Paper

Persuasive Features in Web-Based Alcohol and Smoking Interventions: A Systematic Review of the Literature

Tuomas Lehto[1*], MSc; Harri Oinas-Kukkonen[1*], PhD

Oulu Advanced Research on Software and Information Systems, Department of Information Processing Science, University of Oulu, Oulu, Finland
*all authors contributed equally

Corresponding Author:
Tuomas Lehto, MSc
Oulu Advanced Research on Software and Information Systems
Department of Information Processing Science
University of Oulu
Rakentajantie 3
Oulu, 90570
Finland
Phone: 358 8 553 1900
Fax: 358 8 553 1890
Email: tuomas.lehto@oulu.fi

Abstract

Background: In the past decade, the use of technologies to persuade, motivate, and activate individuals' health behavior change has been a quickly expanding field of research. The use of the Web for delivering interventions has been especially relevant. Current research tends to reveal little about the persuasive features and mechanisms embedded in Web-based interventions targeting health behavior change.

Objectives: The purpose of this systematic review was to extract and analyze persuasive system features in Web-based interventions for substance use by applying the persuasive systems design (PSD) model. In more detail, the main objective was to provide an overview of the persuasive features within current Web-based interventions for substance use.

Methods: We conducted electronic literature searches in various databases to identify randomized controlled trials of Web-based interventions for substance use published January 1, 2004, through December 31, 2009, in English. We extracted and analyzed persuasive system features of the included Web-based interventions using interpretive categorization.

Results: The primary task support components were utilized and reported relatively widely in the reviewed studies. Reduction, self-monitoring, simulation, and personalization seem to be the most used features to support accomplishing user's primary task. This is an encouraging finding since reduction and self-monitoring can be considered key elements for supporting users to carry out their primary tasks. The utilization of tailoring was at a surprisingly low level. The lack of tailoring may imply that the interventions are targeted for too broad an audience. Leveraging reminders was the most common way to enhance the user-system dialogue. Credibility issues are crucial in website engagement as users will bind with sites they perceive credible and navigate away from those they do not find credible. Based on the textual descriptions of the interventions, we cautiously suggest that most of them were credible. The prevalence of social support in the reviewed interventions was encouraging.

Conclusions: Understanding the persuasive elements of systems supporting behavior change is important. This may help users to engage and keep motivated in their endeavors. Further research is needed to increase our understanding of how and under what conditions specific persuasive features (either in isolation or collectively) lead to positive health outcomes in Web-based health behavior change interventions across diverse health contexts and populations.

(J Med Internet Res 2011;13(3):e46) doi:10.2196/jmir.1559

KEYWORDS

Web-based; online; Internet; alcohol; smoking; intervention; behavior change; persuasive; PSD model; review

Introduction

In the past decade, the use of technologies to persuade, motivate, and activate individuals' health behavior change has been a quickly expanding field of research [1-13]. The use of the Web (and related technologies) for delivering interventions has been especially relevant. Automated health behavior interventions have the potential of high reach and low cost [14]. A recent meta-analysis of 75 randomized controlled trials (RCTs) provided support for their effectiveness in changing knowledge, attitudes, and behavior in the health promotion area [15].

There is no consensus on the terminology used to designate the activities conducted over the Internet for mental and physical health purposes. According to Barak et al [16], a Web-based intervention is "a primarily self-guided intervention program that is executed by means of a prescriptive online program operated through a website and used by consumers seeking health- and mental health-related assistance." The intervention in itself is aimed at creating a positive change and/or improving knowledge, awareness, or understanding by providing sound health-related material to the user through an interactive Web-based information system.

Considerable variety exists in terms of types of program content, interactivity, functionality, and level of multimedia of Web-based interventions. In addition, intensity is a major variable; some Web-based interventions are long-term, automated, interactive, tailored, multicomponent programs whereas others are brief online screening instruments with tailored feedback. Finally, participant attrition and exposure rates may vary widely (eg, [18,19]). Technological applications integrating health information with online peer support, decision support, and/or help with behavior change provides an alternative for helping people to achieve better health [20]. However, relatively little is known of designing effective Web-based interventions to support sustained behavior change and improved well-being (eg, [4,21-25]).

The purpose of the present review was to extract and analyze persuasive system features of the included Web-based alcohol and smoking interventions use by applying the PSD model [26]. Suggestions for further research and development are provided for this expanding field of research.

Persuasive Technology and Health Behavior Change

Research on persuasive technology has been introduced relatively recently [27,28]. Briñol and Petty [29] outline persuasion as follows: "In the typical situation where persuasion is possible, a person or a group of people (ie, *the recipient*) receives an intervention (eg, *a persuasive message*) from another individual or group (ie, *the source*) in a particular setting (ie, *the context*)."

Persuasive systems may be defined as computerized software or information systems designed to reinforce, change, or shape attitudes or behaviors or both without using coercion or deception [26]. Successful persuasion takes place when the target of change (eg, attitudes or beliefs) is modified in the desired direction [29].

Changing people's behavior is at the heart of health promotion. An individual's behavior has a significant impact on, for example, cancer and heart disease, which are common causes of premature mortality. The Internet is transforming health care [30] and can be seen as a prime candidate for the application of key behavioral science theories and principles to promote healthier behaviors. There are several advantages in Internet-delivered interventions, for example, reducing cost and increasing convenience for users, reduction of health service costs, reduction of stigma and isolation of users, the need for timely information, and increased user and supplier control of the intervention [31-33]. Internet-based resources, in particular the Web, have many of the characteristics necessary for persuasive communication, and they may provide a channel which integrates the positive attributes of interpersonal and mass communication [34,35]. Web-based systems can give immediate feedback and match the information with the respondent's level of awareness, beliefs, and motivations at that particular time [10]. Additionally, Web-based interventions may overcome isolation of time, mobility, and geography. It has to be noted that Web-based interventions still may be no substitute for face-to-face contact [31,36].

Persuasive Technology: Designing Systems That Aim at Behavior Change

Examples of persuasive technology can be found quite easily, as there are a variety of websites promoting healthier lifestyles. One of the strongest domains of innovation for persuasive technology in the near future will be preventive health care [37]. On a par with health behavior change, persuasive technology has the potential for significant breakthroughs in many areas of human well-being, such as education and environmental conservation. Nevertheless, the use of persuasive technology in the health arena is still in its infancy. While the field is expanding, it is evident that more research is needed to better determine how the persuasiveness of the systems affects users' intended behavior.

According to Fogg [38], attempts to create persuasive systems often fail because many projects are too ambitious, being set up for failure. For example, a design team might select a challenging behavior as the target, for example, smoking cessation, but without having ever before created such a persuasive system, the success rate might be low. Zhang [39] stated: "A large number of health information system projects fail. Most of these failures are not due to flawed technology but rather due to the lack of systematic considerations of human and other nontechnology issues in the design and implementation processes." Thus, designing systems that aim at behavior change requires thorough understanding of the problem domain and the underpinning theories and strategies of persuasive systems design. Usually, an interdisciplinary team of professionals is also needed. The main decision points in developing interventions include defining the primary goal of the intervention, defining the target population, and selecting the messages for the intervention [40].

In the present study, the persuasive systems design model (PSD) [26] was applied as the framework for identifying various persuasive techniques that have been incorporated into the

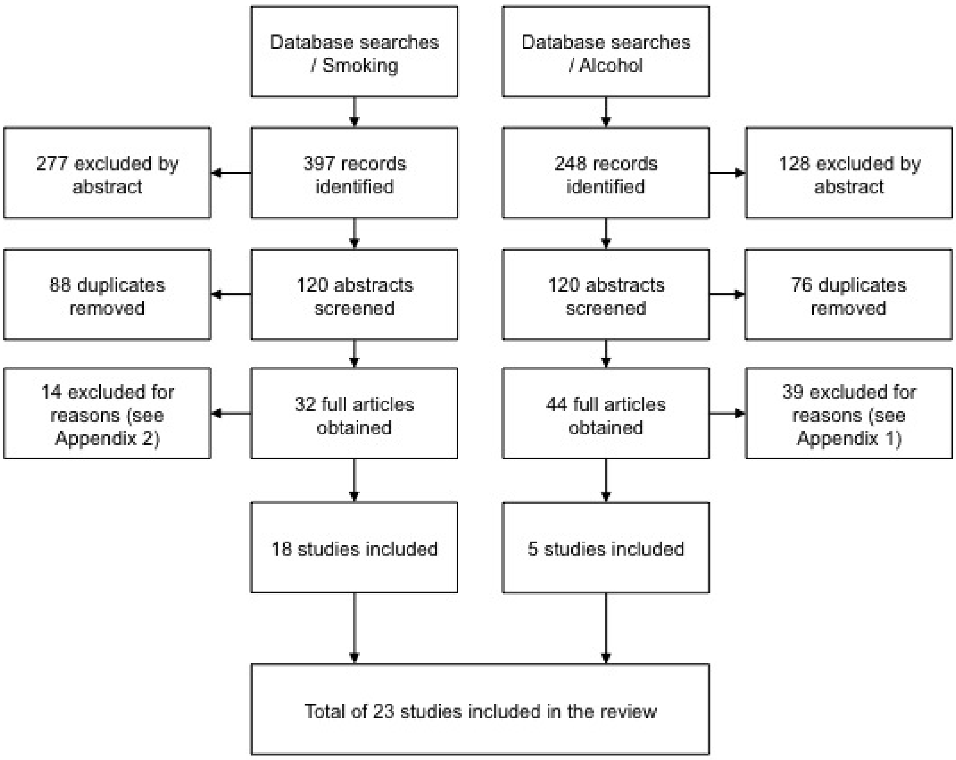
Web-based substance use interventions. The PSD model provides a recent and extensive conceptualization of technology-mediated persuasion.

Methods

Identification of Studies

We conducted electronic literature searches in five databases (Academic Search Premier [EBSCO Publishing, Herts, England], Cochrane Central Register of Controlled Trials, ISI Web of Knowledge, Medline [Ovid], and Scopus) to identify randomized controlled trials of Web-based interventions for substance use published January 1, 2004, through December 31, 2009, in English.

Figure 1. Study identification process



The following terms were used in the database searches: (1) online*, Internet*, web* and (2) intervention, self-help, treatment, trial. To identify alcohol interventions we used the following additional search terms: drink*, alcohol*. For smoking interventions we utilized the terms: smoke*, smoking, cigarette, tobacco, cessation. An asterisk (*) denotes a wildcard. We also screened the bibliographies of relevant articles, including systematic reviews [4,8,10,13,17,41,42] and meta-analyses [5-7,9] to identify potentially relevant studies. The study identification process is depicted in Figure 1.

Inclusion and Exclusion Criteria

Articles were included if they: (1) focused on Web-based alcohol or smoking (tobacco) interventions, (2) assessed behavioral outcomes or program utilization, (3) were randomized controlled trials or quasi-experimental designs, and (4) were peer-reviewed full research articles.

Follow-up studies that used data from the same cohort of participants were excluded. Both Danaher et al [43] and Severson et al [44] drew data from the same source, but they were included. Brief interventions (see [45, 46]) were excluded because of their rather limited content. Moreover, brief interventions have been commonly reported [47], even to the extent that it was deemed appropriate to put more emphasis on analyzing the less frequently reported, more complex Web-based substance use interventions. According to Moyer [45], a challenge in outlining the research literature on brief interventions is the varying definitions used in different studies. Following Babor's [46] definition, we included the studies of interventions providing more than three sessions and/or aiming at more than 60 minutes of individual engagement with the program. Furthermore, we excluded articles explicitly using the term *brief intervention* in the title, abstract, or keywords. (See

Multimedia Appendices 1 and 2 for excluded articles and reasons these were excluded.)

Data Abstraction

In total, 23 studies were included in the review and coded. The methodological quality of the included studies was evaluated by applying the CONSORT (Consolidated Standards of Reporting Trials) 2010 checklist [48]. There were some concerns with the quality of 4 studies [44,49-51], but they were still included. Overall, the information presented in the selected articles was thoroughly examined and evaluated. The first author coded all the included articles using a predefined form (devised by the authors) for evaluating persuasive systems. In addition, the abstracted data included various study characteristics (see Table 2). The resulting entries were checked and commented on by the second author. Any disparities were resolved through discussion.

Persuasive Systems Design (PSD) Model

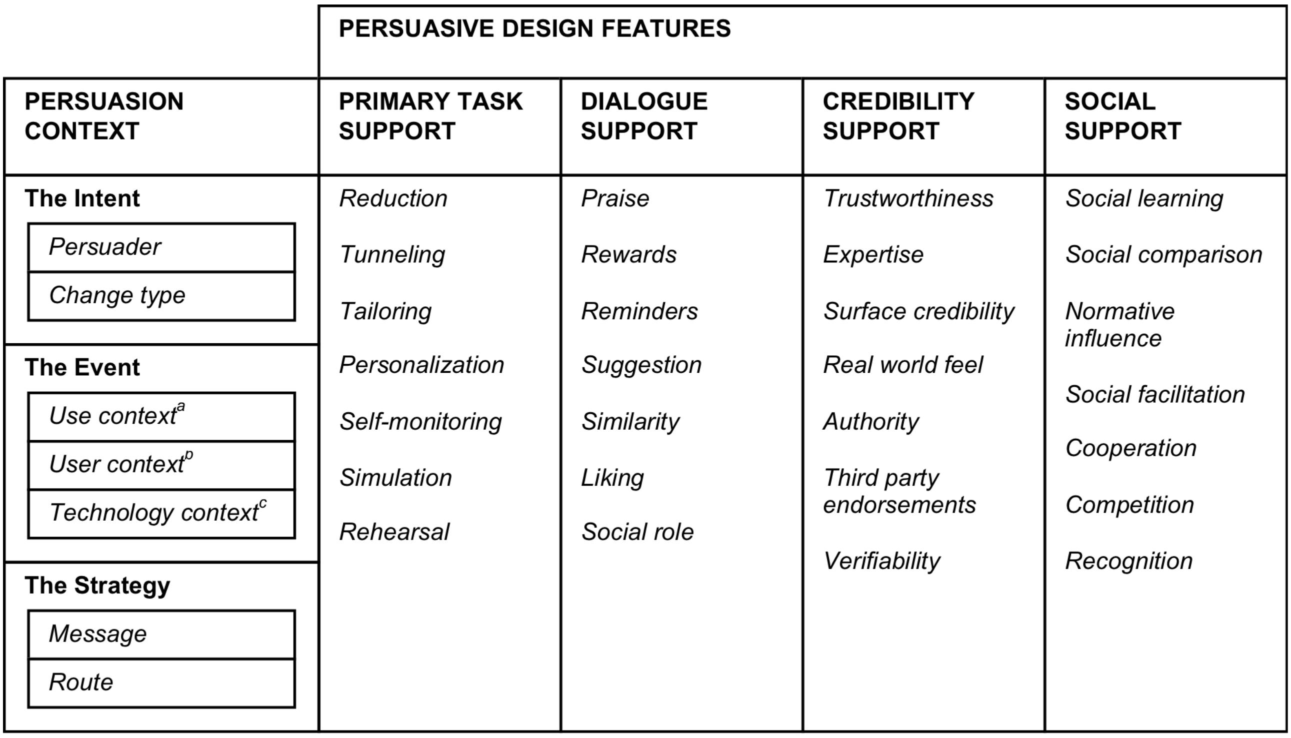
Information technology always influences people's attitudes and behavior in one way or another [52,53]. Oinas-Kukkonen and Harjumaa [26,52] have conceptualized a framework for designing and evaluating persuasive systems, known as the persuasive systems design (PSD) model. The PSD model builds on multiple theoretical constructs, such as goal-setting theory [54], elaboration likelihood model [55], and theory of reasoned action/planned behavior [56]. The PSD model is described in full detail elsewhere [26].

Although the PSD model is yet relatively unknown, we consider its use to be justified for this context. In our view, it is the most sophisticated persuasive design and evaluation method available. Many of the principles in the PSD model have been adopted and modified from the seminal work of Fogg [27]. We acknowledge the existence of similar endeavors, such as Ritterband's behavior model for Internet interventions [22] and Abraham and Michie's [24] taxonomy of behavior change techniques used in interventions. Despite the similarities and potential overlap, these approaches are quite different.

The PSD model presents a way to analyze, design, and evaluate the persuasion context and related techniques. Persuasion context analysis includes recognizing the intent, the event, and the strategy for persuasion. Dey [57] defines context as follows: "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

In the PSD model [26], the categories for persuasive system principles are primary task support (supporting the user's primary task), dialogue support (supporting the interaction between the user and the system), system credibility (the more credible the system is, the more persuasive it is), and social support (the system motivates users by leveraging social influence).

Figure 2. PSD Model (adapted from Oinas-Kukkonen and Harjumaa [26])



	PERSUASIVE DESIGN FEATURES			
PERSUASION CONTEXT	PRIMARY TASK SUPPORT	DIALOGUE SUPPORT	CREDIBILITY SUPPORT	SOCIAL SUPPORT
The Intent *Persuader* *Change type*	*Reduction*	*Praise*	*Trustworthiness*	*Social learning*
	Tunneling	*Rewards*	*Expertise*	*Social comparison*
	Tailoring	*Reminders*	*Surface credibility*	*Normative influence*
The Event *Use context*[a] *User context*[b] *Technology context*[c]	*Personalization*	*Suggestion*	*Real world feel*	*Social facilitation*
	Self-monitoring	*Similarity*	*Authority*	*Cooperation*
	Simulation	*Liking*	*Third party endorsements*	*Competition*
The Strategy *Message* *Route*	*Rehearsal*	*Social role*	*Verifiability*	*Recognition*

[a] Problem domain dependent features
[b] User dependent features e.g. goals, motivation, lifestyles, and others
[c] Technology dependent features

The intent includes the persuader and the target behavior change type that the system is to cause in the user. The persuader is the initiator for the development of the system. The event contains the use, user, and technology contexts.

The use context refers to the problem domain dependent features. The use context, in particular problem domain dependent features, is relevant to the persuasion event. The use context can be determined by answering the following questions [58]: Who are the users as a group? What problem-domain dependent features are to be addressed by the design? Who (or what) else is competing for attention in this space?

The user context refers to the individual users characteristics. The user context includes, but is not limited to, users' (patient, research participant, consumer [22]) characteristics, goals, abilities, and cultural factors. The user context can be clarified by addressing the following questions [58]: What is specific for the users with regard to what they are to be persuaded of? Why is persuasion needed? What constrains their decision?

There is an obvious need for depicting the technology context when describing a Web-based intervention. The technology context refers to the features and requirements of the technological platform and/or application.

The strategy in the PSD model emphasizes two elements, namely the message and the route. The message refers to the form and/or content selected to deliver the intended transformation (eg, behavior or attitude change). The content could be, for instance, statistical data about the health risks of drinking, but the information could be presented to the user in plain text, streaming video, or it could be embedded in a game. The route for persuasion can be direct, indirect, or both. A direct approach provides one or a few solid and convincing arguments, whereas an indirect route is based on a number of facts rather than a single strong argument (compare *central* and *peripheral* routes in ELM). Both routes may be in use simultaneously. A system might represent rational arguments while employing design patterns, which in themselves have been proven persuasive. As an example, an avatar with a specific voice type (angry vs soft spoken or male vs female) to present the message might make the delivery of the content more persuasive for a user; thus, users would be persuaded through the design choices made by the designer.

The persuasive design dimensions and principles are discussed and exemplified below.

Results

Study Characteristics

The characteristics of the included studies are presented in Table 1. Of the 23 articles, 16 targeted smoking, 5, problem drinking, and 2, smokeless tobacco use.

Table 1. Characteristics of the included studies

Study Author (Year)	Problem Domain	Primary Objective of the Study	User Context (Number of Participants)	Use and Technology Context (Intervention)	Summary of Findings
An et al (2008) [59]	Smoking	To determine whether an online intervention with college smokers could increase self-reported 30-day abstinence rates at the end of a 2-semester intervention	College smokers at the University of Minnesota (517)	Online college life magazine that provided personalized smoking cessation messages and peer email support (RealU)	The rate of 30-day abstinence at week 30 was higher for the intervention compared with the control group. (41% vs 23%, *P* < .001)
Bersamin et al (2007) [60]	Problem drinking	To assess whether a new online alcohol misuse prevention course is more effective at reducing alcohol use and related consequences among drinkers and non-drinkers	Incoming college freshmen at a northern California public university (622)	Web-based college alcohol education course (College Alc)	Among freshmen who were regular drinkers before college, College Alc reduced the frequency of heavy drinking (d = 0.15), drunkenness (d = 0.09), and negative alcohol-related consequences (d = 0.18). Freshmen who did not report any past 30-day alcohol use before college, College Alc did not experience any beneficial effects.
Brendryen et al (2008) [61]	Smoking	To assess the long-term efficacy of a fully automated digital multimedia smoking cessation intervention	People willing to quit smoking, aged 18 years or older, smoked 10 or more cigarettes daily, and had access to the Internet, email and a cell phone on a daily basis (290)	Fully automated, digital smoking cessation intervention including Web pages, SMS, interactive voice response, emails (Happy Ending)	Participants in the treatment group reported clinically and statistically significantly higher repeated point abstinence rates than control participants. (20% treatment vs 7% control, odds ratio [OR] = 3.43, 95% confidence interval [CI] = 1.60 - 7.34, *P* = .002)
Brendryen et al (2008) [62]	Smoking	To assess the long-term efficacy of a fully automated digital multimedia smoking cessation intervention	People willing to quit smoking, aged 18 years or older, smoked 10 or more cigarettes daily and had access to the Internet, email, and a cell phone on a daily basis (396)	Fully automated, digital smoking cessation intervention including Web pages, SMS, interactive voice response, emails (Happy Ending)	Participants in the treatment group reported clinically and statistically significantly higher repeated point abstinence rates than control participants. (22.3% treatment vs 13.1% control; OR = 1.91, 95% CI 1.12 - 3.26, *P* = .02)
Buller et al (2008) [63]	Smoking	To reduce smoking by children in grades 6 through 9 by convincing those who had not smoked not to start and persuading those who had already tried smoking to stop	Sixth to ninth graders from Australia and the United States (2077)	Tailored, Web-delivered smoking prevention program for adolescents (Consider This)	No statistically significant differences between groups were found
Danaher et al (2006) [43]	Smokeless tobacco use	To define participant exposure measures to a Web-based program for smokeless tobacco cessation	Recruited smokeless tobacco users (2523)	Web-based smokeless tobacco cessation intervention (ChewFree.com, enhanced)	Participants in the enhanced condition made more visits and spent more time accessing their assigned website than did participants assigned to the basic condition website.
Escoffery et al (2004) [51]	Smoking	To develop and conduct a process evaluation of a Web-based smoking cessation intervention for college smokers	College smokers (70)	Web-based smoking cessation program for college smokers (Kick It!)	No statistically significant differences between groups

Study Author (Year)	Problem Domain	Primary Objective of the Study	User Context (Number of Participants)	Use and Technology Context (Intervention)	Summary of Findings
Etter (2005) [64]	Smoking	To compare the efficacy of two Internet-based, computer-tailored smoking cessation programs	College smokers (11,969)	Web-based, computer-tailored smoking cessation (Stop-tabac.ch)	Statistically significant differences in quit rates in smokers in the contemplation stage favoring the original program. (OR = 1.54, 95% CI = 1.18 - 2.02, P = .002)
Finfgeld-Connett and Madsen (2008) [49]	Problem drinking	To evaluate the effectiveness of a Web-based, self-guided treatment program for women with problem drinking habits who live in rural areas of Missouri	Adult women with problem drinking habits living in Missouri counties (44)	Web-based, self-guided treatment program for problem drinking (intervention name not reported)	No statistically significant results
Hester et al (2009) [65]	Drinking moderation	To evaluate the effectiveness of a Web-based moderation training	Heavy drinkers (84) who responded to a newspaper recruitment ad (Albuquerque, New Mexico) Alcohol Use Disorders Identification Test (AUDIT) scores >7; Drinking >10 standard drinks per week; not currently abstaining; interest in moderating their consumption; aged 21 or older; Internet access at home	Internet-based program and use of the online resources of Moderation Management (MM)	At 3-month follow-up both groups significantly reduced their drinking. Both groups also significantly reduced their alcohol-related problems. Relative to the control, the experimental group had better outcomes on percent days abstinent.
Japuntich et al (2006) [50]	Smoking	To evaluate the impact of the program in an efficacy evaluation context	Smokers (at least 18 years old) motivated to quit smoking, 134 participants were recruited in a research center in Milwaukee, Wisconsin; 150 participated in a research center in Madison, Wisconsin (284)	Web-based smoking cessation and relapse prevention intervention (CHESS SCRP)	No statistically significant differences between groups
Matano et al (2007) [66]	Problem drinking	To pilot test an interactive Web-based intervention for reducing alcohol consumption	Employees of a work site in the Silicon Valley region of California, categorized as low or moderate risk for alcohol problems (145)	Interactive Web-based intervention for reducing alcohol consumption (CopingMatters)	The sample size was inadequate for evaluating treatment effects on drinking [66].
McKay et al (2008) [67]	Smoking	To describe the 6-month follow-up results of an RCT where participants were randomly assigned to either a Web-based smoking cessation program or a Web-based exercise enhancement program	Smokers at least 18 years of age interested in quitting within the next 30 days, willing to engage in moderate physical activity, access to the Internet (2318)	Web-based tailored smoking cessation (Quit Smoking Network; QSN)	No between-condition differences in smoking abstinence were found at 3- and 6-month follow-up assessments.
Muñoz et al (2006) [68]	Smoking	To compare a standard smoking cessation intervention to the same guide plus a mood management intervention	English- or Spanish-speaking smokers; ≥ 18 years old, smoking ≥ 5 cigarettes/day, using email at least once weekly and intending to quit in the next month; recruited from general population in USA (568)	Web-based intervention providing standard cessation information, tailored advice; individually timed educational messages (ITEMs); online mood management (MM) course (Guía)	ITEMs increased the effectiveness of the Guía. However, MM reduced quit rates, at times significantly so.

Study Author (Year)	Problem Domain	Primary Objective of the Study	User Context (Number of Participants)	Use and Technology Context (Intervention)	Summary of Findings
Muñoz et al (2009) [69]	Smoking	To examine abstinence rates of an Internet smoking cessation intervention and whether providing additional elements to a static Internet stop-smoking guide increases them	Spanish- and English-speaking participants (worldwide) were recruited using online campaigns; ≥ 18 years old, smoking ≥ 5 cigarettes/day, using email at least once weekly and intending to quit in the next month (1000)	Condition 1 was the “Gui´a Para Dejar de Fumar,” a static National Cancer Institute evidence-based stop smoking guide; Condition 2 consisted of Condition 1 plus ITEMs; Condition 3 consisted of Condition 2 plus MM; and Condition 4 consisted of Condition 3 plus a “virtual group” (an asynchronous bulletin board)	No significant differences among the four conditions were found.
Patten et al (2006) [70]	Smoking	To test the efficacy of a home-based, Internet-delivered treatment for adolescent smoking cessation	Adolescent smokers aged 11-18 years (139)	Web-based smoking cessation intervention tailored to adolescents (Stomp Out Smokes; SOS)	No statistically significant differences between groups were found.
Riper et al (2007) [71]	Problem drinking	To determine the effectiveness of a self-help intervention for adult problem drinkers	Adult Dutch problem drinkers (261)	Web-based self-help intervention (Drink Less)	At follow-up, 17.2% of the intervention group participants had reduced their drinking within the guideline norms; in the control group this was 5.4% (OR = 3.66, 95% CI = 1.3-10.8, *P* = .006, number needed to treat [NNT] = 8.5). The intervention subjects decreased their mean weekly alcohol consumption significantly more than control subjects, with a difference of 12.0 standardized units (95% CI 5.9 - 18.1, *P* < .001, standardized mean difference 0.40).
Severson et al (2008) [44]	Smokeless tobacco use	To test the impact of an interactive, tailored Web-based intervention versus a more linear, text-based website	Recruited smokeless tobacco users (2523)	Interactive, tailored Web-based intervention (ChewFree.com, enhanced)	Participants in the enhanced condition quit at significantly higher rates (vs basic condition). Abstinence was 40.6% in the enhanced condition vs 21.2% in the basic condition (*P* < .001). Using intent-to-treat analysis, quit rates were 12.6% vs 7.9% (*P* < .001)
Stoddard et al (2008) [72]	Smoking	To determine the use and satisfaction with two versions of a smoking cessation website, one of which included an asynchronous bulletin board	Adult federal employees or contractors to the federal government who responded to an email and indicated a willingness to quit smoking in 30 days (1375)	Smoking cessation website (Smokefree.gov, added bulletin board)	No statistically significant differences between groups were found.(Time spent on the website was significantly longer for the intervention subjects than for the control subjects.)

Study Author (Year)	Problem Domain	Primary Objective of the Study	User Context (Number of Participants)	Use and Technology Context (Intervention)	Summary of Findings
Strecher et al (2005) [73]	Smoking	To assess the efficacy of Web-based tailored behavioral smoking cessation program among nicotine patch users	Smokers in the United Kingdom and Republic of Ireland who purchased a certain patch and connected to a website (3971)	Web-based tailored behavioral smoking cessation (CQ Plan)	Continuous abstinence rates at 6 weeks were 29.0% in the tailored condition vs 23.9% in the nontailored condition (OR = 1.30, *P* < .001), at 12 weeks continuous abstinence rates were 22.8% versus 18.1%, respectively (OR = 1.34, *P* < .001) (Satisfaction with the program was significantly higher in the tailored than in the nontailored condition)
Strecher et al (2008) [74]	Smoking	To determine (1) whether engagement in a Web-based smoking cessation intervention predicts 6-month abstinence, (2) whether certain groups are more likely to have lower engagement, and (3) whether particular program components influence engagement	Smokers, participants from two large health maintenance organizations (1866)	Web-based program for smoking cessation and relapse prevention (intervention name not reported)	The total number of Web sections opened was related to subsequent smoking cessation. More personalized source and high-depth tailored self-efficacy components were related to a greater number of Web sections opened.
Swartz et al (2006) [75]	Smoking	To test the short-term efficacy of an automated behavioral intervention for smoking cessation delivered via a website	18 years or older, smoking cigarettes on a daily basis, considering quitting smoking in the next 30 days, and being able to access the website. (351)	Video-based website (1-2-3-Smokefree)	At 90 days, the cessation rate was 24.1% for treatment group versus 8.2% for the control group (*P* = .002). Using an intent-to-treat model, 12.3% of the treatment group were abstinent versus 5.0% in the control group (*P* = .015)
Woodruff et al (2007) [76]	Smoking	To test an innovative approach to smoking cessation that might be particularly attractive to adolescent smokers	Adolescent smokers in high school (136)	Web-based counseling program, virtual world chat room for adolescent smoking cessation (Breathing Room)	At the immediate postintervention assessment, intervention group participants were significantly more likely to report that they had abstained from smoking during the past week (*P* < .01), smoked fewer days in the past week (*P* < .001), smoked fewer cigarettes in the past week (*P* < .01), and considered themselves former smokers (*P* < .05). At a 1-year follow-up assessment, only the number of times quit was statistically significant (*P* < .05).

Persuasion Context: The Intent

Persuader

All of the articles stated a primary objective of the study (see Table 2), thus revealing a motive to persuade the users of the system. Most commonly, the Web-based interventions seemed to have been established by relatively small teams of people with varying expertise and background.

Change Type

In 8 of the articles, the intended change type was explicated. There were examples of reporting the intended change type in a simple and clear manner. For example,

“The self-help program proceeds in four successive stages: (1) preparing for action; (2) goal setting; (3) behavioral change; and (4) maintenance of gains and relapse prevention” [71]; “Participants were reassessed after 90 days for their alcohol consumption to assess changes in their drinking behavior” [66]; and “The QSN condition provided smoking cessation information and behavior change strategies while the Active Lives condition provided participants with physical activity recommendations and goal setting" [67].

Persuasion Context: The Event

Use Context

Our findings confirm the claim put forward by Griffiths and colleagues [31], who stated as a result of their systematic review that a number of studies gave no reason for using the Internet as the mode of delivery other than stating that the software application exists and needs to be evaluated. None of the articles reported the use context, for example, user groups, problem-domain dependent features, and attention drawers in detail. A high abstraction level in system descriptions makes it difficult to grasp what kind of interaction, and under what circumstances, really takes place through the system and to what extent the potential outcome really is due to the intervention.

User Context

The user context refers to characteristics of the individual user. All articles reported the user context to some extent. The authors of published articles may possess detailed information about users and their individual characteristics, but for some reason it does not seem to be clearly reflected in the dissemination.

Technology Context

The interventions included in the review were Web-based, and all of them aimed at persuading the users in some way. The majority of the research articles (14 out of 23) presented the technology context concisely. Screenshots are worthwhile, even though they do not reveal the backbone of the system, the information architecture, or the flow between parts of the system and its content. Escoffery et al [51] reported a brief flowchart of the development and evaluation process of the intervention, and Swartz et al [75] presented an overview flowchart of the intervention program. In our opinion, these types of charts are beneficial in helping readers to understand what actually takes place (and when) through the interventions.

Persuasion Context: The Strategy

Route

All of the reviewed studies seemed to rely on indirect routes for persuasion. However, due to the limited descriptions of the route utilized, it was not possible to confirm this fully. Of the studies reviewed, 14 revealed the underlying theories or methods behind the intervention to some extent. The most common theories were social cognitive theory (in 5 studies) and stages of change (in 4 studies), while in 5 of the studies the application of cognitive-behavioral methods was reported.

Message

One of the key decision points in developing interventions is the selection of messages for the intervention [40]. According to Ritterband et al [22], the message focuses on the source and style of the content, and it provides important information about who created the content and how it was presented for the users. These are hypothesized to impact user engagement and other mechanisms of change, including the acquisition of knowledge and motivation. In the PSD model, *the source* is called *the persuader*. All of the reviewed articles described the message at least partially.

Persuasive Features: Primary Task Support

The functionalities in the primary task category support the carrying out of the user's primary task. Persuasion techniques in this category [26] include *reduction, tunneling, tailoring, personalization, self-monitoring, simulation, and rehearsal* (compare Fogg [27]). These were reported relatively well in the reviewed studies.

Reduction

All of the reviewed articles described functionality countable as reduction, that is, the system reduces complex behavior into simple tasks helping users to perform the target behavior. This is important because a system that guides users through a process or experience provides opportunities to persuade along the way. For instance, in spite of the alcohol treatment program described in Finfgeld-Connett and Madsen [49], being complex (consisting of 8 reference modules and 15 decision-making modules) its use seemed to have been made easy for the end-users. For another instance, Swartz et al [75] described the modular structure of the intervention as "benefits of stopping smoking, overcoming common barriers to cessation, strategies for avoiding situations that prompt cravings, strategies for dealing with cravings, and setting a quit date."

Self-monitoring

A system keeping track of the user's performance or status supports achieving one's goals. Not surprisingly, this type of self-monitoring functionality was found in all of the included articles. For example, the system described in Escoffery et al [51] allowed participants to review their smoking by sending immediate feedback forms and copies of the personalized assessments to their email accounts. An et al [59] reported that participants received a weekly email invitation to visit the study website to report on health and lifestyle habits for the prior week, whereas Finfgeld-Connett and Madsen [49] described decision-making modules that included daily alcohol monitoring. Brendryen et al [61] stated that a major focus in their application was to ensure that participants comprehended that self-awareness, self-monitoring, active participation, and engagement are essential elements for reaching personal goals.

Simulation

Enabling users to observe the link between the cause and its effect is regarded as simulation. Simulations that educate users about certain topics can leave a lasting impact that transfers to the real world [77]. In all, 14 interventions featured at least some sort of simulation. For instance, Stoddard et al [72] described an interactive smoker's risk tool simulating changes in the risk of death due to smoking based on the smoker's history and time of quitting. In the Woodruff and colleagues [76] study, a virtual pathology lab showed pictures of damaged organs and premature aging, providing a scenario for discussing the short- and long-term health effects of smoking.

Personalization

A persuasive system may offer personalized content and services for its users. In order for the content to be personalized, the user has to disclose some personal information about herself, for example, through registration or by creating a personal profile [78]. The quality of Web personalization depends on how well the content generated by the personalization agent matches the preferences of the user in a particular domain [79]. Features of personalization were described in 13 articles. An et al [59] reported that email message content was personalized by peer coaches using information provided by participants during their weekly visits to the website. The enhanced website in Severson et al [44] included a guided, interactive program ("Personal Quitting Assistant") to help each user create an individual plan for quitting smokeless tobacco and preventing relapse. The intervention in Swartz et al [75] included creating a personalized quit plan calendar with individualized tips.

Tailoring

According to Rimer and Kreuter [80], studies related to tailoring should explore how and under what circumstances tailoring works and how its effects may be optimized. Tailoring the content to meet the potential needs, interests, personality, usage context, and/or other factors relevant to a user group is likely to increase the persuasiveness of the system. In all, 8 articles described tailoring functionality embedded in the actual interventions. For example, Strecher et al [73] reported that information collected in the enrollment questionnaire of their application was used to tailor the Web-based intervention materials. McKay et al [67] described tailoring portions of the program content to match each participant's smoking status in their application. The intervention in Swartz et al [75] consisted of 13 separate versions or strands, including 12 demographically targeted versions and 1 multicultural version, each with the same basic structure and content. The tailored versions were based on user's sex, age, and ethnicity.

Tunneling

Tunneling may enhance the change process since the user is led through a predetermined sequence of steps and receives the most appropriate content, particularly at a proper time [43]. A total of 4 articles described tunneling functionality. Buller et al [63] reported that program progression was controlled in their case by teachers. The *current* and *previous* modules were available to students, but students could not progress forward into subsequent modules until the teacher revealed the password to do so. In McKay et al [67], the intervention condition utilized hybrid information architecture (see [81]) where first-time users were tunneled through a series of tailored pages in order to introduce them to the key concepts and strategies of a behavioral program for quitting smoking. Severson et al [44] reported that in their "Planning to Quit" module users progressed linearly through content addressing readiness to quit, reasons for quitting, and level of dependence. In Hester et al [65], users were able to go through the program's modules sequentially, or they could choose the most relevant modules for them.

Rehearsal

Rehearsing a behavior can enable people to change their attitudes or behavior in the real world. Only 1 article reported this type of functionality; virtual locations (in the "Breathing Room" virtual world) such as a teen dance club and a fast food restaurant provided settings to discuss social influence to smoke and relapse, and an amphitheater was available for additional virtual meetings and discussions [76].

Persuasive Features: Dialogue Support

Dialogue support defines the key principles in keeping the user active and motivated in using the system and helping the user to reach the intended behavior. The principles in the dialogue support category are *praise*, *rewards*, *reminders*, *suggestion*, *similarity*, *liking*, and *social role*. Surprisingly, dialogue support was an underreported area in the reviewed papers.

Reminders

A persuasive system should remind users of their target behavior during the use of the intervention. A recent systematic review showed that the use of periodic prompts could be effective in behavior change interventions [82]. Of the reviewed studies, 10 utilized these kinds of reminders. In Stoddard et al [72], all participants received email reminders. During the first 2 weeks of their study on smoking cessation, participants received four email reminders unless they had previously expressed their wish to discontinue their participation. Each email included advice on quitting, a brief message encouraging use of the intervention, as well as the time frame of the future follow-up assessments. In Brendryen et al [61], if users did not log on to the program or answer the log-off call, they received a reminder call and up to two reminder text messages. The reminders were fully automated. Japuntich et al [50] described nonautomated reminders; if participants went a week without logging onto the system, the staff telephoned them and reminded them to log in.

Social Role

A system adopting a social role (eg, doctor or teacher) may be more persuasive. None of the reported interventions seemed to utilize social role per se. However, 9 of the articles described ask-an-expert service, which is related to social role (but also falls under expertise and authority in credibility support).

Suggestion

A system should provide the user with fitting suggestions at proper moments during the system use. This kind of suggestion was featured in 4 articles. In Japuntich et al [50], the program suggested different articles or other services to the user based on his or her responses to the *check-in* (eg, smokers reporting depression were encouraged to use the cognitive behavioral therapy service). At the start of each session, the program in McKay et al [67] displayed online prompts recommending the review of content that a participant had not yet explored. In Stoddard et al [72], the program provided *Did you know?* messages containing appropriate links and information. The content of the program in Strecher [73] included suggestions for coping with nicotine replacement therapy and encouraging compliance.

Similarity

Individuals are more readily persuaded through systems resembling themselves in some meaningful way. For example, a system aimed at teenagers should employ youthful phrases and imagery. This principle is known as similarity (compare *social identity cues* [78]), and it was found in 4 of the articles. In Escoffery et al [51], college smokers provided ideas for components to be incorporated into the Web-based intervention, for example, graphical presentations of information, providing only small amounts of text, sharing of stories/situations, quizzes, and other interactive features. These were then taken into account in the development of the program content. Similarity was utilized in Bersamin et al [60] as four of the five units in the program for college students ("College Alc") included a streaming video clip depicting college students in an alcohol-related context. In addition, it utilized student-generated harm-prevention plans.

Liking

An attractive system is likely to be more persuasive. This principle is known as *liking,* and it was addressed in 3 articles. Woodruff et al [76] stated that the primary goal of their application was to test an innovative approach for smoking cessation that might be particularly appealing to adolescent smokers. In Buller et al [63], audio narration, graphics, animation, sound effects, and music were utilized to create a rich multimedia environment to stimulate user engagement. In Escoffery et al [51], as a part of the formative research before the development of the program, college smokers were asked about their Internet usage and features that they liked to learn about potential elements to be added into the Web application.

Rewards

The system should reward the user for achieving self-set goals, for example [54]. None of the articles reported this type of automated functionality. An example implementation could be, for instance, that the user would be rewarded with a virtual trophy upon completion of a certain task. However, Severson et al [44] reported that their *Staying Quit* module provided tailored information and behavioral strategies on eight major topics including *Reward Yourself*.

Praise

A system could praise users via words, images, symbols, or sounds based on their behaviors. By offering praise, a system can make users more open to persuasion. Quite surprisingly, this technique was not featured in any of the reviewed studies.

Persuasive Features: Credibility Support

Credibility is a persuasive element (eg, [78,83,84]). Harris et al [85] suggested that perhaps even seemingly superficial design elements of a website can influence responses to health-risk information. In their study, credibility cues affected the engagement with the site and influenced subsequent health behavior and cognition. According to Briñol and Petty [29], confidence in one's thoughts is likely to be undermined if the received message is not credible.

The PSD model describes seven principles for supporting system credibility: *trustworthiness, expertise, surface credibility, real world feel, authority, third party endorsements, and verifiability*. The analysis of these is more a continuum than a dichotomy. For this reason, explicit numbers and percentages are not given here. However, based on the textual descriptions of the interventions, we suggest that most of them incorporated—to a notable degree—expertise (system provides information demonstrating knowledge, experience, and competence), verifiability (system provides means to verify the accuracy of site content via outside sources), authority (system quotes an authority, such as a statement or norms by an authoritative health institute), and trustworthiness (system provides truthful, fair, and unbiased information). Trustworthiness is crucial in website engagement as users will engage with sites they perceive trustworthy and navigate away from those they mistrust [78].

Naturally, evaluating perceived credibility is more or less subjective. People make initial assessments of the system credibility based on a firsthand inspection. This principle is called *surface credibility*. A persuasive system should provide information of (and means to communicate with) the organization and/or actual people behind its content and services. This feature is called *real-world feel*. To fully evaluate both surface credibility and real-world feel would require a hands-on approach on the actual implemented interventions.

A persuasive system could also provide endorsements from respected and renowned sources, for example, a recommendation by an authoritative health organization, an award for excellence in usability, or a privacy seal to ensure confidentiality. However, none of the articles reported utilizing third party endorsements.

Persuasive Features: Social Support

According to Uchino [86], social support may refer to the aspects of the social network (groups or familial ties), specific behaviors (eg, emotional or informational support), or our perceived availability of support resources that may be shaped early in life. In the PSD model, the social support category describes how to design the system so that it motivates users by leveraging social influence. The model operates with the following persuasion techniques: *social learning, social comparison, normative influence, social facilitation, cooperation, compensation,* and *recognition*.

The social learning principle means that individuals may be more motivated to perform a target behavior if they can observe others performing the behavior while using the system. A closely related principle is social comparison: users will be more motivated to perform the target behavior if they can compare their performance with the performance of others. Users are also more likely to perform target behavior if they are able to observe others performing the behavior or are being observed by others. This principle is called *social facilitation.*

The most common means to providing social support were asynchronous peer discussion forums and synchronous chat rooms. A variety of other online social support features were also utilized. For a full description of studies regarding social learning, social comparison social facilitation, and normative influence, see Table 2.

Table 2. Social support in the Web-based interventions

Study Author (Year)	Social Learning, Comparison and/or Facilitation	Normative Influence	Other Support
An et al (2008) [59]	No	Not reported	Email exchange with peer coach
Bersamin et al (2007) [60]	Discussion forum	Streaming video clips, College Alcohol Use unit	Not reported
Brendyen et al (2008) [61,62] (Note: 2 articles are combined here as they both study the Happy Ending system)	No	Not reported	Pre-recorded audio messages for relapse prevention, Interactive Voice Response-based craving helpline
Buller et al (2008) [63]	No	Prevention content focused on social influence and aimed, for example, to correct inexact perceptions of tobacco use norms	Not reported
Danaher et al (2006) [43] Severson et al (2008) [44] (Note: these articles are combined because they study the same website with the same participants)	Peer-to-peer forum	Testimonial videos	Ask-an-expert forum
Escoffery et al (2004) [51]	Stage-matched discussion forums	Shared personal stories area	Not reported
Etter (2005) [64]	Discussion forums, chat rooms	Personal stories written by current and former smokers	Not reported
Finfgeld-Connett and Madsen (2008) [49]	Asynchronous bulletin board, synchronous chat featured	Not reported	Private messaging to the researcher and other users
Hester et al (2009) [65]	Online mutual-help support community	Not reported	Online and face-to-face meetings
Japuntich et al (2006) [50]	Discussion group, chat room (trained counselor available)	The smoking-related topics included facts about smokers, smoking, and cigarette companies	Ask-an-expert service
Matano et al (2007) [66]	No	Individualized feedback with normative data	Not reported
McKay et al (2008) [67]	Peer-to-peer forum	Not reported	Ask-an-expert forum
Muñoz et al (2006, 2009) [68,69]	Asynchronous bulletin board	Not reported	Mood management online course
Patten et al (2006) [70]	Discussion support group	Videos of personal stories	Private email service with an expert
Riper et al (2008) [71]	Moderated peer-to-peer discussion forum	Streaming video (quitting information and testimonials)	Not reported
Stoddard et al (2008) [72]	Asynchronous bulletin board	Not reported	Ask-an-expert forum
Strecher et al (2005) [73]	No	Not reported	Behavioral support email messages
Strecher et al (2008) [74]	No	A section with a narrative success story	Not reported
Swartz et al (2006) [75]	No	Personalized video segments	Not reported
Woodruff et al (2007) [76]	Virtual world chat room, virtual locations	Topics covered, for example, peer influence; virtual locations in “Breathing Room” virtual world	Smokers interacted with each other as well as with the counselor

A system can apply normative influence, in other words, positive peer pressure to enhance the likelihood that an individual will adopt a target behavior. Buller et al [63] and Matano et al [66] reported that users were working individually with the program to create a sense of privacy so they would unfold their smoking intentions.

In addition to the aforementioned ways for providing social support, a system can persuade users to adopt a target attitude or behavior by leveraging human beings' natural drive to compete and cooperate. Competition was reported in 2 and cooperation in 1 of the analyzed papers. In An et al [59], the intervention site actively promoted a "Quit and Win" contest and included links to the online sign-up for this contest. An example of cooperation was found in Woodruff et al [76], as the students were working as a group in the virtual environment.

By offering public recognition to an individual or a group, a system can increase the likelihood that a person and/or group will adopt a target behavior. Only 1 of the articles presented functionality countable as recognition. In Woodruff et al [76] a message, "Guero's [a person's name] been quit for a month!" was presented on billboards throughout the virtual world.

Discussion

Findings

Of the 23 articles included in the review, 20 primarily measured health behavior outcomes. Of these articles, 12 reported statistically significant differences between groups. Overall, 3 articles assessed program utilization, and 2 of them reported positive findings (see Table 1).

The primary task support components were reported relatively widely in the reviewed studies. Reduction, self-monitoring, simulation, and personalization seem to be the most used ways to support accomplishing a user's primary task. This is an encouraging finding because reduction and self-monitoring can be considered as the key elements of primary task support. The utilization of tailoring was surprisingly low. The lack of tailoring may imply that the interventions are targeted for too broad an audience. It would be reasonable to assume that different approaches were needed for different kinds of user groups. Elements of dialogue support were mostly underutilized in the interventions. Leveraging reminders [82] was the most common way to enhance the user-system dialogue.

Credibility issues are crucial in website engagement as users will engage with sites they perceive credible and navigate away from those they do not find credible. Based on the textual descriptions of the interventions, we cautiously suggest that most of them were relatively credible. Nevertheless, it seems to be, as Danaher and Seeley [19] eloquently put it, that "credibility is in the eye of the beholder."

The prevalence of social support in the reviewed interventions was encouraging (see Table 2). Social support provided by the systems is often based on peer support. According to Eysenbach [36], there is not yet very strong evidence for what type of peer-based social support the systems really should provide. In their systematic review, Shahab and McEwen [17] argued that while chat forums could potentially aid in stopping smoking through the providing extra social support, it still seems to be that entirely automated interventions are even more effective. Bennett and Glasgow [14] claimed that there are no examples of trial designs that would enable a systematic investigation of the potential benefits of social networking. In this review, we were not able to confirm the effect of social support components on health behavior outcomes. This is mainly due to the fact that only 1 of the included articles [72] explicitly studied the effect of adding a social component (virtual community) within the intervention. Furthermore, many research and design issues still need to be resolved about the type of online support, for example, whether it should be expert-led versus user-driven, moderated versus unmoderated, synchronous versus asynchronous, or open access versus restricted access, among other issues.

Applying the Persuasive Systems Design Model

Evaluating the effectiveness of specific persuasive features within Web-based interventions is difficult since the features are not usually explicitly tested. According to Kypri and Lee [87], the descriptions and analyses of how interventions are developed are often absent from scientific literature due to space constraints. However, we think that the presentation of detailed information about the theoretical basis, functionality, content, and structure of a Web-based intervention helps to interpret the results and conduct evaluations as on a more finely grained level. Ahern [88] points out that randomization to experimental groups or conditions remains the gold standard for evaluating intervention efficacy but may not provide the most relevant information for dissemination. Furthermore, methodological challenges and latent scientific foundations in researching Internet-based interventions are acknowledged by many researchers (eg, [19]). Glasgow [89] stated that the diversity in content area, disciplines involved, and publication outlets is the reason that the consistency in how Internet-based solutions have been conceptualized, reported, and evaluated has been low. Finally, the brevity of intervention descriptions makes it more difficult to draw generalizable results (see [8,13]).

Due to these issues, the application of the PSD model turned out to be relatively laborious. In the present review, we relied on textual descriptions of the interventions, thus being able to provide only a limited synthesis. Regardless of its wide coverage, the PSD model is not an exhaustive list of persuasive features, and also some of the features are overlapping (eg, social learning/comparison/facilitation and liking/similarity) and thus rather difficult to analyze. New persuasion techniques may also be identified in the future. The PSD model has been built in such a manner that it may evolve, but even as it stands now, it is an important asset for any health behavior change system developer.

Limitations and Strengths of the Review

Analyzing persuasive design is a challenging task. When conducting an analysis such as described in the present review, potential bias lies in the interpretation of the articles. Nevertheless, in extracting and categorizing persuasive features, we rigorously observed if the authors clearly stated the described variables. Obviously, the articles did not necessarily follow the very same terminology as found in the PSD Model. Thus, the analysis was based on interpretive categorization.

This systematic review focused on randomized controlled trials, thus excluding potentially meritorious studies. (Quasi-experimental studies were not found in the search process.) A meta-analysis was not conducted due to the heterogeneity of the studies. Overall, there are already several reviews on Web-based (or similar) interventions. To our knowledge, the present review is the first systematic review to address persuasive system features in Web-based interventions for substance use.

Conclusion

In this review we examined the persuasive system features of the included Web-based interventions. We think that this type of novel approach is useful for current and future research for recognizing what kind of tactics in present systems have been utilized to motivate people in achieving better health.

However, at this point, linking specific persuasive features to outcomes is difficult, relying only on brief textual descriptions of the interventions. Also, it is not possible to determine the (perceived) credibility of a Web-based intervention based on reading an article. We acknowledge that many studies have examined, for example, the role of tailoring in health behavior change interventions [10,11], but the *persuasiveness* of a particular component/application/system is a more complex issue and has yet to be tackled in future endeavors.

We are not implying that the mere presence of persuasive features is enough. The development of Web-based and other similar interventions is a highly elaborate and a multifaceted issue. Still, it is relevant to consider the technological aspects since the Web and related technologies are being used as a delivery channel. Atienza and colleagues [90] pertinently remarked that “health information technology does not occur in a vacuum, but rather technologies exist within social systems.” In order for widespread adoption, dissemination, and extended use of technology-enabled health behavior change interventions to take place, it is necessary to investigate not only how the interventions affect individuals, but also how individuals interact with technology and each other [90]. Further research is also warranted to increase our understanding of how and under what circumstances specific persuasive features (either in isolation or collectively) lead to positive health outcomes in Web-based health behavior change interventions across diverse contexts and populations.

Acknowledgments

We wish to thank the Graduate School on Software Systems and Engineering (SoSE), the Finnish Funding Agency for Technology and Innovation (TEKES) and the RichWeb and Eliksiiri/Salwe projects for funding this research. The second author would like to thank the Academy of Finland. We also express our gratitude to the reviewers for their constructive comments.

Conflicts of Interest

None declared

Multimedia Appendix 1

Excluded articles (N=39): Alcohol

[PDF file (Adobe PDF File), 57 KB - jmir_v13i3e46_app1.pdf]

Multimedia Appendix 2

Excluded articles (N=14): Smoking

[PDF file (Adobe PDF File), 43 KB - jmir_v13i3e46_app2.pdf]

References

1. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
2. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]
3. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345. [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
4. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
5. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937 [FREE Full text] [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
6. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

7. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
8. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Internet Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
9. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
10. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
11. Enwald HP, Huotari ML. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. J Med Internet Res 2010;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
12. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. Clin Psychol Rev 2009 Jul;29(5):410-420. [doi: 10.1016/j.cpr.2009.04.001] [Medline: 19403215]
13. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
14. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr 29;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
15. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16. [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
16. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
17. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
18. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
19. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39. [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
20. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
21. Consolvo S, McDonald DW, Landay JA. Theory-driven design strategies for technologies that support behavior change in everyday life. In: Proceedings of the 27th international conference on Human factors in computing systems. New York, NY, USA: ACM; 2009 Presented at: CHI '09; April 4-9; Boston, MA, USA. [doi: 10.1145/1518701.1518766]
22. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27. [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
23. Oinas-Kukkonen H. Behavior change support systems: A research model and agenda. In: Lecture Notes in Computer Science, Vol. 6137. Berlin / Heidelberg: Springer; 2010 Presented at: The Fifth International Conference on Persuasive Technology; June 7-9, 2010; Copenhagen, Denmark. [doi: 10.1007/978-3-642-13226-1_3]
24. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
25. Kraft P, Drozd F, Olsen E. ePsychology: Designing theory-based health promotion interventions. Communications of the Association for Information Systems 2009;24(1, Article 24).
26. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems 2009;24(1, Article 28).
27. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco, CA: Morgan Kaufmann; 2003.
28. Chatterjee S, Price A. Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inform Assoc 2009;16(2):171-178. [doi: 10.1197/jamia.M2859] [Medline: 19074300]
29. Briñol P, Petty RE. Persuasion: Insights from the self-validation hypothesis. Adv Exp Soc Psychol 2009;41:69-118. [doi: 10.1016/S0065-2601(08)00402-4]
30. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. Int J Med Inform 2006 Jan;75(1):29-41. [doi: 10.1016/j.ijmedinf.2005.07.029] [Medline: 16129659]
31. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]

32. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
33. Evers KE, Prochaska JM, Prochaska JO, Driskell MM, Cummins CO, Velicer WF. Strengths and weaknesses of health behavior change programs on the internet. J Health Psychol 2006;8(1):63-70. [doi: 10.1177/1359105303008001435]
34. Cassell MM, Jackson C, Cheuvront B. Health communication on the Internet: an effective channel for health behavior change? J Health Commun 1998;3(1):71-79. [Medline: 10947375]
35. Baym NK, Zhang YB, Lin MC. Social interactions across media: Interpersonal communication on the internet, telephone and face-to-face. New Media Soc 2004;6(3):299-318. [doi: 10.1177/1461444804041438]
36. 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] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
37. IJsselsteijn W, de Kort Y, Midden C, Eggen B. Persuasive technology for human well-being: Setting the scene. Persuasive Technology, Lecture Notes in Computer Science 2006;3962/2006:1-5. [doi: 10.1007/11755494_1]
38. Fogg BJ. Creating persuasive technologies: An eight-step design process. In: Proceedings of the 4th International Conference on Persuasive Technology. New York, NY: ACM; 2009 Presented at: Persuasive Technology; April 26-29; Claremont, CA. [doi: 10.1145/1541948.1542005]
39. Zhang J. Human-centered computing in health information systems. Part 1: analysis and design. J Biomed Inform 2005 Feb;38(1):1-3. [doi: 10.1016/j.jbi.2004.12.002] [Medline: 15694880]
40. Fishbein M, Cappella JN. The role of theory in developing effective health communications. J Commun 2006 Aug;56(Supplement 1):S1-S17. [doi: 10.1111/j.1460-2466.2006.00280.x]
41. Tait RJ, Christensen H. Internet-based interventions for young people with problematic substance use: a systematic review. Med J Aust 2010 Jun 7;192(11 Suppl):S15-S21. [Medline: 20528701]
42. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. Addiction 2011 Feb;106(2):267-282. [doi: 10.1111/j.1360-0443.2010.03214.x] [Medline: 21083832]
43. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
44. Severson HH, Gordon JS, Danaher BG, Akers L. ChewFree.com: evaluation of a Web-based cessation program for smokeless tobacco users. Nicotine Tob Res 2008 Feb;10(2):381-391. [doi: 10.1080/14622200701824984] [Medline: 18236303]
45. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002 Mar;97(3):279-292. [Medline: 11964101]
46. Babor TF, Higgins-Biddle JC. Alcohol screening and brief intervention: dissemination strategies for medical practice and public health. Addiction 2000 May;95(5):677-686. [Medline: 10885042]
47. Copeland J, Martin G. Web-based interventions for substance use disorders: a qualitative review. J Subst Abuse Treat 2004 Mar;26(2):109-116. [doi: 10.1016/S0740-5472(03)00165-X] [Medline: 15050088]
48. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869. [Medline: 20332511]
49. Finfgeld-Connett D, Madsen R. Web-based treatment of alcohol problems among rural women. J Psychosoc Nurs Ment Health Serv 2008 Sep;46(9):46-53. [Medline: 18823000]
50. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8 Suppl 1:S59-S67. [Medline: 17491172]
51. Escoffery C, McCormick L, Bateman K. Development and process evaluation of a web-based smoking cessation program for college smokers: innovative tool for education. Patient Educ Couns 2004 May;53(2):217-225. [doi: 10.1016/S0738-3991(03)00163-0] [Medline: 15140462]
52. Oinas-Kukkonen H, Harjumaa M. A systematic framework for designing and evaluating persuasive systems. Persuasive Technology, Lecture Notes in Computer Science 2008;5033/2008:164-176. [doi: 10.1007/978-3-540-68504-3_15]
53. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Quarterly 2003;27(3):425-478.
54. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [Medline: 12237980]
55. Petty R, Cacioppo J. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205.
56. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50:179-211.
57. Dey AK. Understanding and using context. Personal and Ubiquitous Computing 2001;5(1):4-7. [doi: 10.1007/s007790170019]
58. Torning K, Oinas-Kukkonen H. Persuasive systems design: State of the art and future directions. In: Proceedings of the 4th International Conference on Persuasive Technology. New York, NY: ACM; 2009 Presented at: Persuasive '09; April 26-29; Claremont, CA. [doi: 10.1145/1541948.1541989]

59. An LC, Klatt C, Perry CL, Lein EB, Hennrikus DJ, Pallonen UE, et al. The RealU online cessation intervention for college smokers: a randomized controlled trial. Prev Med 2008 Aug;47(2):194-199. [doi: 10.1016/j.ypmed.2008.04.011] [Medline: 18565577]
60. Bersamin M, Paschall MJ, Fearnow-Kenney M, Wyrick D. Effectiveness of a Web-based alcohol-misuse and harm-prevention course among high- and low-risk students. J Am Coll Health 2007;55(4):247-254. [Medline: 17319331]
61. Brendryen H, Drozd F, Kraft P. A digital smoking cessation program delivered through internet and cell phone without nicotine replacement (happy ending): randomized controlled trial. J Med Internet Res 2008;10(5):e51 [FREE Full text] [doi: 10.2196/jmir.1005] [Medline: 19087949]
62. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]
63. Buller DB, Borland R, Woodall WG, Hall JR, Hines JM, Burris-Woodall P, et al. Randomized trials on consider this, a tailored, internet-delivered smoking prevention program for adolescents. Health Educ Behav 2008 Apr;35(2):260-281. [doi: 10.1177/1090198106288982] [Medline: 17114331]
64. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [doi: 10.2196/jmir.7.1.e2] [Medline: 15829474]
65. Hester RK, Delaney HD, Campbell W, Handmaker N. A web application for moderation training: initial results of a randomized clinical trial. J Subst Abuse Treat 2009 Oct;37(3):266-276. [doi: 10.1016/j.jsat.2009.03.001] [Medline: 19339137]
66. Matano RA, Koopman C, Wanat SF, Winzelberg AJ, Whitsell SD, Westrup D, et al. A pilot study of an interactive web site in the workplace for reducing alcohol consumption. J Subst Abuse Treat 2007 Jan;32(1):71-80. [doi: 10.1016/j.jsat.2006.05.020] [Medline: 17175400]
67. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008;10(5):e40 [FREE Full text] [doi: 10.2196/jmir.993] [Medline: 19017582]
68. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, et al. Toward evidence-based Internet interventions: A Spanish/English Web site for international smoking cessation trials. Nicotine Tob Res 2006 Feb;8(1):77-87. [doi: 10.1080/14622200500431940] [Medline: 16497602]
69. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine Tob Res 2009 Sep;11(9):1025-1034. [doi: 10.1093/ntr/ntp090] [Medline: 19640833]
70. Patten CA, Croghan IT, Meis TM, Decker PA, Pingree S, Colligan RC, et al. Randomized clinical trial of an Internet-based versus brief office intervention for adolescent smoking cessation. Patient Educ Couns 2006 Dec;64(1-3):249-258. [doi: 10.1016/j.pec.2006.03.001] [Medline: 16616449]
71. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
72. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. J Med Internet Res 2008;10(5):e53 [FREE Full text] [doi: 10.2196/jmir.1124] [Medline: 19097974]
73. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. Addiction 2005 May;100(5):682-688. [doi: 10.1111/j.1360-0443.2005.01093.x] [Medline: 15847626]
74. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
75. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
76. Woodruff SI, Conway TL, Edwards CC, Elliott SP, Crittenden J. Evaluation of an Internet virtual world chat room for adolescent smoking cessation. Addict Behav 2007 Sep;32(9):1769-1786. [doi: 10.1016/j.addbeh.2006.12.008] [Medline: 17250972]
77. Beard L, Wilson K, Morra D, Keelan J. A survey of health-related activities on second life. J Med Internet Res 2009;11(2):e17 [FREE Full text] [doi: 10.2196/jmir.1192] [Medline: 19632971]
78. Sillence E, Briggs P, Harris P, Fishwick L. A framework for understanding trust factors in web-based health advice. International Journal of Human-Computer Studies 2006;64(8):697-713. [doi: 10.1016/j.ijhcs.2006.02.007]
79. Tam KY, Ho SY. Web personalization as a persuasion strategy: An elaboration likelihood model perspective. Inform Syst Res 2005;16(3):271-291. [doi: 10.1287/isre.1050.0058]
80. Rimer B, Kreuter M. Advancing tailored health communication: A persuasion and message effects perspective. J Commun 2006;56(Supplement 1):S184-S201. [doi: 10.1111/j.1460-2466.2006.00289.x]
81. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [doi: 10.2196/jmir.7.2.e12] [Medline: 15914459]
82. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]

83. Pornpitakpan C. The persuasiveness of source credibility: A critical review of five decades' evidence. J Appl Soc Psychol 2004;34(2):243-281. [doi: 10.1111/j.1559-1816.2004.tb02547.x]
84. Wathen CN, Burkell J. Believe it or not: Factors influencing credibility on the web. J Am Soc Inf Sci Technol 2002;53(2):134-144. [doi: 10.1002/asi.10016]
85. Harris PR, Sillence E, Briggs P. The effect of credibility-related design cues on responses to a web-based message about the breast cancer risks from alcohol: randomized controlled trial. J Med Internet Res 2009;11(3):e37 [FREE Full text] [doi: 10.2196/jmir.1097] [Medline: 19709989]
86. Uchino BN. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. J Behav Med 2006 Aug;29(4):377-387. [doi: 10.1007/s10865-006-9056-5] [Medline: 16758315]
87. Kypri K, Lee N. New technologies in the prevention and treatment of substance use problems. Drug Alcohol Rev 2009 Jan;28(1):1-2. [doi: 10.1111/j.1465-3362.2008.00014.x] [Medline: 19320668]
88. Ahern DK. Challenges and opportunities of eHealth research. Am J Prev Med 2007 May;32(5 Suppl):S75-S82. [doi: 10.1016/j.amepre.2007.01.016] [Medline: 17466822]
89. Glasgow RE. Enhancing the scientific foundation of internet intervention research. Ann Behav Med 2009 Aug;38(1):46-47. [doi: 10.1007/s12160-009-9134-3] [Medline: 19816749]
90. Atienza AA, Hesse BW, Gustafson DH, Croyle RT. E-health research and patient-centered care examining theory, methods, and application. Am J Prev Med 2010 Jan;38(1):85-88. [doi: 10.1016/j.amepre.2009.10.027] [Medline: 20117562]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
CI: confidence interval
CONSORT: Consolidated Standards of Reporting Trials
ITEM: individually timed educational message
MM: mood management
NNT: number needed to treat
OR: odds ratio
PSD: persuasive system design
RCT: randomized controlled trial
SoSE: Graduate School on Software Systems and Engineering
TEKES: Finnish Funding Agency for Technology and Innovation

Edited by G Eysenbach; submitted 04.05.10; peer-reviewed by L Yardley, W LaMendola, B Bewick, ML Lustria; comments to author 25.06.10; revised version received 15.03.11; accepted 19.03.11; published 22.07.11.

Please cite as:
Lehto T, Oinas-Kukkonen H
Persuasive Features in Web-Based Alcohol and Smoking Interventions: A Systematic Review of the Literature
J Med Internet Res 2011;13(3):e46
URL: http://www.jmir.org/2011/3/e46/
doi:10.2196/jmir.1559
PMID:21795238

Original Paper

A Systematic Review of the Impact of Adherence on the Effectiveness of e-Therapies

Liesje Donkin[1], BSC, MSc; Helen Christensen[2], PhD; Sharon L Naismith[1], BA(Hons), Dpsych(Neuro); Bruce Neal[3], MB ChB, PhD; Ian B Hickie[1], MD, FRANZCP; Nick Glozier[1,4], MBBS, FRANZCP, PhD

[1]Brain & Mind Research Institute, The University of Sydney, Camperdown, Australia
[2]Centre for Mental Health Research, Australian National University, Canberra, Australia
[3]The George Institute for Global Health, The University of Sydney, Sydney, Australia
[4]Disciplines of Psychiatry and Sleep Medicine, Sydney Medical School, The University of Sydney, Sydney, Australia

Corresponding Author:
Liesje Donkin, BSC, MSc
Brain & Mind Research Institute
The University of Sydney
100 Mallet Street
Camperdown, NSW 2050
Australia
Phone: 61 293510520
Fax: 61 293510855
Email: Liesje.Donkin@sydney.edu.au

Abstract

Background: As the popularity of e-therapies grows, so too has the body of literature supporting their effectiveness. However, these interventions are often plagued by high attrition rates and varying levels of user adherence. Understanding the role of adherence may be crucial to understanding how program usage influences the effectiveness of e-therapy interventions.

Objective: The aim of this study was to systematically review the e-therapy literature to (1) describe the methods used to assess adherence and (2) evaluate the association of adherence with outcome of these interventions.

Methods: A systematic review of e-therapy interventions was conducted across disease states and behavioral targets. Data were collected on adherence measures, outcomes, and analyses exploring the relationship between adherence measures and outcomes.

Results: Of 69 studies that reported an adherence measure, only 33 (48%) examined the relationship between adherence and outcomes. The number of logins was the most commonly reported measure of adherence, followed by the number of modules completed. The heterogeneity of adherence and outcome measures limited analysis. However, logins appeared to be the measure of adherence most consistently related to outcomes in physical health interventions, while module completion was found to be most related to outcomes in psychological health interventions.

Conclusions: There is large variation in the reporting of adherence and the association of adherence with outcomes. A lack of agreement about how best to measure adherence is likely to contribute to the variation in findings. Physical and psychological outcomes seem influenced by different types of adherence. A composite measure encompassing time online, activity completion, and active engagements with the intervention may be the best measure of adherence. Further research is required to establish a consensus for measuring adherence and to understand the role of adherence in influencing outcomes.

(J Med Internet Res 2011;13(3):e52) doi:10.2196/jmir.1772

KEYWORDS

Adherence; persistence; online therapy; e-therapy; systematic review

Introduction

The past two decades have seen a shift from traditional face-to-face consultations to technology-driven interventions or e-therapies. Recent reviews have shown that e-therapies are both effective [1-3] and growing in popularity. This is supported by the level of publications relating to e-therapies: Medline and PsycINFO citations in the subject group “online therapy” rose from 12 citations between 1991 and 2000 to 709 citations from 2001 to September 2010.

A potential difficulty in evaluating these programs is adherence (see [4]). Little is known about the degree to which users' engagement matches the usage pattern for which the websites are designed. Also, little is known about the influence of program adherence on outcomes. Within the medication literature, adherence, “the extent to which a person's behaviour – taking medication, following a diet and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider” [5], and persistence, the act of adhering to treatment recommendations for the prescribed duration of time [6], are widely studied. These behavioral variables significantly influence medical [5] and psychotherapy [7-9] outcomes. In pharmaceutical trials, a dose–response curve is often plotted to understand the optimal level of medication to reach a desired response, and adherence is considered highly influential within this. In e-therapy, adherence may be just as important a consideration.

The eHealth equivalent of failing to persist with therapy is treatment dropout. Treatment dropout refers to when a user prematurely stops using the intervention. Some of these users may remain in the trial, completing the trial assessments, while others may choose to leave the trial. Those who choose to leave the trial early are said to have discontinued and are reflected in trial attrition rates. Such attrition can affect the ability of results to be generalized [10,11] and it undermines the statistical power of the trial. Many authors note that dropout rates are high, particularly in open-access trials [12-14] where the intervention is made available to the public with minimal or no cost. Entry into these trials is therefore open, with users being able to join at any time. Despite each user's choice to engage with the site, only a small proportion of users persist with the trial and associated follow-ups. However, with some programs, these figures may still be comparable to [15], or even lower than [16], the dropout rates found in traditional face-to-face therapies. Little is known about the impact of the degree of persistence on outcome in those who complete trial ratings and outcomes.

While an appreciation of persistence is important in evaluating e-therapy, an understanding of adherence to the program content, such as the completion of program modules or online activities, may be more so. As the field of e-therapy has been growing, so has the interest in potentially modifiable user factors that may influence adherence. Clearly, program content evolves from empirically supported research, but only recently have the exploration and manipulation of program factors become foci. Program usability and feasibility testing is increasingly common. Recent findings have indicated that greater use of computer relational skills, such as the use of empathy and social dialogue in the computer program, leads to increased program usage [17]. Many trials use pilots and usability studies to ensure the program functions as planned [18-22]. Several authors have begun to explore the impact of reminders and trial factors on dropout. Clarke and colleagues' [23] comparison of the Overcoming Depression on the Internet websites found that participants were more likely to use the program as recommended if they received reminders, and that this increased use was present regardless of the type of reminder (telephone versus postcard) that they received. Alternatively, Christensen et al [24] demonstrated that weekly tracking and reminders reduced attrition in a cognitive behavior therapy e-therapy intervention for depression. Christensen et al [25] also found that shorter cognitive behavior therapy e-therapy interventions were not as effective as longer interventions but that attrition rates were lower, potentially indicating an important trade-off between acceptability and effectiveness. It is thought that the variation in adherence and persistence may be due to the participant's discretion in using e-therapy [26], rather than therapy that is prescribed, as in a drug trial [13]. Engagement in an e-therapy trial tends to require more physical, cognitive, and time investments by the participant, compared to the relative ease of taking a medication daily. This may explain why e-therapy users are more prone to dropout and nonadherence.

E-therapy trials have an advantage over traditional trials when measuring persistence and adherence, in that more objective measures of intervention usage are readily available to researchers. Objective measures of persistence include metrics such as the number of times the participant accesses the program, and adherence measures include the time spent online, number of completed activities, and patterns of usage. Subjective measures, such as estimated time spent online, reporting on the completion of behaviorally based homework activities, and use of skills, can also be incorporated. Despite the relative ease of capturing these data in online interventions, few studies report these statistics. Merely reporting on adherence and dropout provides limited insight about the impact of adherence on program outcomes or the translation of program behaviors into daily life. Even fewer studies examine the role of adherence on outcomes. The few that do, by publishing only significant results, may result in publication bias.

Recent studies have begun to address this through the exploration of the relationship between program exposure and outcomes [27-30]. Although more data can often be collected in e-therapy trials, there is a need to consider what data are collected, how they are collected, and how adherence is defined. It is likely that the influence of design, application, and logistics of e-therapies on outcomes can become informed in the same way medication dosing affects outcomes. For instance, single daily dosing or polypills have been suggested as methods of improving adherence and outcomes in medication treatments.

To inform this development, we conducted a systematic review of the literature evaluating adherence in e-therapies. Within the context of this review, adherence is defined as the degree to which the user followed the program as it was designed. The two aims were (1) to review the methods used to assess adherence and (2) to evaluate the association between adherence measures and outcome.

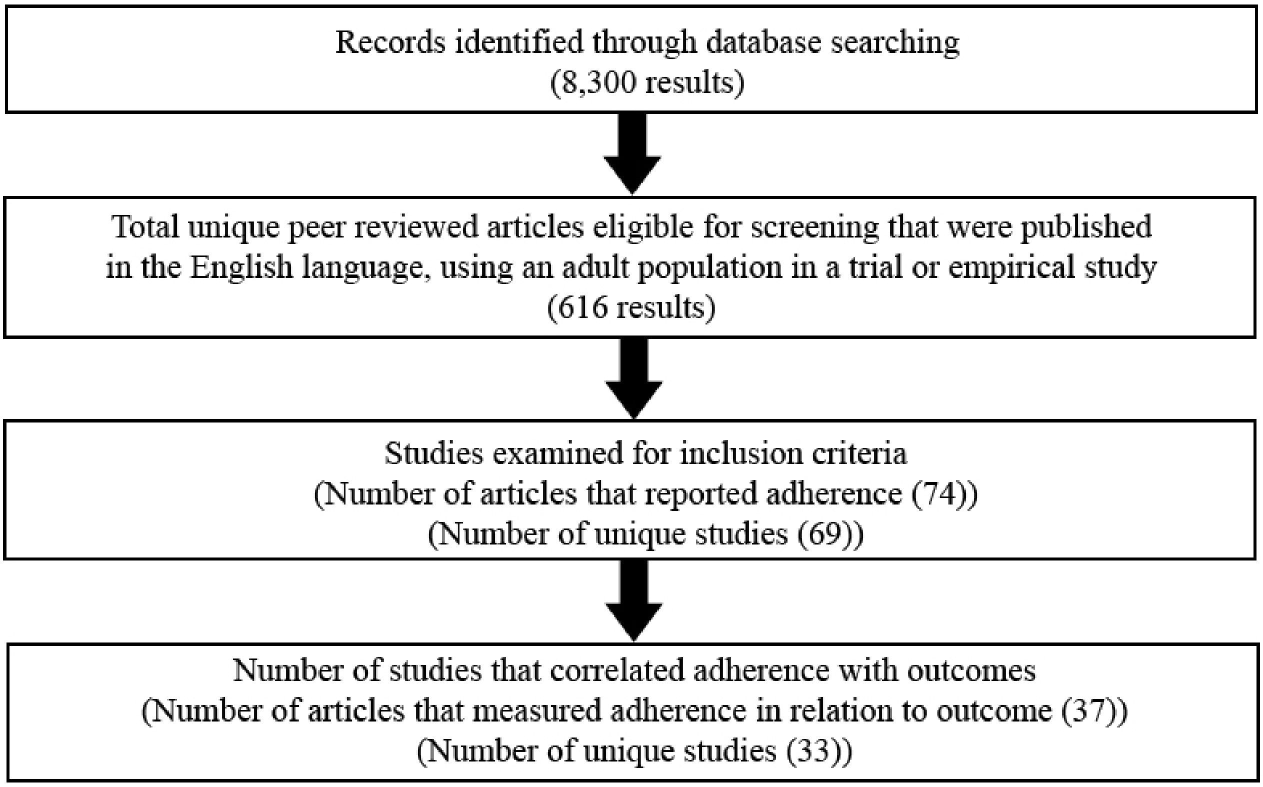
Methods

Systematic Selection Criteria and Search Strategy

A systematic database search was used to identify articles relevant to the aims of the review. Articles published up to and including April 2010, as indexed by Medline, PsycINFO, and the Cochrane Central Register of Controlled Trials databases, were included. Broad keyword search terms were used, favoring sensitivity over specificity, given the relative youth of the field.

Search terms used were the keywords “Internet” OR “web” combined with “therapy” OR “self-help” OR “intervention.” Following this initial search strategy, we identified 8300 articles (see Figure 1 for a flow diagram of article selection).

Figure 1. Flow diagram of the process of article selection for systematic review



The abstracts of the selected articles were reviewed and selected for inclusion by the first author using the following criteria.

Inclusion Criteria

An article was suitable for inclusion if it was published in a peer-reviewed journal and the subjects of the intervention were aged 18 years or older. The article also needed to consist of the evaluation of a “self-help,” user-directed online intervention where the Internet was the primary therapeutic delivery modality. In addition, the intervention also required the user to engage with the material on at least 2 occasions in a structured format (therefore, not a website with tools or information, but requiring a progression through the program). When reporting results, the study needed to report 1 or more adherence measures (defined as a measure of program usage by the participant) as well as the outcome of the intervention.

Exclusion Criteria

Articles that were not written in the English Language were excluded. Design characteristics that led to exclusion were the participant concurrently receiving a psychotherapeutic intervention in addition to the intervention being studied, or that the program involved synchronous communication as part of the program (eg, online chat, teleconferencing, or personalized therapeutic telephone contact from the therapist during the intervention). These studies were excluded to remove the potentially confounding nature of instantaneous feedback, extrinsic motivation, social desirability, and uncontrolled social contact. Studies with telephone contact were included if the calls were reported to be scripted, described as nontherapeutic in nature, and for the purpose of tracking participants.

Additionally, further studies were excluded because the individual completing the program was not the individual displaying the target behavior (eg, teachers of children with behavioral issues, or parents of children with enuresis), the program was not delivered over the Internet (eg, the program was delivered by CD-ROM via a computer), or the program was delivered at a specific location or time (eg, at a hospital clinic or at a specific time each week), therefore requiring the participant’s presence online at that time or place.

Coding of Study Characteristics

Key article characteristics were recorded using a data extraction template designed for this review. Key data captured for analysis consisted of variables that would allow the articles to be reviewed and factors thought to be important in adherence, based on a review of the literature.

These factors consisted of study sample size; the study design, including the nature of the control and intervention; the behavior or state that was the target of the intervention; and a description of the intervention, its therapeutic underpinnings, its intervention style, and its requirements from users.

In addition to the study characteristics, adherence data were captured. These consisted of a measure of the degree to which the individual engaged with the website as determined by program logins, engagement in online activities, time spent onsite, and the number of modules completed. Data were extracted for each adherence measure, including the type of measure and data reported relating to this. Any statistical analyses that examined the role of adherence in outcome were also recorded during data extraction.

The data collected on adherence variables were aggregated and evaluated for strength of association with outcome. Due to the methodological challenges associated with Internet trials (eg, high attrition rates), it was difficult to use a widely used rating system. Given these challenges, an alternative rating system was developed, informed by the Scottish Intercollegiate Guidelines Network rating system [31]. Based on this system, the strength of adherence measure rating was determined, using a 5-category scale based on the consistency of findings from this review (note that only 4 categories are reported here as no measures met the criteria for the 5th category). Details of this system can be found in the footnotes of the corresponding tables.

Results

We initially found a total of 8300 studies by using the above combination of search terms. Once duplicates were removed, the articles were further limited to those that were published in an English-language, peer-reviewed journal (n = 1095). The remaining articles were systematically reviewed (by LD) to ensure they met the inclusion criteria. This resulted in a total of 74 articles describing 69 individual studies. The heterogeneity of adherence and outcome measures precluded a formal meta-analysis.

Sample size varied significantly between the trials, ranging between 20 and 3176 (median 146), with many of the open-access trials having more participants (Table 1). Structured, discrete-period clinical trials tended to have fewer participants. This may be due to the seemingly more intensive researcher input in the form of feedback provided by emails, monitoring of activity completion, and moderation of activity, requiring greater resourcing for smaller numbers of participants. Larger trials tended to evaluate unstructured websites where participants were able to use the website in whatever manner they pleased, rather than using a structured preplanned program.

Mean study discontinuation rate was approximately 23% of all trial participants (range 0%–83%). Studies of physical and psychological target behaviors had similar attrition rates.

Table 1. Descriptive statistics for participant randomization sample size in studies that reported adherence included within systematic review (n = 69)

Descriptive statistic	Total sample size (n)	Physical health sample size (n)	Psychological health sample size (n)
Median	146	190	103
Minimum	20	62	20
Maximum	3176	2523	3176
First quartile	77	91	56
Third quartile	400	958	272
Total number of studies	69	29	40
Total number of participants	34,465	19,147	15,318

Methods Used to Measure Adherence

Adherence data that were captured varied across studies. This included reporting the number of times the participant accessed or logged into the program, completed modules or activities, visits made to forums, posts made to the forum, and pages viewed and printed, as well as self-reported completion of activities away from the program or offline. Despite the commonality and functionality of being able to capture or measure participant logins across trials, only half (33/69) of the studies presented these adherence data in relation to outcome measures. See table 2 for a breakdown of ways in which adherence was measured in the included studies.

Table 2. Methods for measuring adherence to e-therapy as reported by included studies (n = 69)

Measure of adherence	Number of times reported
Logins to program	36
Module completion	31
Time spent online	18
Completion of a predefined activity such or use of an online tool	16
Posts made	9
Pages viewed	5
Replies to emails	6
Forum visits	1
Use of online tools	1
Self-reported completion of offline activities	1
Print requests made	1

Of the 69 studies that reported measuring adherence, approximately half did so by measuring logins and/or completion of modules. Only a quarter of the 69 included studies reported 1 or more of the other potential measures of adherence. The reporting of module completion was more common in studies where the target behavior was psychological health or well-being (25/40, 63%) rather than physical health (6/29, 21%) (n = 69, $\chi^2_1 = 11.9$, $P < .001$). Conversely, login reporting was more common in studies where the target behavior was related to physical health (23/29, 79%) rather than psychological health (13/40, 33%) (n = 69, $\chi^2_1 = 14.8$, $P \leq .001$).

Effect of Adherence on Outcomes

Of the 69 studies included in the review, 33 (48%) analyzed the impact of 1 or more measure of adherence on outcome variables. Complete results of the review can be found in Multimedia Appendix 1.

Logins

A total of 9 studies correlated logins as a measure of adherence to outcomes. Using the previously described rating system, logins were found to be positively associated with outcome in studies targeting food and vegetable consumption [32], physical activity [33-37], and weight management [38-40]. Number of logins was not found to be related to outcomes in studies targeting depression [23,41].

Activities Completed

In 6 studies, activities completed were correlated with outcomes. The number of self-reported activities completed (eg, completing a diary, engaging in online tests, or making forum posts) was found to correlate with outcomes in interventions that targeted cigarette smoking [42], weight management [39,40], and body dissatisfaction [43,44]. In studies targeting physical activity [33,34] or depression [45], the number of activities completed was not found to correlate with outcomes.

Modules Completed

Completed modules were the most commonly reported measure of adherence, and they correlated with outcomes in 16 studies. The number of modules completed was found to correlate with outcomes in interventions that targeted cigarette smoking [42,46,47], depression [45,48-51], and anxiety disorders [52-57].

Time Spent Online

The relationship between time spent online and outcomes was correlated in 4 studies. Time spent online was not correlated to outcomes in studies that targeted depression [58] and anxiety [59], but was correlated to outcomes in studies of infertility related distress [60] and body dissatisfaction [61].

Pages Opened

The number of pages opened was explored in 3 studies. The number of pages opened was negatively correlated to outcome in 1 study targeting depression [58] but positively related to outcomes in 2 studies targeting body dissatisfaction [62,63].

Website Exposure

The relationship between website exposure and outcomes was explored in 3 studies. Website exposure and program usage was positively correlated with outcomes in interventions targeting smokeless tobacco [64], depression [65], and physical activity [66].

Table 3 summarizes adherence measures by intervention target, reported by target behavior. (For a more in-depth table, please see Multimedia Appendix 1.)

Table 3. Summary of the strength of aggregated adherence measure data by target behavior[a]

Target behavior	Strength of adherence–outcome association (number of studies)					
	Logins	Activities completed	Modules completed	Time spent online	Pages opened	Website exposure
Physical health						
Fruit and vegetable consumption	+[b] (n = 1)					
Physical activity	+[b] (n = 3)	0[c] (n = 1)				+[b] (n = 1)
Weight management	++[d] (n = 2)	++ [d] (n = 3)				
Smoking		+[b] (n = 1)	+[b] (n = 4)			
Smokeless tobacco						+[b] (n = 1)
Psychological health						
Depression	0[c] (n = 2)	0[c] (n = 1)	+[b] (n = 7)	–[e] (n = 1)	–[e] (n = 1)	+[b] (n = 1)
Anxiety			+[d] (n = 6)	0[c] (n = 1)		
Body dissatisfaction		+[b] (n = 2)		+[b] (n = 1)	+[b] (n = 2)	
Fertility-related distress				+[b] (n = 1)		

[a] The number of studies that were aggregated to form the strength rating in the review is indicated in parentheses following the rating indicator (+ = positive; – = negative; 0 = null). (For a complete breakdown of the studies that contributed to the aggregate results see Multimedia Appendix 1.)

[b] 1 study or mixed evidence with predominantly positive relationships found between adherence measures and outcome.

[c] No relationship found between adherence and outcome measures.

[d] At least 2 studies finding a positive correlation between increased adherence and outcome measures.

[e] 1 study or mixed evidence with predominantly negative relationships found between adherence measures and outcome.

Discussion

Understanding adherence to e-therapies is important in understanding how these therapies may benefit individuals who need intervention. The impact of adherence on outcome appears to vary. The review demonstrates that however it is measured, adherence is associated positively when reported with intervention outcomes targeting physical health. However, in the most commonly targeted outcome of e-therapy, depression, the number of logins, self-reported activities, time online, and pages opened showed no such association. Only 2 measures of adherence—the degree of completion of the modules within the program and a summative "website exposure" outcome—were associated with better depression outcomes.

The association between the number of logins and outcomes in e-therapies targeting physical health is similar to the positive relationship observed between session attendance and outcome in physical rehabilitation [67-70] and to the dose–response curve seen in medication therapy. This association may largely be due to the number of logins being representative of the participant's willingness to use the program through their return to the website. The number of logins may be more indicative of program usage than are self-reported activities completed or forums posted. This is particularly the case in programs where there is structured program use (therefore needing completion of an activity or module prior to moving to another) but no time restrictions placed on progression. In this style of program, a participant may complete several modules or activities during a single login. Therefore, the benefits received in completing a module, processing it over a time period, and using its skills before building on this with the next module may be lost.

For participants involved in trials targeting weight changes as an outcome [38-40], all measures of adherence were correlated with outcomes. Therefore, more adherent participants had a higher level of weight loss. This is consistent with the current literature, which shows that the more closely people follow dietary plans, the better their outcomes [71]. In the case of physical outcomes, one interpretation of the results could be that adherence is a marker of an unmeasured factor such as personality or self-efficacy that predisposes participants to a better outcome. A more extreme interpretations is that the content of the intervention is irrelevant; it is the application of the participant to "something" that produces the outcomes.

The most common use of e-therapy is to intervene in depression and anxiety. However, measures of logins, self-reported activity completions, and time online were not associated with outcome—only evidence of actual module completion either on its own or in the composite "program exposure" measure appeared to be related to outcome. This suggests a more nuanced interpretation, in that a participant's interaction with the module content leads to change, rather than improvement merely reflecting a greater propensity to adhere. This reflects the face-to-face literature, in that the structured psychotherapies, with module (task) completion and review, are generally more effective than supportive psychotherapy [72-74]. The lack of association of many adherence measures with outcome might be explained by lack of power in these studies. This is particularly supported by more consistent findings in physical health interventions, which had a median sample size nearly twice that of psychological interventions. However, it is of note that many psychological health studies did have large sample sizes and, for some adherence measures, the results were not null, but negative.

The definition of adherence and its consequent analysis also varied considerably across studies. Several studies categorized participants into adherent or not [49,54,57,75], or usage categories such as low, medium, and high [60], while others used continuous variables [23,24,33-36,38-41,43-45,50-52,55,56,58,59,61,76] or a combination of these [42,46,53,64-66,77]. Such variation is likely to produce mixed findings. It is also worth noting that several studies reported collecting multiple measures of adherence but reported on the outcome of only a few selected items. In many, it is unclear why some measures of adherence were reported over others, though this may presumably reflect a reporting bias and may therefore inaccurately portray the role of adherence. It is possible that unreported relationships may be nonsignificant, but may also have been useful in understanding the important components of program engagement and may build understanding of the user-engagement factors of e-therapy.

The variation in measurement of adherence in online interventions makes it difficult to accurately determine the impact of adherence on outcome. Despite having several seemingly objective measures such as number of logins, time spent online, and activities completed, there are still difficulties in determining the “dose” a participant receives. Specifically, time spent online is likely to be influenced by factors such as processing speed, cognitive ability, reading aptitude, and familiarly with using computers, several of which are likely to be influenced by psychological health. Likewise, a person who is unfamiliar with using a computer may write out their activities, rather than engage in the functionality presented in the online environment. This needs to be explored further by studies that randomly assign participants to receive different doses of programs, such as that by Christensen et al [25], which explored exposure to modules against change in outcome measures. Future interventions that intend to measure the impact of adherence on outcomes would benefit from clearly defining their adherence variables and exploring all relationships between the potential adherence variables and outcomes. In addition to this, measures of inactivity need to be included where temporal measures are used. Therefore, when a measure of adherence is the amount of time spent online, there needs to be a time-out function of need to engage readily with the program, to show that it is being used. This will allow the exclusion of individuals who have left their computer on and who are no longer engaging with the program material.

From this review, the mechanism of how adherence influences Internet treatment outcomes is unclear. Medical studies have shown that adherence is associated with better outcomes, regardless of whether patients received placebo or active interventions [78,79]. Within this review, it is unclear whether adherent patients generally do better, regardless of condition. It has been suggested that the mechanisms underlying a generalized adherence effect may be similar to the mechanisms that underlie the placebo effect [80], whereby the positive outcomes achieved from taking a placebo are the result of internalized beliefs about therapy, such as the expectations that the individual holds or the belief that the therapy will be effective. The placebo effect may also indicate a phenomenon of regression to the mean, in that outlying behavior returns closer to the mean over time. Alternatively, it has been suggested that adherence may be a more general indicator of orientation to healthy behaviors. Therefore, individuals who are adherent may be more likely to follow healthy lifestyle practices and therefore have improved outcomes [81,82]. Further exploration within the area of e-therapy needs to determine whether adherence influences outcomes through expectation beliefs or through participants being generally adherent to treatment recommendations. To do this, control group adherence behaviors need to be reported more.

Limitations

We recognize that several key studies that were included in the original sample of 8300 were excluded as a result of selection criteria. Specifically, some trials that potentially would have met the review criteria were excluded before analysis as a result of database coding and indexing. While Cochrane-style systematic reviews have used rigorous data selection and extraction templates to select trials, the varied nature of the design and high attrition rate of e-therapies made it difficult to use such templates. Given this, the limits within the databases were used to ensure that the methodology of the review was as systematic and replicable as possible.

As the studies contained within this review were heterogeneous in terms of the populations studied, nature of the interventions used, lengths of follow-up, and outcome measures, we summarized the findings by adherence measure only, rather than attempting a meta-analysis. This has limited the ability to combine and report the data beyond the format used. Similarly, the large number of small studies included in the review may have been underpowered for finding significant results.

Implications for Future Work

There are many difficulties in determining the role of adherence on outcomes in e-therapies. While objective data can be captured with relative ease, this may not truly reflect the user’s experience and dose. This is particularly pertinent when the user is required to complete activities such as homework exercises between online sessions, or when the user can print off material to review. In both of these scenarios, “objective measures” may underestimate the usage of the program. Similarly, user aspects such as processing speed, and familiarity with Web-based platforms and user interfaces are likely to influence the time spent online. Therefore, understanding the contributing factors of adherence is likely to be as important as understanding adherence per se.

From the findings of this review, it appears that the number of logins is the measure of adherence best correlated with physical outcomes, while module completion correlated with psychological outcomes. This suggests that program persistence and adherence may be important for physical and psychological interventions in different ways. However, these results need to be considered with caution, given the limitations of this review and the potential biases in the data. We therefore recommend that this be explored in future studies where adherence is the focus. It may mean that an aggregated adherence measure such as program exposure could provide a more meaningful measure of adherence by incorporating logins, time spent online, and number of activities and modules completed. However, it is not always clear how such a composite measure was derived in the studies and how each factor contributes to such a score. Similarly, variations existed within composite measure across each trial, making the comparison difficult. Furthermore, the design of interventions could take this into account by maximizing the likelihood of a participant undertaking the behavior most likely to enhance outcomes. For instance, maximizing module completion for depression and anxiety interventions, even at the expense of number of logins, may be a good trade-off. Understanding the differential effects of different measures of adherence will be important in future content and platform development, as well as in evaluating applicability and health service issues.

Acknowledgments

Liesje Donkin receives a scholarship from the University of Sydney to complete her PhD with additional support from the Cardiovascular Disease and Depression Strategic Research Program (Award Reference No. G08S 4048) and beyondblue. Thanks also to Sun Bin who retrieved articles to be included in this review.

Conflicts of Interest

None declared

Multimedia Appendix 1

Association of each type of adherence measure with outcome grouped by target of intervention

[PDF file (Adobe PDF File), 274 KB - jmir_v13i3e52_app1.pdf]

References

1. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009 Dec;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
2. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. Clin Psychol Rev 2009 Jul;29(5):410-420. [doi: 10.1016/j.cpr.2009.04.001] [Medline: 19403215]
3. Barak A, Hen L, Boniel-Nissim M, Shapira NA. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. J Technol Hum Serv 2008;26(2-4):109-160. [doi: 10.1080/15228830802094429]
4. Eysenbach G. Issues in evaluating health websites in an Internet-based randomized controlled trial. J Med Internet Res 2002 Dec;4(3):E17 [FREE Full text] [doi: 10.2196/jmir.4.3.e17] [Medline: 12554548]
5. Sabate E. Adherence to Long-Term Therapies: Evidence for Action. Geneva: World Health Organization; 2003.
6. Cramer JA, Roy A, Burrell A, Fairchild CJ, Fuldeore MJ, Ollendorf DA, et al. Medication compliance and persistence: terminology and definitions. Value Health 2008;11(1):44-47. [doi: 10.1111/j.1524-4733.2007.00213.x] [Medline: 18237359]
7. Coon DW, Thompson LW. The relationship between homework compliance and treatment outcomes among older adult outpatients with mild-to-moderate depression. Am J Geriatr Psychiatry 2003;11(1):53-61. [Medline: 12527540]
8. Neimeyer RA, Kazantzis N, Kassler DM, Baker KD, Fletcher R. Group cognitive behavioural therapy for depression outcomes predicted by willingness to engage in homework, compliance with homework, and cognitive restructuring skill acquisition. Cogn Behav Ther 2008;37(4):199-215. [doi: 10.1080/16506070801981240] [Medline: 18608311]
9. Thase ME, Callan JA. The role of homework in cognitive behavior therapy of depression. J Psychother Integr 2006;16(2):162-177. [doi: 10.1037/1053-0479.16.2.162]
10. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol 2005 Feb;34(1):79-87 [FREE Full text] [doi: 10.1093/ije/dyh300] [Medline: 15561753]
11. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet 2002 Mar 2;359(9308):781-785. [doi: 10.1016/S0140-6736(02)07882-0] [Medline: 11888606]
12. Batterham PJ, Neil AL, Bennett K, Griffiths KM, Christensen H. Predictors of adherence among community users of a cognitive behavior therapy website. Patient Prefer Adherence 2008;2:97-105 [FREE Full text] [Medline: 19920949]
13. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

14. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
15. Wierzbicki M, Pekarik G. A meta-analysis of psychotherapy dropout. Prof Psychol Res Pr 1993;24(2):190-195. [doi: 10.1037/0735-7028.24.2.190]
16. Strauss JL, Guerra VS, Marx CE, Meade Eggleston A, Calhoun PS. Adherence and psychotherapy. In: Bosworth HB, Oddone EZ, Weinberger M, editors. Improving Patient Treatment Adherence: Concepts, Interventions, and Measurement. New York, NY: Springer Verlag; 2010:215-240.
17. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
18. Bewick BM, Trusler K, Mulhern B, Barkham M, Hill AJ. The feasibility and effectiveness of a web-based personalised feedback and social norms alcohol intervention in UK university students: a randomised control trial. Addict Behav 2008 Sep;33(9):1192-1198. [doi: 10.1016/j.addbeh.2008.05.002] [Medline: 18554819]
19. Bosak KA, Yates B, Pozehl B. Feasibility of an internet physical activity intervention. West J Nurs Res 2009 Aug;31(5):648-661. [doi: 10.1177/0193945909334433] [Medline: 19420280]
20. Goessens BM, Visseren FL, de Nooijer J, van den Borne HW, Algra A, Wierdsma J, SMART Study Group. A pilot-study to identify the feasibility of an Internet-based coaching programme for changing the vascular risk profile of high-risk patients. Patient Educ Couns 2008 Oct;73(1):67-72. [doi: 10.1016/j.pec.2008.06.004] [Medline: 18639410]
21. Harvey-Berino J, Pintauro SJ, Gold EC. The feasibility of using Internet support for the maintenance of weight loss. Behav Modif 2002 Jan;26(1):103-116. [Medline: 11799651]
22. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
23. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
24. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
25. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
26. Advocat J, Lindsay J. Internet-based trials and the creation of health consumers. Soc Sci Med 2010 Feb;70(3):485-492. [doi: 10.1016/j.socscimed.2009.10.051] [Medline: 19926185]
27. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
28. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008 Nov;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]
29. Manwaring JL, Bryson SW, Goldschmidt AB, Winzelberg AJ, Luce KH, Cunning D, et al. Do adherence variables predict outcome in an online program for the prevention of eating disorders? J Consult Clin Psychol 2008 Apr;76(2):341-346. [doi: 10.1037/0022-006X.76.2.341] [Medline: 18377129]
30. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
31. Scottish Intercollegiate Guidelines Network. SIGN 50: A Guideline Developers Handbook. Edinburgh: NHS Quality Improvement Scotland (NHS QIS); 2008.
32. Alexander GL, McClure JB, Calvi JH, Divine GW, Stopponi MA, Rolnick SJ, MENU Choices Team. A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. Am J Public Health 2010 Feb;100(2):319-326. [doi: 10.2105/AJPH.2008.154468] [Medline: 20019315]
33. Carr LJ, Bartee RT, Dorozynski C, Broomfield JF, Smith ML, Smith DT. Internet-delivered behavior change program increases physical activity and improves cardiometabolic disease risk factors in sedentary adults: results of a randomized controlled trial. Prev Med 2008 May;46(5):431-438. [doi: 10.1016/j.ypmed.2007.12.005] [Medline: 18207228]
34. Carr LJ, Bartee RT, Dorozynski CM, Broomfield JF, Smith ML, Smith DT. Eight-month follow-up of physical activity and central adiposity: results from an Internet-delivered randomized control trial intervention. J Phys Act Health 2009 Jul;6(4):444-455. [Medline: 19842458]
35. Lewis B, Williams D, Dunsiger S, Sciamanna C, Whiteley J, Napolitano M, et al. User attitudes towards physical activity websites in a randomized controlled trial. Prev Med 2008 Nov;47(5):508-513. [doi: 10.1016/j.ypmed.2008.07.020] [Medline: 18773915]

36. Marcus BH, Lewis BA, Williams DM, Dunsiger S, Jakicic JM, Whiteley JA, et al. A comparison of Internet and print-based physical activity interventions. Arch Intern Med 2007 May 14;167(9):944-949 [FREE Full text] [doi: 10.1001/archinte.167.9.944] [Medline: 17502536]
37. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008;10(5):e40 [FREE Full text] [doi: 10.2196/jmir.993] [Medline: 19017582]
38. 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] [doi: 10.1001/jama.289.14.1833] [Medline: 12684363]
39. 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]
40. Webber KH, Tate DF, Michael Bowling J. A randomized comparison of two motivationally enhanced Internet behavioral weight loss programs. Behav Res Ther 2008 Sep;46(9):1090-1095. [doi: 10.1016/j.brat.2008.06.008] [Medline: 18675402]
41. Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, et al. Overcoming depression on the Internet (ODIN): a randomized controlled trial of an Internet depression skills intervention program. J Med Internet Res 2002 Dec;4(3):e14 [FREE Full text] [doi: 10.2196/jmir.4.3.e14] [Medline: 12554545]
42. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine Tob Res 2009 Sep;11(9):1025-1034. [doi: 10.1093/ntr/ntp090] [Medline: 19640833]
43. 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]
44. 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]
45. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. Br J Psychiatry 2005 Nov;187:456-461 [FREE Full text] [doi: 10.1192/bjp.187.5.456] [Medline: 16260822]
46. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
47. Seidman DF, Westmaas JL, Goldband S, Rabius V, Katkin ES, Pike KJ, et al. Randomized controlled trial of an interactive internet smoking cessation program with long-term follow-up. Ann Behav Med 2010 Feb;39(1):48-60. [doi: 10.1007/s12160-010-9167-7] [Medline: 20177844]
48. Spek V, Nyklícek I, Smits N, Cuijpers P, Riper H, Keyzer J, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 2007 Dec;37(12):1797-1806. [doi: 10.1017/S0033291707000542] [Medline: 17466110]
49. de Graaf LE, Gerhards SA, Arntz A, Riper H, Metsemakers JF, Evers SM, et al. Clinical effectiveness of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. Br J Psychiatry 2009 Jul;195(1):73-80 [FREE Full text] [doi: 10.1192/bjp.bp.108.054429] [Medline: 19567900]
50. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
51. Spek V, Cuijpers P, Nyklícek I, Smits N, Riper H, Keyzer J, et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. Psychol Med 2008 May;38(5):635-639. [doi: 10.1017/S0033291707002590] [Medline: 18205965]
52. Tillfors M, Carlbring P, Furmark T, Lewenhaupt S, Spak M, Eriksson A, et al. Treating university students with social phobia and public speaking fears: Internet delivered self-help with or without live group exposure sessions. Depress Anxiety 2008;25(8):708-717. [doi: 10.1002/da.20416] [Medline: 18729147]
53. Titov N, Andrews G, Choi I, Schwencke G, Mahoney A. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. Aust N Z J Psychiatry 2008 Dec;42(12):1030-1040. [doi: 10.1080/00048670802512107] [Medline: 19016091]
54. Titov N, Andrews G, Schwencke G. Shyness 2: treating social phobia online: replication and extension. Aust N Z J Psychiatry 2008 Jul;42(7):595-605. [doi: 10.1080/00048670802119820] [Medline: 18612863]
55. Carlbring P, Nilsson-Ihrfelt E, Waara J, Kollenstam C, Buhrman M, Kaldo V, et al. Treatment of panic disorder: live therapy vs. self-help via the Internet. Behav Res Ther 2005 Oct;43(10):1321-1333. [doi: 10.1016/j.brat.2004.10.002] [Medline: 16086983]
56. Furmark T, Carlbring P, Hedman E, Sonnenstein A, Clevberger P, Bohman B, et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. Br J Psychiatry 2009 Nov;195(5):440-447 [FREE Full text] [doi: 10.1192/bjp.bp.108.060996] [Medline: 19880935]

57. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
58. Clarke G, Kelleher C, Hornbrook M, Debar L, Dickerson J, Gullion C. Randomized effectiveness trial of an Internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. Cogn Behav Ther 2009 Dec;38(4):222-234. [doi: 10.1080/16506070802675353] [Medline: 19440896]
59. Kenardy J, McCafferty K, Rosa V. Internet-delivered indicated prevention for anxiety disorders: a randomized controlled trial. Behav Cogn Psychother 2003;31(3):279-289. [doi: 10.1017/S1352465803003047]
60. Cousineau TM, Green TC, Corsini E, Seibring A, Showstack MT, Applegarth L, et al. Online psychoeducational support for infertile women: a randomized controlled trial. Hum Reprod 2008 Mar;23(3):554-566 [FREE Full text] [doi: 10.1093/humrep/dem306] [Medline: 18089552]
61. Low KG, Charanasomboon S, Lesser J, Reinhalter K, Martin R, Jones H, et al. Effectiveness of a computer-based interactive eating disorders prevention program at long-term follow-up. Eat Disord 2006;14(1):17-30. [doi: 10.1080/10640260500403816] [Medline: 16757446]
62. Taylor CB, Bryson S, Luce KH, Cunning D, Doyle AC, Abascal LB, et al. Prevention of eating disorders in at-risk college-age women. Arch Gen Psychiatry 2006 Aug;63(8):881-888 [FREE Full text] [doi: 10.1001/archpsyc.63.8.881] [Medline: 16894064]
63. 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. [doi: 10.1037/0022-006X.72.5.914] [Medline: 15482051]
64. Severson HH, Gordon JS, Danaher BG, Akers L. ChewFree.com: evaluation of a Web-based cessation program for smokeless tobacco users. Nicotine Tob Res 2008 Feb;10(2):381-391. [doi: 10.1080/14622200701824984] [Medline: 18236303]
65. de Graaf LE, Huibers MJ, Riper H, Gerhards SA, Arntz A. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. J Affect Disord 2009 Aug;116(3):227-231. [doi: 10.1016/j.jad.2008.12.009] [Medline: 19167094]
66. 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]
67. Nathan PA, Keniston RC, Meadows KD. Outcome study of ulnar nerve compression at the elbow treated with simple decompression and an early programme of physical therapy. J Hand Surg Br 1995 Oct;20(5):628-637. [Medline: 8543870]
68. Rives K, Gelberman R, Smith B, Carney K. Severe contractures of the proximal interphalangeal joint in Dupuytren's disease: results of a prospective trial of operative correction and dynamic extension splinting. J Hand Surg Am 1992 Nov;17(6):1153-1159. [Medline: 1430959]
69. Lyngcoln A, Taylor N, Pizzari T, Baskus K. The relationship between adherence to hand therapy and short-term outcome after distal radius fracture. J Hand Ther 2005;18(1):2-8. [doi: 10.1197/j.jht.2004.10.008] [Medline: 15674780]
70. Groth GN, Wilder DM, Young VL. The impact of compliance on the rehabilitation of patients with mallet finger injuries. J Hand Ther 1994;7(1):21-24. [Medline: 8012481]
71. Sacks FM, Bray GA, Carey VJ, Smith SR, Ryan DH, Anton SD, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. N Engl J Med 2009 Feb 26;360(9):859-873 [FREE Full text] [doi: 10.1056/NEJMoa0804748] [Medline: 19246357]
72. Cottraux J, Note I, Albuisson E, Yao SN, Note B, Mollard E, et al. Cognitive behavior therapy versus supportive therapy in social phobia: a randomized controlled trial. Psychother Psychosom 2000;69(3):137-146. [Medline: 10773778]
73. Tolin DF. Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. Clin Psychol Rev 2010 Aug;30(6):710-720. [doi: 10.1016/j.cpr.2010.05.003] [Medline: 20547435]
74. Rowa K, Antony MM. Psychological treatments for social phobia. Can J Psychiatry 2005 May;50(6):308-316. [Medline: 15999944]
75. Warmerdam L, van Straten A, Twisk J, Riper H, Cuijpers P. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 2008;10(4):e44 [FREE Full text] [doi: 10.2196/jmir.1094] [Medline: 19033149]
76. Taylor Y, Eliasson A, Andrada T, Kristo D, Howard R. The role of telemedicine in CPAP compliance for patients with obstructive sleep apnea syndrome. Sleep Breath 2006 Sep;10(3):132-138. [doi: 10.1007/s11325-006-0059-9] [Medline: 16565867]
77. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. Web-based smoking-cessation programs: results of a randomized trial. Am J Prev Med 2008 May;34(5):373-381. [doi: 10.1016/j.amepre.2007.12.024] [Medline: 18407003]
78. Avins AL, Pressman A, Ackerson L, Rudd P, Neuhaus J, Vittinghoff E. Placebo adherence and its association with morbidity and mortality in the studies of left ventricular dysfunction. J Gen Intern Med 2010 Dec;25(12):1275-1281. [doi: 10.1007/s11606-010-1477-8] [Medline: 20706875]

79. Granger AL, Fehnel SE, Hogue SL, Bennett L, Edin HM. An assessment of patient preference and adherence to treatment with Wellbutrin SR: a web-based survey. J Affect Disord 2006 Feb;90(2-3):217-221. [doi: 10.1016/j.jad.2005.08.018] [Medline: 16360216]
80. Czajkowski SM, Chesney MA, Smith AW. Adherence and the placebo effect. In: Shumaker SA, Ockene JK, Riekert KA, editors. The Handbook of Health Behavior Change. New York, NY: Springer; 2009:713-734.
81. Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. N Engl J Med 1980 Oct 30;303(18):1038-1041. [doi: 10.1056/NEJM198010303031804] [Medline: 6999345]
82. Horwitz RI, Viscoli CM, Berkman L, Donaldson RM, Horwitz SM, Murray CJ, et al. Treatment adherence and risk of death after a myocardial infarction. Lancet 1990 Sep 1;336(8714):542-545. [Medline: 1975045]

Edited by G Eysenbach; submitted 14.02.11; peer-reviewed by B Cugelman, W Brouwer; comments to author 10.03.11; revised version received 24.05.11; accepted 05.05.11; published 05.08.11.

Please cite as:
Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N
A Systematic Review of the Impact of Adherence on the Effectiveness of e-Therapies
J Med Internet Res 2011;13(3):e52
URL: http://www.jmir.org/2011/3/e52/
doi:10.2196/jmir.1772
PMID:21821503

Original Paper

Modeling the Cost-Effectiveness of Health Care Systems for Alcohol Use Disorders: How Implementation of eHealth Interventions Improves Cost-Effectiveness

Filip Smit[1,2], PhD; Joran Lokkerbol[3], MSc; Heleen Riper[4,5], PhD; Maria Cristina Majo[1,6], PhD; Brigitte Boon[7], PhD; Matthijs Blankers[8], MSc

[1]Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Centre of Prevention and Brief Intervention, Utrecht, Netherlands

[2]VU University Medical Centre, Department of Epidemiology and Biostatistics, EMGO+ Institute for Health and Health Care Research, Amsterdam, Netherlands

[3]University of Amsterdam, Department of Business and Industrial Statistics, Amsterdam, Netherlands

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

[5]GGZinGeest, Regional Mental Health Service Centre, VU University Medical Centre, Amsterdam, Netherlands

[6]Tor Vergata University, Department of Financial Economics and Quantitative Methods, Rome, Italy

[7]Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Centre of Brief Intervention and Prevention, Utrecht, Netherlands

[8]Amsterdam Institute for Addiction Research, Academic Medical Centre, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Filip Smit, PhD
Trimbos Institute (Netherlands Institute of Mental Health and Addiction)
Centre of Prevention and Brief Intervention
P.O. Box 725
Utrecht, 1400 AS
Netherlands
Phone: 31 629735268
Fax: 31 302971111
Email: fsmit@trimbos.nl

Abstract

Background: Informing policy decisions about the cost-effectiveness of health care systems (ie, packages of clinical interventions) is probably best done using a modeling approach. To this end, an alcohol model (ALCMOD) was developed.

Objective: The aim of ALCMOD is to estimate the cost-effectiveness of competing health care systems in curbing alcohol use at the national level. This is illustrated for scenarios where new eHealth technologies for alcohol use disorders are introduced in the Dutch health care system.

Method: ALCMOD assesses short-term (12-month) incremental cost-effectiveness in terms of reductions in disease burden, that is, disability adjusted life years (DALYs) and health care budget impacts.

Results: Introduction of new eHealth technologies would substantially increase the cost-effectiveness of the Dutch health care system for alcohol use disorders: every euro spent under the current system returns a value of about the same size (€ 1.08, ie, a "surplus" of 8 euro cents) while the new health care system offers much better returns on investment, that is, every euro spent generates € 1.62 in health-related value.

Conclusion: Based on the best available evidence, ALCMOD's computations suggest that implementation of new eHealth technologies would make the Dutch health care system more cost-effective. This type of information may help (1) to identify opportunities for system innovation, (2) to set agendas for further research, and (3) to inform policy decisions about resource allocation.

(J Med Internet Res 2011;13(3):e56) doi:10.2196/jmir.1694

KEYWORDS

Alcohol-related disorders; early intervention; health care systems; cost-effectiveness

Introduction

Alcohol use disorders are a leading cause of disease burden [1,2] and are associated with substantial economic costs [3-5]. Therefore, curbing alcohol use has long been recognized as an important public health objective [6,7]. Health care systems play a crucial role in achieving this objective, but most health care systems offer room for improvement in terms of greater efficiency. This begs the question what type of health care system (ie, what mix of interventions) is optimal. We could provisionally define an optimal health care system in terms of meeting the following criteria: the health care system needs to be acceptable to its recipients, scalable to absorb increasing demands for health care, effective to generate the required health gains, and economically affordable to become sustainable over time. Public health planners need ways to design health care systems that optimize these criteria, compare the relative advantage of newly designed systems with the current one, and choose the most cost-effective system. This is a daunting but important task.

However, this task might be facilitated with a simulation model, which can compare a “base-case” scenario (eg, the current mix of clinical interventions) with an alternative (hypothetical) scenario consisting of new interventions or a different mix of interventions. In order for it to be helpful, the model should be able to evaluate the relative advantage of one system over another in terms of incremental cost-effectiveness and be used as an aid to decision-making.

With these aims in mind, we developed an alcohol model (ALCMOD) that can address the above issues. Developing ALCMOD was conducted within the framework of the World Health Organization's International Action Plan on Implementing eHealth Technologies for Substance Abuse. In this context, we wanted to shed light on the population-level cost-effectiveness of health care systems for alcohol use disorders before and after the introduction of new eHealth technologies in Belarus, Brazil, India, Mexico, and the Netherlands. ALCMOD is programmed in Microsoft Excel 2007, because Excel is available on most computers.

The purpose of this paper is to describe ALCMOD's input and output and to take an in-depth look at the model's throughput: its computational strategies, the underlying assumptions, and its limitations. One such limitation is ALCMOD's focus on short-term impacts. Restricting the time horizon to 1 year was a conscious choice because there are several alcohol use disorders (heavy, hazardous, and harmful use and alcohol dependence; see Textbox 1 for definitions) and a lack of empirical data that help to quantify the longer-term treatment effects and relapse rates for each of the alcohol use disorders. By contrast, simulation of short-term health and budget impacts is straightforward and requires fewer assumptions. Strengths of ALCMOD include its ability to evaluate combinations of interventions, its adaptability to different populations and settings, its capacity to handle uncertainty in input parameters, and the way it incorporates coverage and adherence rates for each of the modeled interventions.

We illustrate ALCMOD's computations for the base-case of usual care in the Netherlands versus an alternative scenario consisting of usual care augmented with three eHealth interventions: the DrinkTest, DrinkingLess, and an online therapist-led treatment for problem drinking, termed OnlineTreatment henceforth. The DrinkTest is a brief online intervention consisting of screening one's alcohol use followed by automated personalized advice. DrinkingLess is an online four-step cognitive behavioral intervention. The steps in DrinkingLess are: (1) exploring one's alcohol use, (2) goal setting, (3) behavioral change, and (4) maintenance. Both the DrinkTest and DrinkingLess have been evaluated in randomized trials and meta-analytically and were found to be effective in curbing alcohol use [9-12]. Both the DrinkTest and DrinkingLess are pure self-help interventions, but OnlineTreatment is a therapist-led intervention. Communication between participant and therapist is conducted over the Internet in seven synchronous written chat sessions of 45 minutes each. The sessions are thematically structured and cover themes such as goal setting, self-control techniques, monitoring, recognizing situations that incur a risk of relapse, and relapse prevention techniques.

Textbox 1. Alcohol use disorders

Alcohol use disorders from the lexicon of alcohol and drug terms published by the World Health Organization [8]

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

OnlineTreatment has been evaluated in a randomized trial [13]. Preliminary results (not yet published) indicate that OnlineTreatment is effective and cost-effective. It is worth noting that the three eHealth interventions increase in intensity and could be used in a stepped-care framework, thus starting with the least intensive intervention, the DrinkTest, and moving up to the more intensive levels of DrinkingLess and OnlineTreatment, if so required.

The emergence of evidence-based eHealth technologies offers opportunities for innovation in existing health care systems. The new technologies may help to reach population segments that were hitherto not reached because they live in hard to reach rural areas or because they may have shied away from face-to-face delivered health services out of fear of stigma. The new technologies are also very scalable, thus allowing people to access health care services in an unprecedented way. In addition, the new eHealth technologies could be cost-effective, especially when offered as well-structured self-help interventions or as interventions with (minimal) therapist support. Considering the global health gap with regard to the alcohol use disorders [6], these developments could become quite important. However, to date there is only limited evidence for the cost-effectiveness of eHealth interventions [14]. For these reasons, it is opportune to conduct a population-level health-economic evaluation of the possible health gains and budget impacts of adding new eHealth technologies to the existing health care system for alcohol use disorders.

Methods

Target Population

By way of input, ALCMOD requires data that describe key characteristics of the target population. Selecting the name of the country will automatically trigger ALCMOD to upload the age and gender distribution of the population of the selected country and the corresponding mortality rates. ALCMOD also needs to know the size of the target population, and in the Netherlands, the target population consists of 993,200 men and 222,800 women aged 18 to 69 years who could be classified as *problem drinkers* [15]. Other required input is the preintervention profile of the target population based on the Alcohol Use Disorders Identification Test, the AUDIT [16-19]. The decision to base ALCMOD on the AUDIT was motivated by the idea that the AUDIT is globally used. Moreover, new eHealth interventions will screen participants with the AUDIT. Thus, even when a country has no AUDIT data yet, these data are likely to become available via eHealth interventions in the near future. In the Netherlands, data from the AUDIT are available and can be automatically uploaded in ALCMOD.

Intervention Packages

ALCMOD allows a description of the intervention mix representing the base-case scenario and the designing of an alternative scenario with a different mix of interventions or new interventions added to existing ones. In ALCMOD's default setting, a range of interventions—both face-to-face interventions and eHealth interventions—are shown for heavy, hazardous, and harmful alcohol use and alcohol dependence. Two parameters need to be set for each of the interventions: the coverage rate and the adherence rate.

Coverage Rate

When some of the interventions shown in ALCMOD's default setting are not available in a country, then their coverage rate has to be set to 0%. This is equivalent to saying that the intervention is not offered in a country. Other interventions might be available for every person belonging to the target population, and the coverage rate is then set to 100% (universal coverage). However, due to the many obstacles to full implementation, the coverage rate of most interventions is somewhere between 0 and 100% and can be set accordingly in ALCMOD.

Adherence Rate

Recipients of interventions might be less than willing or unable to fully comply with the intervention, and the degree of adherence is likely to moderate treatment response. Therefore, the adherence rate is an important parameter when evaluating the effectiveness of interventions. Adherence rates may be obtained from the literature, experts, or via focus groups in the target population.

The idea is that health care scenarios can be developed by changing the level of coverage for a series of interventions. Table 1 shows the settings for the three scenarios that we modeled: (1) the current Dutch health care system for alcohol use disorders without eHealth interventions (base-case scenario), (2) the Dutch health care system augmented with the eHealth interventions (alternative scenario 1), and (3) the Dutch health care system where face-to-face interventions have been substituted for 50% by the new eHealth interventions (alternative scenario 2).

Table 1. Modeled scenarios: coverage rates (%) for each of the interventions

Target Group Alcohol Use Disorder	Intervention	Base-Case Scenario	Alternative Scenario 1	Alternative Scenario 2
Heavy	Brief face-to-face intervention[a]	10	10	5
	Online brief intervention [b]	0	5	5
Hazardous	Brief face-to-face intervention[a]	10	10	5
	Online brief intervention[b]	0	5	5
	Behavioral intervention[c]	6	6	3
	Online behavioral intervention[d]	0	6	3
Harmful	Behavioral intervention[c]	9	9	3
	Online behavioral intervention[d]	0	9	3
	Online therapist-led treatment[e]	0	9	3
	Detox and acamprosate[f]	5	5	5
	Aftercare and rehab with AA[g]	5	5	5
Dependence	Behavioral intervention[c]	5	5	2.5
	Online therapist-led treatment[e]	0	5	2.5
	Detox and acamprosate[f]	5	5	5
	Aftercare and rehab with AA[g]	5	5	5

[a] Brief face-to-face is modeled as a brief intervention consisting of screening followed by personalized feedback by a physician usually in a single session (< 10 minutes), occasionally in two sessions (one for screening, the other for personalized feedback).

[b] Online brief intervention is modeled as online screening and automated personalized feedback (DrinkTest).

[c] Behavioral intervention is modeled as eight to ten sessions of individual cognitive behavioral therapy (CBT) under the guidance of a therapist, followed by one booster session.

[d] Online self-help intervention (DrinkingLess) is modeled as four (range 3 to 12) sessions of online interactive CBT-based self-help preceded by referral by a general practitioner (GP).

[e] Online therapist-led intervention is modeled as eight sessions of online therapist-led CBT.

[f] Detox is modeled as 1-week ambulatory detoxification followed by clinical management with acamprosate.

[g] Aftercare and rehabilitation is modeled as participation in Alcoholics Anonymous (AA) over 12 months.

The choice of intervention mix was informed by Room et al and Benegal et al [17,19] and the Dutch multidisciplinary guideline for the treatment of alcohol use disorders [20]. The choice of interventions was also motivated by two additional considerations: availability of evidence of the intervention's effectiveness in the meta-analytical literature [21] and the nonoverlapping *independent* nature of the interventions such that each intervention could be added to other interventions without creating overlap for a specific alcohol use disorder. Finally, the scenarios have been simplified by assuming that all interventions are associated with an adherence rate of 50%. This was done to ensure that differences in the cost-effectiveness ratios are due to fundamental differences in health technologies and not simply an effect of greater or lesser treatment adherence. However, it is possible to adjust adherence rates in ALCMOD. After all, some interventions might be associated with better or poorer adherence, and adherence itself might be amenable to intervention such as motivational enhancement. Changing the adherence parameters allows evaluation of these issues.

Cost and Effect Parameters

In the ALCMOD default settings, some of the intervention parameters have been preset and need not be changed but can be changed if so required. These parameters are the costs and the effects of the interventions.

Costs

ALCMOD's default setting makes use of the full economic cost price of each of the interventions. To be precise, the costs are the per-participant costs of delivering an intervention expressed in euros (€) for the Netherlands in the reference year 2009 (see Table 2). The costs are based on the amount of resources (labor, facilities, and supplies) used for offering the intervention during its postimplementation stage. We made our own costing tool to estimate the costs (in euros) of interventions in a systematic and uniform way that is in agreement with the Dutch guideline for costing health care interventions [22]. For other countries, the per-participant costs of offering an intervention need to be assessed. These assessments can be carried out with the help of an auxiliary costing tool, for example Cost It, available from WHO's CHOICE website. Neither costs nor effects are discounted because ALCMOD takes a short-term (12-month) perspective.

Effects

Intervention effects are expressed as the standardized mean difference, also known as Cohen's d. This metric indicates how many standard units (on a scale of standard deviations) the experimental group has improved relative to a control group on a relevant outcome such a change in drinking behavior. The effect size d is often reported in the meta-analytical literature and gives access to a large body of scientific evidence. We extracted effect sizes at 6- or 12-months follow-up for all the interventions from the meta-analytical literature and our own research (see Table 3) and these values were used to populate ALCMOD with its default parameter settings.

Table 2. Per-patient intervention costs in 2009 euros (€) within uncertainty range (based on 1000 simulations)

Target Group Alcohol Use Disorder	Intervention	Costs in Euros	Uncertainty Range (Euros)	
			Low	High
Heavy	Brief face-to-face intervention[a]	58	52	75
	Online brief intervention[b]	10	9	10
Hazardous	Brief face-to-face intervention[a]	58	52	75
	Online brief intervention[b]	10	9	10
	Behavioral intervention[c]	2024	1702	2550
	Online self-help intervention[d]	207	198	224
Harmful	Behavioral intervention[c]	2024	1702	2550
	Online self-help intervention[d]	207	198	224
	Online therapist-led intervention[e]	764	227	1451
	Detox and acamprosate[f]	1800	1620	2232
	Aftercare and rehab with AA[g]	500	250	750
Dependence	Behavioral intervention[c]	2024	1702	2550
	Online therapist-led intervention[e]	1276	979	1408
	Detox and acamprosate[f]	1800	1620	2232
	Aftercare and rehab with AA[g]	500	250	750

[a] Brief face-to-face intervention modeled as screening at € 5.70 followed by 1 or 2 (Poisson distributed) 10-minute contacts with GP at € 32.03 per contact

[b] Online brief intervention (DrinkTest) modeled as 40% of target population (N = 1,255,000) reached with information about the website, 8% responding to AUDIT screener and receiving automated personalized feedback. Per-participant annual costs include website upgrading at € 50,000 research at € 50,000 and hosting at € 25,000.

[c] Behavioral intervention is modeled as 8 to 14 (Poisson distributed) sessions of cognitive behavioral therapy (CBT) under guidance by a therapist, including referral, intake, and one booster session

[d] Online self-help intervention (DrinkingLess) is modeled as 15% of target population (N = 1,255,000) reached with information about the website 5% uptake rate, and 4 sessions (range 3 to 12) of online CBT-based self-help preceded by referral by a GP. Per-participant annual costs include € 75,000 for website upgrading, € 50,000 for research, € 25,000 for hosting, plus € 75,000 for moderating forum and technical assistance.

[e] Online therapist-led intervention is modeled as an average of 4 sessions (range 1 to 9) of 45 minutes each of online therapist-led CBT, preceded by GP referral. Per-participant costs include per annum costs of € 8000 for website upgrading, € 5000 for hosting, plus € 2000 for technical assistance.

[f] Detox is modeled as a 1-week ambulatory detoxification followed by clinical management with acamprosate under the supervision of a substance use disorder treatment specialist and a physician over 3 months.

[g] Aftercare and rehabilitation is modeled as participation in Alcoholics Anonymous at an average of € 500 (range € 250 to € 750) per patient for a year.

It is worth noting that ALCMOD uses two types of effects: the standardized mean difference, d, which was just discussed, and the impact of an intervention in terms of the percent reduction of pure alcohol intake in grams per day (g/day). The former effect (d) impacts on health-related quality of life (QOL) via changes in disorder severity. ALCMOD uses the percent reduction of pure alcohol intake to model treatment effects on mortality (see below for details). Although ALCMOD can handle different alcohol reduction rates for each of the modeled interventions, we have assumed a pre-post reduction of alcohol intake by 20% for all interventions [21,23], because reduction of alcohol intake was not always reported in the literature. This should not overly distort outcomes because the short-term effects of alcohol use on mortality are small, thus limiting their impact on disease burden as measured by disability-adjusted life years (DALYs). We say this on the understanding that alcohol-related mortality becomes an important, even a dominant, factor when disease burden is modeled out to full life expectancy, especially in the more severe alcohol use disorders.

Here we need to address a final point about the required input for ALCMOD. ALCMOD can be operated in two modes: deterministic and stochastic. In deterministic mode, ALCMOD does not take into account the uncertainty in parameters such as costs and effects. ALCMOD conducts all computations, but only once, and these calculations are primed on the mean value of all parameters. Much of ALCMOD's output, which is based on uncertainty, is then disabled. However, in stochastic mode, ALCMOD can handle uncertainty surrounding the cost (in euros) and effect (d) parameters. Our costing tool assesses the uncertainty in costs with the help of simulations of resource use (with 1000 iterations), and both randomized trials and meta-analyses of trials often report 95% confidence intervals of the effect size d. Thus we assume that costs are surrounded by an uncertainty range, and effects, by a 95% confidence interval, both having a lower and an upper limit. ALCMOD assumes a gamma distribution for costs and the normal distribution for the effect size d. Both distributions can be specified in ALCMOD such that the distributions fit within the lower and upper limits of costs and effects. In stochastic mode, ALCMOD then proceeds with drawing random values from these distributions, conducts all the computations, and repeats this process many times (maximum 10,000 times). This helps to capture uncertainty in the input parameters.

Table 3. Effectiveness of the interventions: standardized mean differences, 95% confidence interval for d (95% CI), difference in pure alcohol intake (mg/day) and references

Target Group Alcohol Use Disorder	Intervention	d	95% CI
Heavy	Brief face-to-face intervention[a]	0.26	0.20 to 0.32
	Online brief intervention[b]	0.19	-0.02 to 0.40
Hazardous	Brief face-to-face intervention[c]	0.32	0.23 to 0.42
	Online brief intervention[b]	0.19	-0.02 to 0.40
	Behavioral intervention[d]	0.34	0.12 to 0.56
	Online self-help intervention[e]	0.31	-0.69 to 1.30
Harmful	Behavioral intervention[d]	0.34	0.12 to 0.56
	Online self-help intervention[e]	0.31	-0.69 to 1.30
	Online therapist-led intervention[f]	0.58	0.29 to 0.88
	Detox and acamprosate[g]	0.21	0.14 to 0.29
	Aftercare and rehab with AA[h]	0.28	0.20 to 0.37
Dependence	Behavioral intervention[i]	0.32	0.05 to 0.59
	Online therapist-led intervention[f]	0.59	0.30 to 0.90
	Detox and acamprosate[g]	0.21	0.14 to 0.29
	Aftercare and rehab with AA[h]	0.28	0.20 to 0.37

[a] Moyer et al's [21] meta-analysis of brief face-to-face interventions in approximately 4300 users meeting criteria of at least heavy drinking.

[b] Randomized trial of 450 participants presenting with either excessive alcohol consumption (> 20 units weekly) and/or binge drinking (> 5 units on a single occasion on least one day per week) in the past 6 months [9].

[c] Reanalysis of Beich et al’s [24] meta-analysis of brief face- to-face interventions in 2989 users meeting criteria of hazardous drinking.

[d] Walters’ [25] meta-analysis based on approximately 320 harmful users.

[e] Randomized trial of 261 excessive drinkers from the general population [11] where odds ratio (OR) converted into d using Chinn's equation [26].

[f] Randomized trial of 250 adults with mean AUDIT score of 20 at baseline with intervention was online treatment versus waiting list at 3 months and the AUDIT as outcome [13].

[g] Mann et al's [27] meta-analysis of 1670 people receiving acamprosate after detoxification where odds ratios converted into d using Chinn's method [26].

[h] Tonigan et al's [28] meta-analysis of 2097 harmful and dependent users where effect size r converted into d.

[i] Walters’ [25] meta-analysis based on approximately 210 dependent users.

ALCMOD's Throughput

Differences in Costs

Modeling cost differences between two health care systems is straightforward once the per-participant costs of delivering all modeled interventions have been estimated and when the coverage rates of the interventions have been established. The number of people in the target group (stratified by alcohol use disorder) is then multiplied by the coverage rate of each intervention and multiplied by the appropriate per-participant full economic cost price. The cost analyses are always conducted for both the base-case and alternative scenario, such that the cost difference between two modeled health care systems can be computed and expressed as incremental costs.

Differences in Disease Burden

The disability adjusted life year (DALY) is a measure of disease burden in a population. It combines two components of disease burden: morbidity and mortality. The first is related to lesser quality of life due to disability. Mortality arises when illness is associated with premature death. Thus, a DALY can be computed as the sum of years lost due to disability (YLD) plus years of life lost (YLL) due to mortality, hence, DALY = YLD + YLL.

The first term in the DALY equation, YLD, can be computed as the number of cases manifesting with an alcohol use disorder, N (point prevalence), weighted by a disability weight, DW. Thus, YLD = N × DW. DWs range from 0 to 1, where 0 is *no burden* (good health) and 1 refers to a health condition as undesirable as death. Although the literature offers advice for choosing DWs for the alcohol use disorders [29-33], ALCMOD makes no use of DWs that are directly associated with each of the disorder-specific health states. Instead, it computes the (downward) shift in DW as a consequence of the treatment effect d. As said, d is the standardized mean difference indicating how many standard units the treatment group has moved away from the group that received no care. Thus, d is essentially a “health improvement shift” due to intervention. The task at hand, then, is to “translate” the health improvement shift (of size d) into a corresponding shift in DW. This strategy has been developed by Sanderson et al [34] who used a panel of experts for obtaining a conversion factor of 0.18 (95% CI 0.16-0.20) to translate a shift in d into a shift in DW in alcohol use disorders. The change in DW is then multiplied by the appropriate number of people to arrive at an estimate of the number of YLD avoided. When running in stochastic mode, ALCMOD automatically conducts extensive uncertainty analyses around Sanderson's conversion factor.

The second term in the DALY equation, YLL, is calculated as the difference in life expectancy when people reduce drinking levels. We obtained estimates of the gender-specific relative risk, RR, of all-cause mortality attributable to pure ethanol intake (in g/day) using the expression [35], $\ln(RR) = b_1*(\ln(x+1)) + b_2*\ln(x) + e$, where x = grams of pure ethanol intake per day and b_1 and b_2 are -0.1030 and 0.0035 for men and -0.0645 and 0.0029 for women, respectively (our own estimates from Gmel's paper [35]). Exponentiating ln(RR) gives the relative risk, RR, and the RRs are then combined with the gender and age-specific mortality rates of the country for which the outcomes are modeled. This produces estimates of changes in life expectancy due to changes in alcohol intake. Because ALCMOD takes a short-term perspective, treatment induced impacts on life expectancy were calculated as the number of avoided deaths in the present year.

The difference in YLD and YLL between the scenarios determines the difference in the disease burden as measured by DALY between two modeled health systems, the so-called incremental effects.

ALCMOD offers the use of a (downward) *attenuation factor* that reduces the carry-over effects from lesser drinking to lesser mortality and better health-related quality of life. After all, it can be assumed that former drinkers still have a higher risk of dying and poorer quality of life than people who never drank before or have been consistent moderate drinkers [36,37]. In other words, returning to less risky drinking levels is assumed to be beneficial but not as beneficial as a history of no drinking or moderate drinking. Hence this attenuation factor, which can be used to conduct sensitivity analyses for further evaluation of this issue. In all subsequent analyses we used a downward adjustment of 20% to be on the conservative side.

Combining Costs and Effects

Once the treatment costs and the reductions in DALY disease burden have been computed for each scenario, then it is a small step to also compute the difference between the costs of both health care systems as $\Delta(C) = C_1 - C_0$ and the difference between the effects as $\Delta(E) = E_1 - E_0$. The ratio, $\Delta(C)/\Delta(E)$, is the incremental cost-effectiveness ratio (ICER), which tells us if the alternative health care system offers better value for money than the current health care system.

Handling Uncertainty

As indicated, in stochastic mode, ALCMOD takes parameter uncertainty into account. The uncertainty is captured by drawing values from the cost and effect distributions of all interventions in a random way and basing the calculations on these randomly drawn values. This can be repeated n times (in practice 500 times appears to be sufficient) and the outcomes of each of the iterations is stored in vectors of size n of the costs and effects of each of the scenarios, their differences, and the ICER. Following standard health economic modeling routines [37], the vectors are then used to produce ALCMOD's output, such as the mean and the median of the outcomes and several ICER plots and graphs.

Results

Preintervention Target Group

In November 2009 we obtained data from the DrinkingLess monitoring system on 1083 women and 2538 men who participated in DrinkingLess. Their mean age was 44.7 years (SD 10.7). Table 4 presents the observed AUDIT distribution for this population and the relative risk (RR) of premature death due to alcohol, as computed by ALCMOD.

Table 4. Preintervention characteristics of the target population

AUDIT Score	Tentative Label	Men, % (n = 987,000)	Women, % (n = 267,000)	RR (Death) Men	Women
0 - 1	Abstinent	0.1	0.1	1.00	1.00
2 - 7	Moderate	1.6	3.6	0.86	0.96
8 - 15	Heavy	18.4	23.5	0.95	0.99
16 - 19	Hazardous	22.2	23.5	0.99	1.05
20 - 29	Harmful	50.1	43.4	1.10	1.12
30 - 40	Dependence	7.6	5.9	1.36	1.28

Comparing Current Care With New eHealth Interventions Added

We begin by comparing the current health care system (base-case scenario) with an alternative scenario where eHealth interventions are added to conventional care. In this comparison, it was assumed that the new eHealth interventions would attract a different segment from the target population—a segment that would otherwise not have been the recipient of conventional care. Making this (unrealistic) assumption is a conscious choice, and we will return to it in the “Discussion” section. The results are as follows.

The total health care costs in the base-case scenario are € 233 million. Adding new eHealth interventions would raise the health care expenditure to € 319 million, an increase of € 86 million. Under the base-case scenario, 5022 DALYs are averted; under the new scenario, this is doubled to 10,319 averted DALYs, an additional 5296 averted DALYs (including 32 alcohol-related deaths that are avoided under the new scenario). Thus, the alternative health care system delivers more population health albeit at higher costs. Figure 1 provides a corresponding visualization: the scatter of simulated ICERs (due to uncertainty in the input parameters) corresponding to the alternative scenario is placed more to the north (more costs) and more to the east (more health) than the scatter belonging to the current health care system. Now, investing € 86 million for averting 5296 extra DALYs (ie, € 16,053/DALY) raises the question whether that would be money spent wisely. In the Netherlands, the willingness to pay for one averted DALY is about € 80,000 with a lower bound of € 50,000. An even more conservative willingness-to-pay ceiling is customarily set at € 20,000/DALY for nonfatal and mild disorders. It follows that the estimated € 16,053/DALY falls well below any of the usual willingness-to-pay ceilings.

Figure 1. Total costs and effects in millions of euros (base-case scenario versus eHealth interventions added)

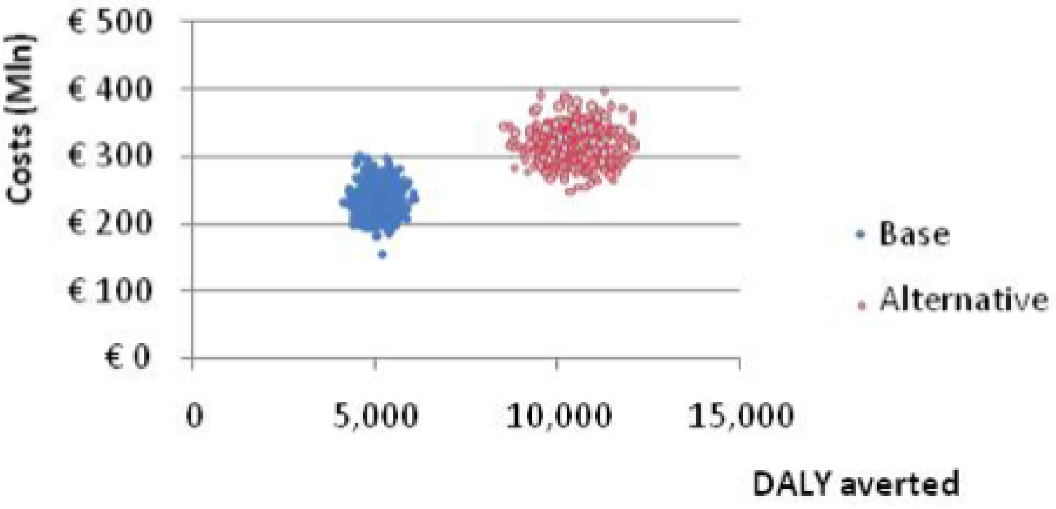
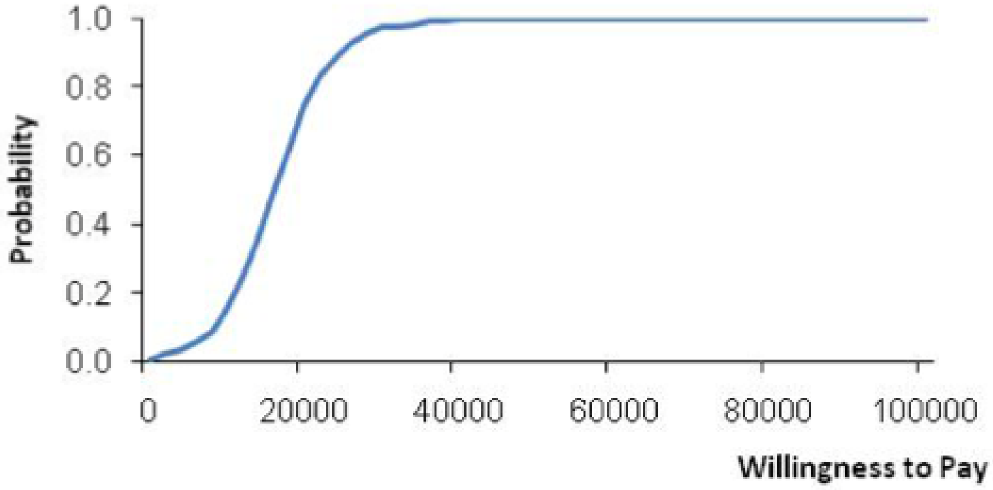


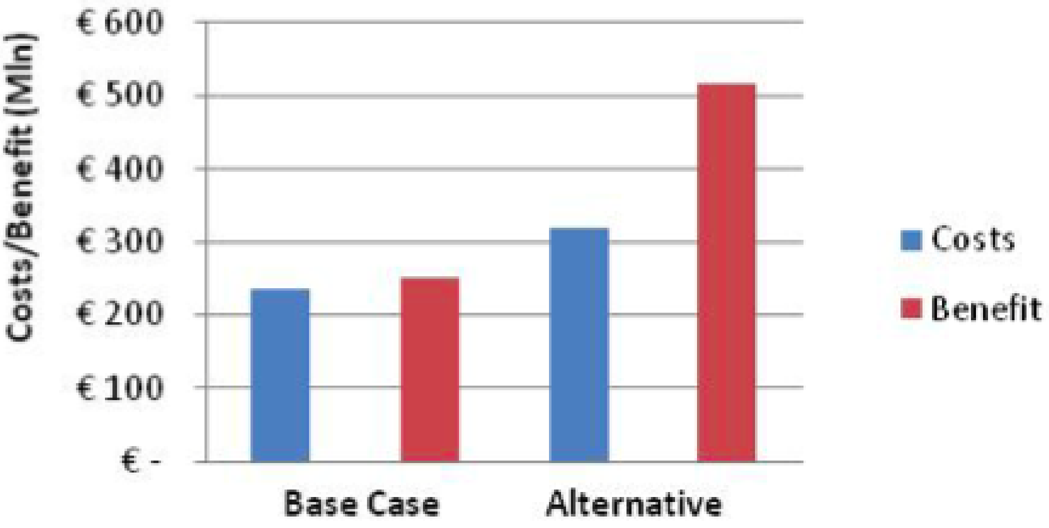
Figure 2, the ICER acceptability curve, represents a slightly different approach to the same issue. It depicts the probability that we must conclude that the new health care system is more cost-effective than the current system (vertical axis) for a range of willingness-to-pay ceilings (horizontal axis). For the simulated data, Figure 2 shows that the likelihood that the new health care system must be regarded as cost-effective increases sharply with increasing willingness-to-pay ceilings: the probability equals 0% when the willingness to pay for an additional health gain of one DALY averted is € 0, increases to 50% at € 16,000 and to 75% at € 20,000. Beyond € 30,000 the probability approaches certainty, and the conclusion that we must regard the new system as more cost-effective is no longer affected by higher willingness–to-pay levels. Again, accepting the threshold of € 20,000/DALY implies that the new health system compares favorably with the current system in terms of cost-effectiveness.

Figure 2. ICER acceptability curve (base-case scenario versus eHealth interventions added)



Assuming for a moment that the willingness to pay for averting one DALY is €50,000, then we could directly compare the costs of the health care system (in euros) with health gains (also expressed in euros) by multiplying the averted DALYs by €50,000. Figure 3, a cost/benefit chart, shows ALCMOD's simulation results. The chart shows that costs and benefits are just balanced under the current health care system, while the benefits clearly outweigh the costs under the new system. To be more precise, every euro spent under the current system returns a value of about the same size (€1.08, ie, a “surplus” of 8 euro cents), while the new health care system offers much better returns on investment: every euro spent generates €1.62 in health-related value.

Figure 3. Cost-benefit chart in millions of euros (base-case scenario versus eHealth interventions added)



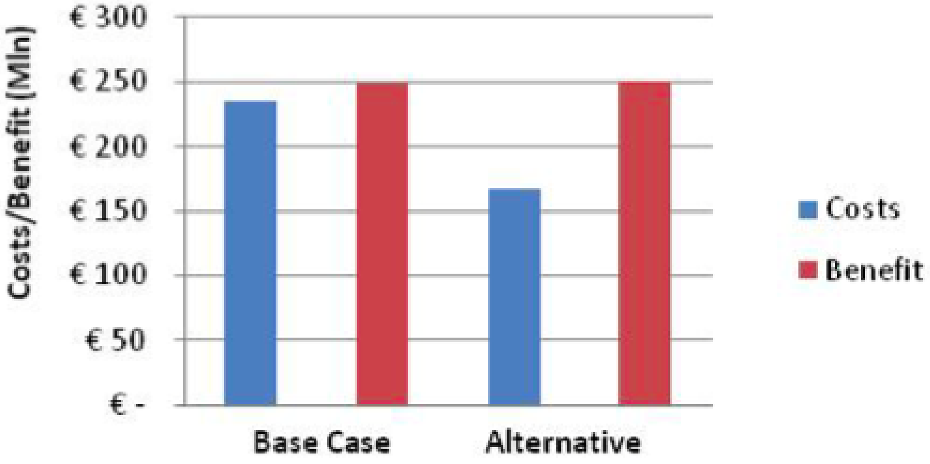
To summarize, the new health care system, with eHealth interventions added, is associated with higher health care delivery costs overall, but it would be a health care system which is more efficient than the current one, offering better value for money.

Comparing Current Care With New eHealth Interventions With Partial Substitution

We also simulated another comparison, this one between a base-case scenario that represents the current system and an alternative scenario with eHealth interventions added. In this case, however, the conventional face-to-face interventions are partly substituted by the new eHealth interventions. In this scenario, the coverage rate remains the same before and after the introduction of the eHealth interventions. Such a situation would arise if the eHealth interventions were to tap into the same target population seeking professional help, whether face-to-face, eHealth, or otherwise. In this scenario, the number of people who receive health care remains the same before and after the introduction of the new health technologies, and interventions are competing for the same target population and therefore partially substitute each other.

In this scenario, ALCMOD computes that the number of DALYs averted under both systems is virtually the same: 4984 DALYs under the current system and (exactly) 5000 DALYs under the new system. In other words, partial substitution of conventional face-to-face interventions by eHealth interventions does not have any appreciable impact on population health. However, the overall cost of the new system is much lower at €166 million than the cost of the current system of €234 million, resulting in a cost saving of € 68 million. Figure 4 relays the same information. Again assuming a willingness to pay of €50,000/DALY, the cost-benefit ratio indicates that for every euro invested the generated health revenues are worth €1.06 (ie, 6 euro cents surplus for every euro invested) under the current health care system. This improves to 52 euro cents surplus for every euro invested under the new scenario where the face-to-face interventions have been partly substituted by new eHealth interventions.

Figure 4. Cost/benefit chart in millions of euros (base-case scenario versus partial substiution scenario)



Discussion

Main Findings

The main rationale for introducing eHealth technologies is to increase timely access to health services, to reduce the costs of delivering health care, and to make more efficient use of the health care workforce. Indeed, ALCMOD's simulation results suggest that widespread implementation of eHealth interventions for alcohol use disorders would help to substantially increase population health in the Netherlands, albeit at higher system costs, when eHealth interventions are added to the existing health care system and more people become the recipients of the expanded system. The cost-effectiveness of the Dutch health system would also substantially improve if the new eHealth interventions were partially replacing some of the current face-to-face interventions. Then, adding eHealth interventions becomes a cost-effective option, because it will produce the same level of population health for a significantly smaller health care budget. The “truth” might be found somewhere between both extremes, because it is unlikely that the new eHealth interventions will exclusively recruit people that would otherwise not have been the recipients of conventional health care (as assumed in the first comparison), while it is also unlikely that the new eHealth interventions will tap into exactly the same pool of health care users (as assumed in second comparison). At any rate, both extreme scenarios carry the message that widespread introduction of eHealth technologies would help to substantially increase the efficiency of the Dutch health care system overall, with a more favorable cost-benefit ratio either way.

Strengths and Limitations of ALCMOD

One of the benefits of a simulation model is that it helps to organize vast fields of knowledge across several disciplines. In the case of ALCMOD, these disciplines encompass addiction epidemiology and health economics, while the evidence that supports effect parameters is drawn from randomized clinical trials, meta-analyses, and evidence-based clinical guidelines. In addition, a model makes all the necessary information available in a dynamic form, permitting “what if” analyses. This could be of assistance to policy formulation. ALCMOD is therefore best seen as a decision-making support tool, capable of giving almost instant feedback on policy-makers' attempts to find an optimal solution in the context of constrained decision-making in a complex environment. ALCMOD can also be employed for setting research agendas. After all, it helps to identify those parameters that have an impact on health gains and costs. When some of these parameters are surrounded by a nonnegligible degree of uncertainty, then empirical research is recommended, with the aim of reducing uncertainty in those parameters. Furthermore, ALCMOD can assist in identifying opportunities for system innovation by simulating hypothetical interventions, for example, an adjunctive intervention that helps to enhance treatment adherence. Among other strengths of ALCMOD are its adaptability to other countries, settings, and target groups and its capability to explicitly model treatment coverage and adherence rates. Finally, ALCMOD conducts automated multivariate uncertainty analyses to quantify uncertainty in costs, effects, and related parameters.

ALCMOD is subject to several limitations that need to be taken into account when interpreting ALCMOD's outcomes. First, ALCMOD's outcomes are modeled as *steady-state* population averages, and it is not clear when a health care system finds equilibrium after the introduction of new health technologies. This is unlikely to occur instantaneously and might take as long as several years. Second, it should be borne in mind that the introduction of new health technologies entails costs of its own, but the costs of introducing new technologies are not incorporated in ALCMOD's output. In fact, ALCMOD's output captures only the costs of offering a package of interventions once the interventions have been fully implemented. However, it will always take effort, time, and expenditure before the results of an improved health care system become visible in real life. Third, introduction of eHealth technologies may have unforeseen consequences that may increase longer-term health care costs, for example, by supply-induced demand for health care, thus attracting people to the health care system who otherwise would not have become dependent on (expensive, face-to-face delivered) health care. Fourth, it should be understood that ALCMOD focuses on short-term health impacts. Thus, ALCMOD ignores the longer-term impacts on quality of life, mortality, and health care utilization and it should be understood that longer-term impacts depend, in part, on a wide range of alcohol-related disorders that usually occur later in life. Since these longer-term effects are mainly related to the more severe alcohol use disorders, ALCMOD is unlikely to capture the full benefits of interventions for the severe disorders and may thus give undue weight to the less severe disorders. Fifth, ALCMOD is limited in that it only models clinical interventions while disregarding other alcohol-control options, such as banning alcohol advertising, taxing, restricting access to alcoholic beverages, and improving road safety, although these nonclinical interventions are likely to be (very) cost-effective [38]. In the same vein, ALCMOD regards only the cost impacts incurred by the health care system, while disregarding costs and cost-offsets outside the health care system, such as patients' out-of-pocket payments to access services, changes in labor productivity, and costs incurred by the criminal justice system. To summarize ALCMOD's basic assumptions: ALCMOD only models incremental health gains and health care delivery costs over the shorter time horizon, assuming a steady state in the modeled health care systems. See Textbox 2 for a summary of ALCMOD's assumptions and their justifications.

Textbox 2. ALCMOD's assumptions and justifications

General assumptions:

- ALCMOD disregards the longer-term downstream costs, cost offsets, and health effects due to less drinking because the empirical literature rarely reports treatment effects beyond 12 months.
- Per-participant costs are assumed to follow a gamma distribution [38].
- Treatment effects, expressed in standardized mean difference scores, d, are assumed to follow the standard normal distribution, because d is almost equivalent to a *z*-score.
- The YLD (quality of life) differential is based on Sanderson et al’s conversion factor [34], which translates a change in disorder severity of size d induced by an intervention into a corresponding shift in the disability weight (DW) used in the YLD calculations.
- The YLL (mortality) differential is based on Gmel et al’s [35] relative risk of all-cause mortality stratified for level of pure ethanol intake (g/day).
- Costs and DALY outcomes have not been discounted because the focus is on short-term (<12 months) postimplementation (steady state) health and budget impacts.

Additional assumptions for the current simulations:

- The AUDIT distribution obtained from DrinkingLess is representative for the target population because this is a population of (former) problem drinkers still at risk of an alcohol use disorder and willing to seek treatment.
- Adherence rate is 50% for all interventions because a constant figure would help to obtain a clear view on cost-effectiveness due to fundamental changes in health care technologies.
- Alcohol intake is reduced by 20% after all interventions because the short-term contribution of YLL to the DALY disease burden is virtually negligible.
- All treatment effects on YLD and YLL are attenuated by 20% because the detrimental health effects of problem drinking are likely to linger on—even after return to moderate drinking or abstinence.

Conclusion

It is not immediately clear if our findings are valid for countries other than the Netherlands. After all, in low-income countries, labor might be less costly than the capital inputs required for the new eHealth technologies. Also the population's access to the Internet could be an issue. Moreover, one could encounter cultural obstacles to using the Internet for alcohol use disorders. Such factors might impinge on coverage and adherence rates and mitigate impacts on population health, ultimately diminishing the cost-effectiveness of new health technologies. To illustrate, in the Netherlands, close to 90% of the population has access to the Internet, and Internet usage is distributed fairly evenly across demographic groups, but in other countries, Internet usage might be concentrated in only some population segments. In addition, it is worth noting that the emergence of mobile technologies may offer an opportunity to offer eHealth interventions for population segments that otherwise might be hard to reach. Therefore, the question as to whether eHealth will deliver the same benefits to other countries is best addressed per country, per setting, and per target group. Ante hoc assessment of the cost-effectiveness of innovations in health care systems may help to inform policy decisions. ALCMOD was created with exactly these aims in mind.

We recommend that ALCMOD be used in an iterative consensus building process that encompasses all pertinent stakeholders (eg, health care users, health care providers, health care financiers, and health policy planners) who can review and make amendments to modeled scenarios. Recently, we had an encouraging experience with such an approach while using a similar model for the treatment of depressive disorder. In any case, we would advise against using ALCMOD as an autopilot for policy making. After all, setting priorities for health care delivery is about acceptability and equity, as well as about cost-effectiveness considerations. As always, we need to base decisions on the best judgments and evidence available, but the evidence that informed ALCMOD points toward the conclusion that eHealth interventions can help to bridge the mental health gap by bringing scalable and cost-effective health services within reach of all who have access to the Internet—literally at their fingertips.

Acknowledgments

This work was conducted within the framework of World Health Organization's International Action Plan on Implementing eHealth Technologies for Substance Abuse. The authors acknowledge, with many thanks, financial support from the Netherlands' Ministry of Health, technical support by the Economics Advisory of PricewaterhouseCoopers, Amsterdam, and are grateful for the expertise that was kindly lent by the staff of WHO Headquarters. The views expressed in this paper are those of the authors alone.

Conflicts of Interest

None declared

References

1. Ezzati M, Lopez AD, Rodgers A, Murray CJL. Comparative Quantification of Health Risks: Global and Regional Burden of Diseases Attributable to Selected Major Risks. Geneva, Switzerland: World Health Organization; 2004.
2. Rehm J, Taylor B, Room R. Global burden of disease from alcohol, illicit drugs and tobacco. Drug Alcohol Rev 2006 Nov;25(6):503-513. [doi: 10.1080/09595230600944453] [Medline: 17132570]
3. Mohapatra S, Patra J, Popova S, Duhig A, Rehm J. Social cost of heavy drinking and alcohol dependence in high-income countries. Int J Public Health 2010 Jun;55(3):149-157. [doi: 10.1007/s00038-009-0108-9] [Medline: 20024666]
4. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009 Jun 27;373(9682):2223-2233. [doi: 10.1016/S0140-6736(09)60746-7] [Medline: 19560604]
5. Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de Graaf R, Beekman A. Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ 2006 Dec;9(4):193-200. [Medline: 17200596]
6. World Health Organization. WHO Mental health Gap Action Programme: scaling up care for mental, neurological, and substance use disorder. Geneva, Switzerland: World Health Organization; 2008.
7. Beaglehole R, Bonita R. Alcohol: a global health priority. Lancet 2009 Jun 27;373(9682):2173-2174. [doi: 10.1016/S0140-6736(09)61168-5] [Medline: 19560583]
8. World Health Organization. Lexicon of alcohol and drug terms published by the World Health Organization URL: http://www.who.int/substance_abuse/terminology/who_lexicon/en/ [accessed 2011-07-19] [WebCite Cache ID 60I82RKgg]
9. Meijer S, Smit F, Schoemaker C, Cuijpers P. Gezond verstand: evidence-based preventie van psychische stoornissen Common sense: evidence-based prevention of mental disorders De Bilt: RIVM, centrum Volksgezondheid Toekomst Verkenningen en Trimbos-instituut, centrum Preventie en Kortdurende Interventie. In: Gezond verstand: evidence-based preventie van psychische stoornissen. Bilthoven: RIVM; 2006.
10. Boon B, Risselada A, Huiberts A, Riper H, Smit F. Curbing alcohol use in male adults through computer generated personalized advice: randomized controlled trial. J Med Internet Res 2011;13(2):e43 [FREE Full text] [doi: 10.2196/jmir.1695] [Medline: 21719412]
11. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
12. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-Self-help Interventions for Curbing Adult Problem Drinking: A Meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]

13. Blankers M, Koeter MW, Schippers GM. Internet therapy versus internet self-help versus no treatment for problematic alcohol use: A randomized controlled trial. J Consult Clin Psychol 2011 Jun;79(3):330-341. [doi: 10.1037/a0023498] [Medline: 21534652]
14. Tate DF, Finkelstein EA, Khavjou O, Gustafson A. Cost effectiveness of internet interventions: review and recommendations. Ann Behav Med 2009 Aug;38(1):40-45. [doi: 10.1007/s12160-009-9131-6] [Medline: 19834778]
15. Van Laar M, Cruts AAN, Van Ooyen-Houben MMJ, Meijer RF, Brunt T. Netherlands National Drug Monitor: NDM Annual Report 2009. Utrecht, the Netherlands: Trimbos Institute; 2010.
16. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The Alcohol Use Disorders Identification Test: Guidelines for use in primary care. 2nd edition. Geneva, Switzerland: World Health Organization; 2001.
17. Room R, Babor T, Rehm J. Alcohol and public health. Lancet 2005;365(9458):519-530. [doi: 10.1016/S0140-6736(05)17870-2] [Medline: 15705462]
18. Reid MC, Fiellin DA, O'Connor PG. Hazardous and harmful alcohol consumption in primary care. Arch Intern Med 1999;159(15):1681-1689 [FREE Full text] [Medline: 10448769]
19. Higgins CB, Roeske WR, Karliner JS, O'Rourke R, Berk RN. Predictive factors and mechanism of arrhythmias and myocardial ischaemic changes in elderly patients during barium enema. Br J Radiol 1976 Dec;49(588):1023-1027. [Medline: 1000170]
20. Landelijk Kwaliteitsinstituut CBO & Trimbos Instituut. Multidisplinaire richtlijn stoornissen in het gebruik van alcohol Multidisciplinary Guideline Alcohol Use Disorders. Utrecht, the Netherlands: Trimbos Instituut; 2009.
21. Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002 Mar;97(3):279-292. [Medline: 11964101]
22. Oostenbrink JB, Bouwmans CAM, Koopmanschap MA. Manual for Costing: Methods and Standard Costs for Economic Evaluations in Health Care (in Dutch). Diemen, the Netherlands: Health Insurance Board; 2004.
23. Davis WT, Campbell L, Tax J, Lieber CS. A trial of "standard" outpatient alcoholism treatment vs. a minimal treatment control. J Subst Abuse Treat 2002 Jul;23(1):9-19. [Medline: 12127464]
24. Beich A, Thorsen T, Rollnick S. Screening in brief intervention trials targeting excessive drinkers in general practice: systematic review and meta-analysis. BMJ 2003 Sep 6;327(7414):536-542 [FREE Full text] [doi: 10.1136/bmj.327.7414.536] [Medline: 12958114]
25. Walters GD. Behavioral self-control training for problem drinkers: a meta-analysis of randomized control studies. Behavior Therapy 2000;31(1):135-149.
26. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000 Nov 30;19(22):3127-3131. [Medline: 11113947]
27. Mann K. Pharmacotherapy of alcohol dependence: a review of the clinical data. CNS Drugs 2004;18(8):485-504. [Medline: 15182219]
28. Tonigan JS, Toscova R, Miller WR. Meta-analysis of the literature on Alcoholics Anonymous: sample and study characteristics moderate findings. J Stud Alcohol 1996 Jan;57(1):65-72. [Medline: 8747503]
29. Johns B, Baltussen R, Hutubessy R. Programme costs in the economic evaluation of health interventions. Cost Eff Resour Alloc 2003 Feb 26;1(1):1 [FREE Full text] [Medline: 12773220]
30. Stouthard M, Essink-Bot M, Bonsel G, Barendregt J, Kramers P, Van de Water H, et al. Disabilty weights for diseases in the Netherlands. Rotterdam: Department of Public Health, Erasmus University. Rotterdam, the Netherlands: Department of Public Health, Erasmus University; 1997.
31. Murray CJL, Lopez A, Mathers CD, Stein C. The Global Burden of Disease 2000 Project: Aims, Methods and Data Sources. Geneva, Switzerland: World Health Organization; 2001.
32. Kraemer KL, Roberts MS, Horton NJ, Palfai T, Samet JH, Freedner N, et al. Health utility ratings for a spectrum of alcohol-related health states. Med Care 2005 Jun;43(6):541-550. [Medline: 15908848]
33. Smit F, Van Laar M, Croes E, Busschbach J. Ziektelastgewichten voor misbruik, schadelijk gebruik en afhankelijkheid van alcohol en drugs Disability weights for abuse, harmful use and dependence of alcohol and drugs. Utrecht, the Netherlands: Trimbos Institute; 2008.
34. Sanderson K, Andrews G, Corry J, Lapsley H. Using the effect size to model change in preference values from descriptive health status. Qual Life Res 2004 Sep;13(7):1255-1264. [Medline: 15473504]
35. Gmel G, Gutjahr E, Rehm J. How stable is the risk curve between alcohol and all-cause mortality and what factors influence the shape? A precision-weighted hierarchical meta-analysis. Eur J Epidemiol 2003;18(7):631-642. [Medline: 12952136]
36. Barbosa C, Taylor B, Godfrey C, Rehm J, Parrott S, Drummond C. Modelling lifetime QALYs and health care costs from different drinking patterns over time: a Markov model. Int J Methods Psychiatr Res 2010 Jun;19(2):97-109. [doi: 10.1002/mpr.306] [Medline: 20506446]
37. Briggs A, Claxton K, Sculpher MJ. Decision Modelling for Health Economic Evaluation. Oxford, England: Oxford University Press; 2006.
38. Chisholm D, Rehm J, Van Ommeren M, Monteiro M. Reducing the global burden of hazardous alcohol use: a comparative cost-effectiveness analysis. J Stud Alcohol 2004 Nov;65(6):782-793. [Medline: 15700517]

Abbreviations

AA: Alcoholics Anonymous
ALCMOD: alcohol model
AUDIT: Alcohol Use Disorders Identification Test
CBT: cognitive behavioral therapy
DALY: disability adjusted life year
ICER: incremental cost-effectiveness ratio
DW: disability weight
GP: general practitioner
OR: odds ratio
QOL: quality of life
RR: relative risk
YLD: years lost due to disability
YLL: years of life lost

Edited by G Eysenbach; submitted 23.11.10; peer-reviewed by D McDaid; comments to author 20.12.10; revised version received 06.01.11; accepted 14.03.11; published 11.08.11.

Please cite as:
Smit F, Lokkerbol J, Riper H, Majo MC, Boon B, Blankers M
Modeling the Cost-Effectiveness of Health Care Systems for Alcohol Use Disorders: How Implementation of eHealth Interventions Improves Cost-Effectiveness
J Med Internet Res 2011;13(3):e56
URL: http://www.jmir.org/2011/3/e56/
doi:10.2196/jmir.1694
PMID:21840836

Original Paper

Bringing Loyalty to E-health: Theory Validation Using Three Internet-Delivered Interventions

Rik Crutzen[1], PhD; Dianne Cyr[2], PhD; Nanne K de Vries[1], PhD

[1]Maastricht University/CAPHRI, Maastricht, Netherlands
[2]Simon Fraser University, Surrey, BC, Canada

Corresponding Author:
Rik Crutzen, PhD
Maastricht University/CAPHRI
P.O. Box 616
Maastricht, 6200MD
Netherlands
Phone: 31 433882828
Fax: 31 433671032
Email: rik.crutzen@maastrichtuniversity.nl

Abstract

Background: Internet-delivered interventions can effectively change health risk behaviors, but the actual use of these interventions by the target group once they access the website is often very low (high attrition, low adherence). Therefore, it is relevant and necessary to focus on factors related to use of an intervention once people arrive at the intervention website. We focused on user perceptions resulting in e-loyalty (ie, intention to visit an intervention again and to recommend it to others). A background theory for e-loyalty, however, is still lacking for Internet-delivered interventions.

Objective: The objective of our study was to propose and validate a conceptual model regarding user perceptions and e-loyalty within the field of eHealth.

Methods: We presented at random 3 primary prevention interventions aimed at the general public and, subsequently, participants completed validated measures regarding user perceptions and e-loyalty. Time on each intervention website was assessed by means of server registrations.

Results: Of the 592 people who were invited to participate, 397 initiated the study (response rate: 67%) and 351 (48% female, mean age 43 years, varying in educational level) finished the study (retention rate: 88%). Internal consistency of all measures was high (Cronbach alpha > .87). The findings demonstrate that the user perceptions regarding effectiveness ($\text{beta}_{\text{range}}$ .21–.41) and enjoyment ($\text{beta}_{\text{range}}$ .14–.24) both had a positive effect on e-loyalty, which was mediated by active trust ($\text{beta}_{\text{range}}$ .27–.60). User perceptions and e-loyalty had low correlations with time on the website (r_{range} .04–.18).

Conclusions: The consistent pattern of findings speaks in favor of their robustness and contributes to theory validation regarding e-loyalty. The importance of a theory-driven solution to a practice-based problem (ie, low actual use) needs to be stressed in view of the importance of the Internet in terms of intervention development. Longitudinal studies are needed to investigate whether people will actually revisit intervention websites and whether this leads to changes in health risk behaviors.

(J Med Internet Res 2011;13(3):e73) doi:10.2196/jmir.1837

KEYWORDS

e-Loyalty; adherence; attrition; user perceptions; theory; Internet; interventions

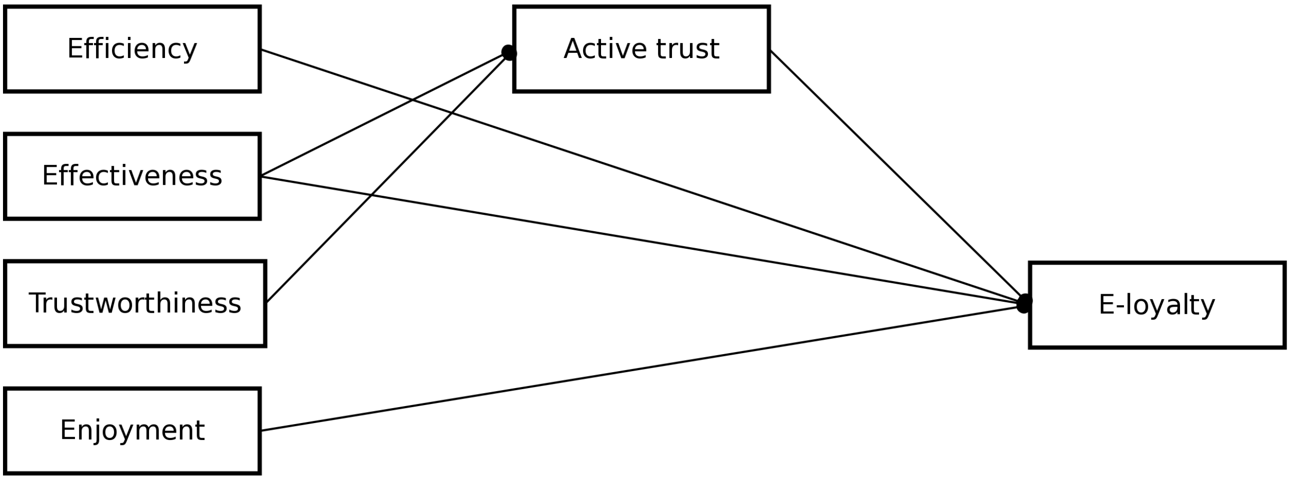
Introduction

Internet-delivered interventions can effectively change health risk behaviors (eg, lack of physical activity, low consumption of fruit, cigarette smoking, and excessive alcohol consumption) [1]. However, the actual use of these interventions by the target group once they access the website is very low [2,3]. For example, server statistics of an intervention promoting heart-healthy behaviors showed that 285,146 visitors from unique internet protocol (IP) addresses landed on the home page in a 36-month period, but 56.3% of these left the intervention website within 30 seconds [4]. This finding touches on the critical issue in Internet-delivered interventions: how can behavior ever be changed if people are not exposed or are only briefly exposed to the actual intervention? Therefore, it is relevant and necessary to focus on factors related to use of an intervention once people arrive at the intervention website. These factors relate to the *visitor* (eg, people's motivation to be healthy [5,6]) as well as the *intervention* website (eg, visual complexity of the homepage). Two recently published systematic reviews provide a detailed overview of factors used by current *interventions* to stimulate use of intervention websites [7,8]. Our study focused on perceptions of *visitors* resulting in a user experience [9,10].

User experience refers to what a person thinks and feels during and after exposure to a website [11]. The main idea is that a positive user experience leads to increased website use. User experience consists of cognitive and affective perceptions [12]. Cognitive perceptions are rational in nature and are induced by utilitarian or cognitive motives. Affective perceptions are emotional in nature and are induced by hedonic or affective motives [13]. Previous studies demonstrated the importance of these perceptions regarding intention to use a technology [14] and to visit a website again [12]. We designed our study on the basis of these findings and applied them to loyalty regarding intervention websites in the field of eHealth (ie, e-loyalty). Besides visiting an Internet-delivered intervention again, e-loyalty also consists of recommending an Internet-delivered intervention to others. The latter is based on previous research indicating that word-of-mouth is an effective strategy to improve use of Internet-delivered interventions [15,16]. A background theory for e-loyalty is still lacking for Internet-delivered interventions. Although previous studies did explicitly describe the theory used to develop the content of Internet-delivered interventions, these theories primarily related to behavior determinants or behavior change [8,17]. Theory development regarding e-loyalty is highly needed to increase the public health impact of Internet-delivered interventions. Therefore, in this study we propose and validate a conceptual model.

To systematically constitute the proposed conceptual model, we describe conceptual definitions and their relationship with e-loyalty [18]. Terminology that is used within the conceptual model (ie, key user perceptions) is derived from other fields such as e-commerce. Although these terms can have a different meaning within public health, we chose to use the same terminology as in previous studies in other fields to avoid further confusion. The key user perceptions in the conceptual model are efficiency, effectiveness, trustworthiness, enjoyment, and active trust [12]. *Efficiency* refers to easy search of and access to the information provided, and *effectiveness* refers to the quality of that information (eg, in terms of relevance) [19]. These cognitive perceptions have parallels with perceived ease of use and perceived usefulness in the technology acceptance model, but are applicable in a broader context [20]. The positive effect of these cognitive perceptions on e-loyalty has been demonstrated in, for example, e-service environments [21]. The idea that a positive user experience leads to e-loyalty applies not only to cognitive perceptions, but also to affective perceptions [5,22]. These affective perceptions are often referred to as *enjoyment* [23] and have been demonstrated to have a positive effect on e-loyalty in, for example, e-commerce [24]. *Trustworthiness* is defined as the believability of the provided information and refers to both cognitive and affective perceptions: it is based on a cognitive process (eg, rational reasons) and an emotional base (eg, a strong positive affect for the trustee) [25]. It has been demonstrated to have a positive effect on e-loyalty in, for example, online shopping [26,27]. *Active trust* might be a working mechanism leading to e-loyalty [28]. Whereas trustworthiness refers to the believability (eg, "I trust the information presented on this website"), active trust refers to the confidence in acting on the provided information (eg, "I would act on the information presented on this website"). Active trust has been proven to be the primary intermediate associated with e-loyalty [26,29]. In line with the study of Cugelman and colleagues [28], we expected active trust to mediate the impact of trustworthiness and effectiveness on e-loyalty. This resulted in the following hypotheses to be tested in a new context lacking a background theory: the field of eHealth (Figure 1).

Figure 1. Conceptual model.



Hypothesis 1

H1a Efficiency has a positive effect on e-loyalty.

H1b Effectiveness has a positive effect on e-loyalty.

H1c Enjoyment has a positive effect on e-loyalty.

Hypothesis 2

> *H2a Active trust mediates the relationship between effectiveness and e-loyalty.*
>
> *H2b Active trust mediates the relationship between trustworthiness and e-loyalty.*

Methods

To improve the external and ecological validity, we included 3 generally available, Internet-delivered interventions. The interventions were certified according to the guidelines of the Dutch recognition system for health promotion interventions [30]. The quality assessment of health promotion interventions is supervised by the Netherlands Institute for Public Health and the Environment (interventions aimed at adults) and the Netherlands Youth Institute (interventions aimed at youth) [31]. We included interventions from all levels of recognition (theoretically sound, probable effectiveness, and established effectiveness; inspired by the UK Medical Research Council's evaluation framework for complex interventions) in this study. The first intervention, registered by the Consumer and Safety Foundation (Netherlands), was theoretically sound and is concerned with prevention of sports injuries (intervention 1 [32]). The second intervention, registered by the Netherlands Institute of Mental Health and Addiction, was probably effective and is concerned with drinking less alcohol (intervention 2 [33]). The third intervention, registered by the Netherlands Institute of Mental Health and Addiction, was effective and is intended for people feeling gloomy or having mild depressive complaints (intervention 3 [34]). We must stress that these were all primary prevention interventions aimed at the general public. In other words, these interventions were not targeted at diagnosing (secondary prevention) or treating (tertiary prevention) health problems related to health risk behaviors, but at people who did not yet have these problems. Hence, these interventions were deemed of interest to the general public.

Participants

Participants were recruited through a research panel of a Dutch Internet research agency [35]. From this panel, we invited through email a stratified sample of 592 potential participants to take part in this study. This sample was representative of the Dutch population above 18 years, taking into account gender, age, and level of education. Of those invited, 397 clicked on the link in the invitation email to start the study (response rate: 67%) and 351 finished the study (retention rate: 88%). There was no selective dropout regarding gender ($n = 592$, $\chi^2_1 = 3.2$, $P = .08$), but those who dropped out were somewhat younger (40 vs 43 years, $t_{590} = 2.86$, $P = .004$) and differed in terms of level of education ($n = 592$, $\chi^2_1 = 10.9$, $P = .004$). The final sample consisted of 48% (169/351) women; the average age was 43 (SD 13) years. In terms of level of education, 30% (107/351) of the participants had a low level of highest completed education, 35% (122/351) an intermediate level, and 35% (122/351) a high level (according to the definitions of Statistics Netherlands).

Procedure

The study consisted of 3 blocks (ie, 3 intervention websites and related measurements) that were presented at random to each participant. In each block participants were exposed to 1 of the 3 intervention websites described above, and subsequently participants completed the measures described in the measurements section. Participants were asked to assess several websites. It was stressed that there were no right or wrong answers and they could base their opinion on their first impression. The reason behind this was to prevent participants from thoroughly studying the intervention website, and to mimic a real-life situation in which the time being exposed to and willing to investigate an intervention website is often limited [4]. On average, participants took 17 minutes to complete the full study (eg, exploring the intervention websites and completing related measurements). Participants received credit points for participating in the study, for a value of €1.95.

Measurements

Directly after exposure to each intervention website, participants indicated whether they had seen the website before. For each intervention website, data from participants who indicated that they had seen website before were removed, because their perceptions and loyalty might have been based on the previous exposure to the intervention website. This concerned 8 (different) participants per intervention website and results did not differ if their data were included. Subsequently, participants completed the following validated measures after being exposed to each intervention website.

E-loyalty: intention to visit the website again (eg, “It is likely that I will visit the website again in the future”) and whether participants would recommend the website to others (“It is likely that I will recommend this website to others ”) were assessed by 3 items each [36]. Items were answered on a 7-point Likert scale ranging from “strongly agree” to “strongly disagree.”

User perceptions: efficiency (eg, “I was able to access the information quickly on this website”), effectiveness (eg, “The website provided me with relevant information about...”), trustworthiness (eg, “I trust the information presented on this website”), enjoyment (eg, “I found my visit to this website enjoyable”), and active trust (eg, “I would act on the information presented on this website if needed”) were assessed by 3 items each [21,37]. Items were answered on a 7-point Likert scale ranging from “strongly agree” to “strongly disagree.”

Two native speakers (RC and an assistant) translated all items into Dutch and discussed semantic similarity until reaching a consensus. Besides these self-reported measures, time on each intervention website was assessed by means of server registrations.

Analyses

First, using Predictive Analytics SoftWare Statistics (version 18.0; IBM Corporation, Somers, NY, USA), we conducted correlation and reliability analyses for each intervention website separately. Subsequently, using Mplus (version 5; Muthén & Muthén, Los Angeles, CA, USA), we constructed structural equation models to test the hypotheses per intervention website.

First of all, we tested the hypothesized conceptual model: intention to visit again and recommending to others were regressed on efficiency, effectiveness, and enjoyment; active trust was regressed on effectiveness and trustworthiness. Subsequently, we added paths to the conceptual model based on modification indices, which are chi-square distributed, implying that a modification index larger than 3.84 indicates that adding the suggested path will significantly improve model fit. The reason to include paths beyond the hypotheses was to explore whether unanticipated relationships might explain variance in e-loyalty and, hence, contribute to theory development. The criterion for accepting or rejecting a hypothesis was a significant pattern across all 3 models. A level of significance of .05 was used for the relationships within the model.

Comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) were used as fit indices for each model. CFI and TLI are goodness-of-fit indices, where larger values signal better fit. Values over .95 indicate close fit. RMSEA and SRMR are goodness-of-fit indices, where larger values signal worse fit. Rules of thumb for close fit are RMSEA ≤ .05 and SRMR ≤ .09 [38,39].

Results

Tables 1–3 show the results of correlation and reliability analyses. Internal consistency of all measures was high (Cronbach alpha > .87). Overall, correlations between user perceptions and e-loyalty were high (r_{range} .44–.84). User perceptions and e-loyalty have low correlations with time on the website (r_{range} .04–.18).

Table 1. Correlation matrix intervention 1 (N = 343)

	Alpha	Mean	SD	1	2	3	4	5	6	7	8
1. Efficiency	.98	4.7	1.6	–	.68	.72	.71	.61	.58	.60	.04
2. Effectiveness	.95	4.4	1.6		–	.79	.77	.82	.75	.76	.13
3. Trustworthiness	.97	4.7	1.4			–	.73	.72	.64	.66	.16
4. Enjoyment	.99	4.1	1.7				–	.79	.74	.75	.09
5. Active trust	.94	4.2	1.7					–	.76	.76	.10
6. Intention to visit again	.89	3.6	1.7						–	.84	.09
7. Recommend to others	.95	4.0	1.7							–	.14
8. Time on website (minutes)	–	3:06									–

Table 2. Correlation matrix intervention 2 (N = 343)

	Alpha	Mean	SD	1	2	3	4	5	6	7	8
1. Efficiency	.97	4.8	1.5	–	.62	.63	.60	.57	.49	.58	.12
2. Effectiveness	.91	4.3	1.5		–	.71	.67	.76	.63	.70	.16
3. Trustworthiness	.96	4.6	1.4			–	.66	.68	.50	.61	.14
4. Enjoyment	.98	4.0	1.5				–	.71	.62	.67	.09
5. Active trust	.91	4.1	1.6					–	.71	.74	.16
6. Intention to visit again	.87	3.3	1.7						–	.77	.18
7. Recommend to others	.94	4.0	1.7							–	.16
8. Time on website (minutes)	–	1:28									–

Table 3. Correlation matrix intervention 3 (N = 343)

	Alpha	Mean	SD	1	2	3	4	5	6	7	8
1. Efficiency	.98	4.7	1.5	–	.63	.61	.57	.57	.44	.57	.09
2. Effectiveness	.95	3.9	1.7		–	.73	.72	.77	.72	.78	.14
3. Trustworthiness	.97	4.4	1.4			–	.70	.70	.56	.69	.07
4. Enjoyment	.98	4.0	1.6				–	.79	.70	.75	.11
5. Active trust	.94	3.8	1.7					–	.74	77	.09
6. Intention to visit again	.91	3.3	1.7						–	.82	.10
7. Recommend to others	.96	3.7	1.8							–	.14
8. Time on website (minutes)	–	3:10									–

Table 4 shows the results of the structural equation models when testing the conceptual model. H1a was rejected; efficiency did not have a positive effect on e-loyalty. H1b and H1c were confirmed; both effectiveness and enjoyment had a positive effect on e-loyalty. H2a was also confirmed; active trust mediated the relationship between effectiveness and e-loyalty. Results for H2b were mixed, because the relationship between trustworthiness and active trust differed in terms of being significant and standardized betas [40,41]. Therefore, the relationship between trustworthiness and active trust was included when adding paths to the conceptual model based on modification indices. The only path that was added to the conceptual model was the relationship between enjoyment and active trust. Modification indices (respective values of 50.27, 39.15, and 72.62) suggested the addition of this path to each model representing an intervention website. Table 5 shows the results of the structural equation models when testing this extended model. The results were similar to the conceptual model: H1a was rejected and H1b, H1c, and H2a were confirmed. H2b, however, was rejected; active trust did not mediate the relationship between trustworthiness and e-loyalty, because there was no relationship between trustworthiness and active trust. Unanticipatedly, but consistently, the positive effect of enjoyment was mediated by active trust. All fit indices indicated good fit for the extended model. Figure 2 shows the extended model resulting from the analyses for all 3 intervention websites.

Table 4. Results of conceptual model (figures are standardized betas of paths within the model)

	Intervention					
	1		2		3	
Path[a]	INT[b]	REC[c]	INT	REC	INT	REC
EFI → e-loyalty	ns[d]	ns	ns	ns	–.16	ns
EFE → e-loyalty	.28	.25	ns	.22	.42	.43
ENJ → e-loyalty	.26	.29	.19	.21	.20	.25
ACT → e-loyalty	.39	.33	.58	.40	.36	.26
EFE → ACT	.81		.71		.66	
TRU → ACT	ns		.16		.21	
R^2	.73	.68	.56	.63	.65	.70
CFI[e]	.95		.95		.95	
TLI[f]	.94		.94		.94	
RMSEA[g]	.10		.09		.09	
SRMR[h]	.05		.06		.06	

[a] EFI = efficiency; EFE = effectiveness; TRU = trustworthiness; ENJ = enjoyment; ACT = active trust.

[b] Intention to visit again.

[c] Recommend to others.

[d] Not significant; all other paths are significant at the $P = .05$ level.

[e] Comparative fit index.

[f] Tucker-Lewis index.

[g] Root mean square error of approximation.

[h] Standardized root mean square residual.

Table 5. Results of extended model (figures are standardized betas of paths within the model)

	Intervention					
	1		2		3	
Path[a]	INT[b]	REC[c]	INT	REC	INT	REC
EFI → e-loyalty	ns[d]	ns	ns	ns	–.15	ns
EFE → e-loyalty	.27	.24	ns	.21	.40	.41
ENJ → e-loyalty	.23	.27	.14	.18	.17	.24
ACT → e-loyalty	.41	.35	.60	.42	.39	.27
EFE → ACT	.57		.51		.40	
TRU → ACT	ns		ns		ns	
ENJ → ACT	.37		.34		.47	
R^2	.73	.68	.57	.64	.65	.70
CFI[e]	.95		.96		.96	
TLI[f]	.94		.95		.95	
RMSEA[g]	.09		.09		.09	
SRMR[h]	.05		.05		.05	

[a] EFI = efficiency; EFE = effectiveness; TRU = trustworthiness; ENJ = enjoyment; ACT = active trust.

[b] Intention to visit again.

[c] Recommend to others.

[d] Not significant; all other paths are significant at the $P = .05$ level.

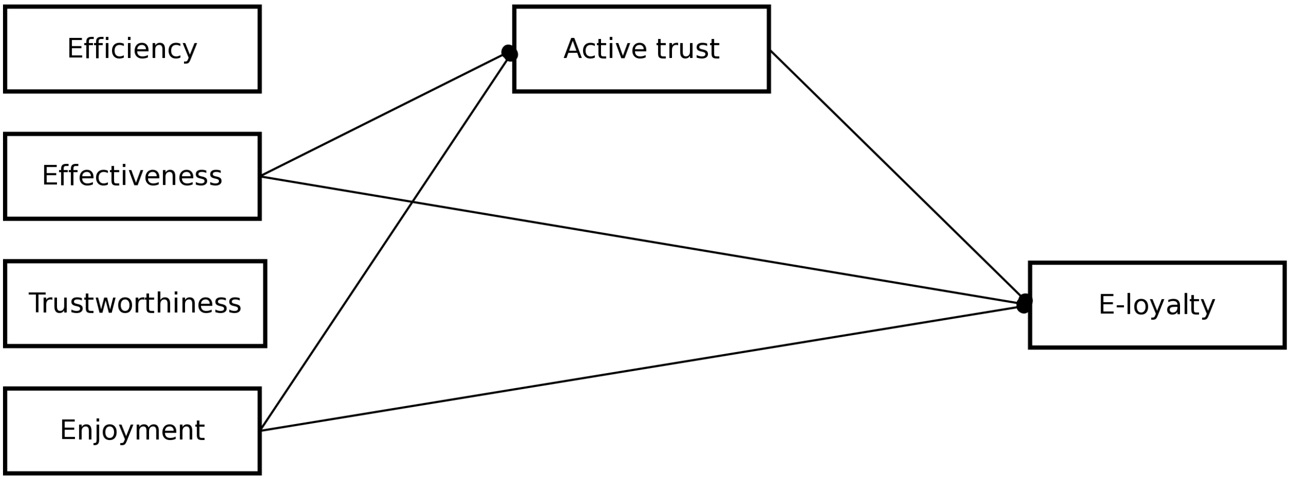
[e] Comparative fit index.

[f] Tucker-Lewis index.

[g] Root mean square error of approximation.

[h] Standardized root mean square residual.

Figure 2. Extended Model.



Discussion

Our findings consistently demonstrate that effectiveness and enjoyment both had a positive effect on e-loyalty, which was mediated by active trust. The findings regarding effectiveness were anticipated and in line with previous research [28]. Mediation of the positive effect of enjoyment by active trust, however, was unanticipated. An explanation can be based on previous research demonstrating that enjoyment is related to cognitive perceptions [42]. Thus, cognitive perceptions might be a working mechanism for the positive effect of enjoyment on e-loyalty. Future research is needed to shed more light on the plausibility of this explanation, since this relationship can also be reversed: affective perceptions as a working mechanism leading to e-loyalty [43].

Rejection of the hypothesis regarding the positive effect of efficiency on e-loyalty can be explained by the procedure used in this study. Efficiency refers to easy search of and access to the information provided. Participants, however, were not necessarily looking for information regarding the topic of the intervention websites to which they were exposed. Although participants could fill out the items regarding efficiency based on *whether* the intervention website at hand would be easy to search and access *if* they were looking for information at that intervention website, the lack of a need for information might explain the absence of evidence for a positive effect of efficiency. This could be solved by giving participants an assignment for which they have to study the intervention website thoroughly. The reason why we did not do this in the current study was to mimic a real-life situation in which people might review an intervention website when time limitations prevail [4]. This was reflected in this study as well, given the average time on website (range 1:28–3:10 minutes).

The lack of a relationship between trustworthiness and active trust in the structural equation models is puzzling. It might be that active trust by itself captures all the variance in e-loyalty that could be explained by trustworthiness. Since previous research demonstrated that active trust is the primary intermediate associated with e-loyalty [26,29], it might be that active trust reduces the possible impact of trustworthiness. This is contrary to previous research [28], however, and still does not explain the absence of a relationship between trustworthiness and active trust.

A final finding that deserves attention is that user perceptions and e-loyalty had low correlation with time on website. This can be explained by a confirmation bias [44]: since participants were told that they had to assess several websites, they might have been looking for evidence in line with their first impression, regardless of whether their impression was negative or positive. So, in the current setting the time spent on an intervention website is independent of user perceptions. Time on website may be related to user perceptions and e-loyalty if people explore an intervention website without any instructions.

In sum, although not all hypotheses were confirmed, this study clearly demonstrates that user perceptions (ie, effectiveness, enjoyment, and active trust) regarding e-loyalty are not important just in fields such as e-commerce, but also in the context of eHealth. The next question is how to improve user perceptions of intervention websites. To answer this question, characteristics of intervention websites need to be systematically manipulated, and these manipulations should be linked to user perceptions, and subsequently to e-loyalty. A possible variable to be manipulated is user control, defined as the voluntary and instrumental actions of a website visitor that influence the user experience [45,46]. The ability to control information flow increases one’s ability to explore and understand the structure of a website [47]. Nevertheless, one of the most common issues faced by visitors of websites is lack of user control [48]. This is awkward, given the wealth of literature (eg, McMillan and Hwang provide an overview [49]) documenting the importance of user control in shaping user experience [45,50]. Furthermore, previous research identified the role of user control (ie, freedom of choice) in attitude change [51] and intention to use [12,14]. The effect of user control on e-loyalty is in line with previous studies and is expected to be mediated through user perceptions [12,52]. Another characteristic to be manipulated in future research might be the use of tailoring strategies (eg, personalization, feedback) that have been shown to have a positive effect on intervention outcomes (in terms of health behaviors), which is related to intervention use [53].

Acknowledgments

This study was supported by the Innovational Research Incentives Scheme Veni from NWO-MaGW (Netherlands Organisation for Scientific Research - Division for the Social Sciences) accredited to the first author. There has been no involvement by the funding body in the review or approval of the manuscript for publication. The authors would like to thank Nienke Beekers for her help in translating items into Dutch and the members of “het voedingsoverleg” for their valuable comments on an earlier version of the manuscript.

Conflicts of Interest

None declared

References

1. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16. [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
2. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009 Apr 29;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
3. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
4. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595. [doi: 10.1093/her/cyp063] [Medline: 19897515]
5. Van 't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: a longitudinal study. J Med Internet Res 2010;12(3):e37 [FREE Full text] [doi: 10.2196/jmir.1345] [Medline: 20813716]

6. Crutzen R, de Nooijer J, Candel MJ, de Vries NK. Adolescents who intend to change multiple health behaviours choose greater exposure to an internet-delivered intervention. J Health Psychol 2008 Oct;13(7):906-911. [doi: 10.1177/1359105308095064] [Medline: 18809641]
7. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
8. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
9. Van Oppen CAML. From Rags to Richness. Maastricht: Maastricht University; 2007.
10. Sundar SS. Theorizing interactivity's effects. Inf Soc 2004;20:385-389. [doi: 10.1080/01972240490508072]
11. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. A conceptual framework for understanding and improving adolescents' exposure to Internet-delivered interventions. Health Promot Int 2009 Sep;24(3):277-284 [FREE Full text] [doi: 10.1093/heapro/dap018] [Medline: 19515716]
12. Cyr D, Head M, Ivanov A. Perceived interactivity leading to e-loyalty: development of a model for cognitive-affective user responses. Int J Hum Comput Stud 2009;67:850-869. [doi: 10.1016/j.ijhcs.2009.07.004]
13. Park CW, Young SM. Consumer response to television commercials: the impact of involvement and background music on brand attitude formation. J Marketing Res 1986;23:11-24.
14. Lee T. The impact of perceptions of interactivity on customer trust and transaction intentions in mobile commerce. J Electron Commer Res 2005;6(3):165-180.
15. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries N. Effectiveness of online word of mouth on exposure to an Internet-delivered intervention. Psychol Health 2009 Jul;24(6):651-661. [doi: 10.1080/08870440802521094] [Medline: 20205018]
16. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Internet-delivered interventions aimed at adolescents: a Delphi study on dissemination and exposure. Health Educ Res 2008 Jun;23(3):427-439 [FREE Full text] [doi: 10.1093/her/cym094] [Medline: 18209115]
17. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
18. Wacker JG. A definition of theory: research guidelines for different theory-building research methods in operations management. Journal of Operations Management 1998;16:361-385. [doi: 10.1016/S0272-6963(98)00019-9]
19. Keeney R. The value of Internet commerce to the customer. Manag Sci 1999;45:533-542. [doi: 10.1287/mnsc.45.4.533]
20. Benbasat I, Barki H. Quo vadis, TAM? J Assoc Inf Syst 2007;8(4):211-218.
21. Cyr D, Hassanein K, Head M, Ivanov A. The role of social presence in establishing loyalty in e-service environments. Interact Comput 2007;19:43-56. [doi: 10.1016/j.intcom.2006.07.010]
22. Sun Y, Wang N, Peng Z. Working for one penny: understanding why people would like to participate in online tasks with low payment. Comput Hum Behav 2011;27:1033-1041. [doi: 10.1016/j.chb.2010.12.007]
23. Van der Heijden H. Factors influencing the usage of websites: the case of a generic portal in the Netherlands. Inf Manag 2003;40:541-549. [doi: 10.1016/S0378-7206(02)00079-4]
24. Cyr D, Head M, Ivanov A. Design aesthetics leading to m-loyalty in mobile commerce. Inf Manag 2006;43:950-963. [doi: 10.1016/j.im.2006.08.009]
25. Komiak S, Benbasat I. Understanding customer trust in agent-mediated electronic commerce, web-mediated electronic commerce, and traditional commerce. Inf Technol Manag 2004;5:181-207. [doi: 10.1023/B:ITEM.0000008081.55563.d4]
26. Flavián C, Guinalíu M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Inf Manag 2006;43:1-14. [doi: 10.1016/j.im.2005.01.002]
27. Gefen D, Karahanna E, Straub D. Trust and TAM in online shopping: an integrated model. MIS Q 2003;27(1):51-90.
28. Cugelman B, Thelwall M, Dawes P. The dimensions of web site credibility and their relation to active trust and behavioural impact. Commun Assoc Inf Syst 2009;24:455-472.
29. Bart Y, Shankar V, Fareena S, Urban G. Are the drivers and roles of online trust the same for all web sites and consumers? A large-scale exploratory empirical study. J Marketing 2005;69:133-152. [doi: 10.1509/jmkg.2005.69.4.133]
30. National Institute for Public Health and the Environment (Netherlands). Centre for Healthy Living (Centrum Gezond Leven). 2011. I-database URL: http://www.loketgezondleven.nl/interventies/i-database [accessed 2011-04-22] [WebCite Cache ID 5y8KAjgh2]
31. Brug J, van Dale D, Lanting L, Kremers S, Veenhof C, Leurs M, et al. Towards evidence-based, quality-controlled health promotion: the Dutch recognition system for health promotion interventions. Health Educ Res 2010 Dec;25(6):1100-1106 [FREE Full text] [doi: 10.1093/her/cyq046] [Medline: 20841318]
32. Voorkom Blessures met Gezondverstand. Stichting Consument en Veiligheid. 2011 URL: http://www.voorkomblessures.nl/csi/websitesportblessure.nsf/ [accessed 2011-04-22] [WebCite Cache ID 5y8K3hUVI]
33. Netherlands Institute of Mental Health and Addiction. Trimbos Instituut. 2011. MinderDrinken URL: http://www.minderdrinken.nl/ [accessed 2011-04-22] [WebCite Cache ID 5y8KJYxhZ]

34. Netherlands Institute of Mental Health and Addiction. Mentalshare. 2011. 'Kleur je leven' Online Zelfhulpcursus URL: https://www.kleurjeleven.nl/ [accessed 2011-04-22] [WebCite Cache ID 5y8KNJQ7U]
35. FlyCatcher. FlyCatcher Internet Research. 2011 URL: http://www.flycatcher.eu/ [accessed 2011-04-22] [WebCite Cache ID 5y8KQdYbW]
36. Cyr D, Bonanni C, Bowes J, Ilsever J. Beyond trust: website design preferences across cultures. J Glob Inf Manag 2005;13:25-52. [doi: 10.4018/jgim.2005100102]
37. Teo H, Oh L, Liu C, Wei K. An empirical study of the effects of interactivity on web user attitude. Int J Hum Comput Stud 2003;58:281-305. [doi: 10.1016/S1071-5819(03)00008-9]
38. Iacobucci D. Structural equations modeling: fit indices, sample size, and advanced topics. J Consum Psychol 2010;20:90-98. [doi: 10.1016/j.jcps.2009.09.003]
39. Kline RB. Principles and Practice of Structural Equation Modeling. New York, NY: Guilford Press; 2005.
40. Crutzen R. A systematic review on computer-based education for patients with hypertension: what about effect sizes? Health Educ J 2010;69:365-366. [doi: 10.1177/0017896910376870]
41. Crutzen R. Adding effect sizes to a systematic review on interventions for promoting physical activity among European teenagers. Int J Behav Nutr Phys Act 2010;7:29 [FREE Full text] [doi: 10.1186/1479-5868-7-29] [Medline: 20398361]
42. Kim J, Fiore AM. Lee H-H. Influences of online store perception, shopping enjoyment, and shopping involvement on consumer patronage behavior towards an online retailer. J Retailing Consum Serv 2007;14:95-107. [doi: 10.1016/j.jretconser.2006.05.001]
43. De Wulf K, Schillewaert N, Muylle S, Rangarajan D. The role of pleasure in web site success. Inf Manag 2006;43:434-446. [doi: 10.1016/j.im.2005.10.005]
44. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 1998;2:175-220. [doi: 10.1037/1089-2680.2.2.175]
45. Liu Y, Schrum LJ. What is interactivity and is it always such a good thing? Implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. J Advertising 2002;31:53-64.
46. Lowry PB, Spaulding T, Wells T, Moody G, Moffit K, Madariaga S. A theoretical model and empirical results linking website interactivity and usability satisfaction. Syst Sci 2006;123. [doi: 10.1109/HICSS.2006.33]
47. Ariely D. Controlling the information flow: effects on consumers' decision making and preferences. J Consum Res 2000;27:233-248. [doi: 10.1086/314322]
48. Srinivasan SS, Anderson R, Ponnavolu K. Customer loyalty in e-commerce: an exploration of its antecedents and consequences. J Retailing 2002;78:41. [doi: 10.1016/S0022-4359(01)00065-3]
49. McMillan SJ. Hwang J-S. Measures of perceived interactivity: an exploration of communication, user control, and time in shaping perceptions of interactivity. J Advertising 2002;31:29-42.
50. Liu Y, Schrum LJ. A dual-process model of interactivity effects. J Advertising 2009;38:53-68.
51. Chiou WB. Induced attitude change on online gaming among adolescents: an application of the less-leads-to-more effect. Cyberpsychol Behav 2008 Apr;11(2):212-216. [doi: 10.1089/cpb.2007.0035] [Medline: 18422416]
52. Hassenzahl M, Tractinsky N. User experience: a research agenda. Behav Inf Technol 2006;25:91-97. [doi: 10.1080/01449290500330331]
53. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]

Edited by G Eysenbach; submitted 22.04.11; peer-reviewed by J Brown, L Yardley; comments to author 12.05.11; revised version received 25.05.11; accepted 26.05.11; published 24.09.11.

Please cite as:
Crutzen R, Cyr D, de Vries NK
Bringing Loyalty to E-health: Theory Validation Using Three Internet-Delivered Interventions
J Med Internet Res 2011;13(3):e73
URL: http://www.jmir.org/2011/3/e73/
doi:10.2196/jmir.1837
PMID:21946128

Original Paper

Factors Influencing the Use of a Web-Based Application for Supporting the Self-Care of Patients with Type 2 Diabetes: A Longitudinal Study

Nicol Nijland[1], PhD; Julia EWC van Gemert-Pijnen[1], PhD; Saskia M Kelders[1], MSc; Bart J Brandenburg[1,2], MD; Erwin R Seydel[1], PhD

[1]Department of Psychology, Health and Technology/Center for eHealth Research and Disease Management, Faculty of Behavioural Sciences, University of Twente, Enschede, Netherlands

[2]Medicinfo, Tilburg, Netherlands

Corresponding Author:
Nicol Nijland, PhD
Department of Psychology, Health and Technology/Center for eHealth Research and Disease Management
Faculty of Behavioural Sciences
University of Twente
Drienerlolaan 5
PO Box 217
Enschede, 7500 AE
Netherlands
Phone: 31 651071559
Fax: 31 534892388
Email: n.nijland@utwente.nl

Abstract

Background: The take-up of eHealth applications in general is still rather low and user attrition is often high. Only limited information is available about the use of eHealth technologies among specific patient groups.

Objective: The aim of this study was to explore the factors that influence the initial and long-term use of a Web-based application (DiabetesCoach) for supporting the self-care of patients with type 2 diabetes.

Methods: A mixed-methods research design was used for a process analysis of the actual usage of the Web application over a 2-year period and to identify user profiles. Research instruments included log files, interviews, usability tests, and a survey.

Results: The DiabetesCoach was predominantly used for interactive features like online monitoring, personal data, and patient–nurse email contact. It was the continuous, personal feedback that particularly appealed to the patients; they felt more closely monitored by their nurse and encouraged to play a more active role in self-managing their disease. Despite the positive outcomes, usage of the Web application was hindered by low enrollment and nonusage attrition. The main barrier to enrollment had to do with a lack of access to the Internet (146/226, 65%). Although 68% (34/50) of the enrollees were continuous users, of whom 32% (16/50) could be defined as hardcore users (highly active), the remaining 32% (16/50) did not continue using the Web application for the full duration of the study period. Barriers to long-term use were primarily due to poor user-friendliness of the Web application (the absence of “push” factors or reminders) and selection of the “wrong” users; the well-regulated patients were not the ones who could benefit the most from system use because of a ceiling effect. Patients with a greater need for care seemed to be more engaged in long-term use; highly active users were significantly more often medication users than low/inactive users ($P = .005$) and had a longer diabetes duration ($P = .03$).

Conclusion: Innovations in health care will diffuse more rapidly when technology is employed that is simple to use and has applicable components for interactivity. This would foresee the patients’ need for continuous and personalized feedback, in particular for patients with a greater need for care. From this study several factors appear to influence increased use of eHealth technologies: (1) avoiding selective enrollment, (2) making use of participatory design methods, and (3) developing push factors for persistence. Further research should focus on the causal relationship between using the system’s features and actual usage, as such a view would provide important evidence on how specific technology features can engage and captivate users.

(J Med Internet Res 2011;13(3):e71) doi:10.2196/jmir.1603

KEYWORDS

Internet; technology; eHealth; email; communication; primary care; self-care; diabetes

Introduction

The prevalence of diabetes is rising quickly. Diabetes among adults—aged 20–79 years—affected 285 million adults in 2010 (6.4%) and is estimated to increase worldwide to 439 million adults by 2030 (7.7%) [1]. Between 2010 and 2030, the number of adults with diabetes will increase 69% in developing countries and by 20% in developed countries. Most people with diabetes fall within the 60- to 79-year-old age group, and approximately 90% have diabetes mellitus type 2. Improving diabetes care management has therefore become a priority for health care facilities and patients' organizations worldwide. The ultimate goal of diabetes care management is to optimize self-care in order to reduce mortality, morbidity, and health care costs [2,3].

The introduction of the Internet into clinical practice has brought about many opportunities for self-care [2-7], as it can be used as a powerful medium for promoting a healthy lifestyle and for increasing understanding about the condition. However, to be effective in empowering patients' self-awareness and engagement, Web applications should be designed to allow individuals to tailor the program to their own specific needs, because patients are increasingly demanding convenient access to a high level of personalized health care [8,9]. To promote self-care, interactive eHealth applications have been developed for continuous self-monitoring, feedback, and information exchange.

From previous studies we know that interactive eHealth technologies contribute positively to health care for patients with a chronic illness, realizing increased patient–provider communication, positive impact on metabolic control and behavior change, improved therapy adherence, and cost reductions [6,7,10-14]. However, the uptake of eHealth in general is still rather low [15,16]. Therefore, more research should be directed toward the factors that provide insights into the actual usage and the accompanying reasons for use and nonuse of eHealth technologies.

Expanding the uptake of eHealth requires, first and foremost, a better understanding of the obstacles that prevent access (initial use) [15-19] and, secondly, a better understanding of the factors that influence the long-term use of eHealth technologies [20-23], since many projects still fail to survive beyond the pilot phase, and user attrition is a typical problem ("Law of Attrition" [22]). To this end, we performed a longitudinal study.

The aim of the study was to explore the factors that influenced the initial and long-term use of a Web-based application for supporting the self-care of patients with type 2 diabetes. A mixed-methods research design was applied to trace the usage over time (log files), along with the reasons for (non)usage (usability tests, interviews, and content analysis of email messages), and to identify user profiles (survey).

Methods

Description of the Web Application

DiabetesCoach, a Web-based application for supporting self-care among patients with type 2 diabetes, was developed to encourage patients to play a more active role in their own care. The Web application is a low-tech solution for a large group of patients and was provided free of charge as a supplement to regular diabetes care. The application was developed by Medicinfo (Tilburg, Netherlands) in close collaboration with general practitioners, nurses, patients, behavioral scientists, and vendors (ie, health insurance companies). Initial development costs were relatively limited, and the running costs of the application were low. Therefore, a rise in use would not lead to an exponential rise in costs.

The following are the core features of the DiabetesCoach:

- *My personal data*: documentation of personal details (eg, treatment plan, medication use).
- *Online monitoring*: registration of metabolic values: weight, blood glucose level, blood pressure, and cholesterol.
- *Email contact*: secured possibility for patient–nurse email communication (response within 5 working days).
- *Online education*: diabetes information and instructions.
- *Calendar*: a place to write down comments, appointments, and personal goals.
- *Lifestyle coach*: self-tests to support lifestyle changes.

The patients' self-monitored data were made available to the nurses with alerts signaling alarming metabolic values. Each nurse had access to each of her own patients' DiabetesCoach details via a private account (protected by username and password). The Web application (not integrated with the nurse's medical record) enabled nurses to set individual goals for their patients, add selected lifestyle programs, and highlights the appropriate chapter of the e-learning program. The patients received no particular instructions with regard to how often they should log on to DiabetesCoach. Patients measured metabolic values both at home and at the primary care practice during office visits. Nurses were allowed to have two extra consultation sessions per patient to compensate for the extra time needed to participate in the study. The information and guidelines provided in DiabetesCoach were in accordance with diabetes care standards and protocols in the Netherlands.

Participant Recruitment

A primary health care foundation in the Netherlands consisting of 10 primary health care practices and a home care organization employing the diabetes nurses (n = 6) agreed to become partners in the pilot. Three primary health care practices volunteered to take part in the DiabetesCoach project.

The selection criteria for patient enrollment were (1) patients with type 2 diabetes mellitus (the primary focus was on fostering lifestyle changes), (2) patients being motivated to perform self-care activities, and (3) patients having access to the Internet and being sufficiently skilled to use the Internet. Through a recruitment letter, 350 patients were invited by the caregivers to use DiabetesCoach. Patients were informed about the purpose and possibilities of the Web application both through the letter and during the office visit. In total, 50 of the 350 invited patients (14%) enrolled in the project and filled out the informed consent forms.

Training sessions (offline) were set up for the enrollees. During the training sessions the participants received instructions on how to use the application, plus a user manual. Also, an email functionality was created for technical support.

Research Design

We used a mixed-methods research design [24,25] to explore the conditions for long-term use of a Web application among patients with type 2 diabetes. Through usability tests and interviews we were able to explain the actual usage, and the survey provided insights into who uses the technology. All of the results combined provided an insight into the usage pattern and preferences of individual users for specific technology features. Log files enabled us to assess the actual and long-term usage (24 months) of the technology features. Table 1 presents an overview of the research instruments and the accompanying characteristics of the study. Figure 1 presents a chronology of the data collection process.

Directly after collecting the responses to the invitation letter, the nurses interviewed 226 of the 300 nonenrollees (patients who chose not to participate) during the office visit. Using an open-ended question the nurses asked the nonenrollees about their reasons for nonenrollment.

A *paper-based survey* was administered at baseline during the training sessions of the enrollees (n = 50) to assess patients' demographics and health-related characteristics (user profiles). The survey consisted of seven closed questions on age, gender, education, health status, diabetes duration, diabetes treatment (medication use), and treatment satisfaction. In total, 42 patients returned the survey completely filled out.

Making use of *log files* (n = 50), we measured the 2-year usage pattern by patients, the number of log-ins by patients, the mean number of hits by patients of the system's core features, and the content of the patient–nurse email messages.

We performed *usability tests* (n = 20) after 3 months of usage to investigate patients' experiences with using the Web application. Part 1 of the test consisted of a semistructured interview with open-ended questions aimed at assessing the patients' eHealth literacy, their reasons for using the Web application, and their positive or negative experiences with using the system based on the critical incidents technique [26,27]. Part 2 contained several tasks related to each feature of the Web application to identify the problems that occurred during real-time use. A trained observer (NN) watched users communicating with the interface of the application while doing simulated tasks and thinking aloud [28]. The participants' activities were recorded with audiovisual equipment (MORAE version 2.1; TechSmith Corporation, Okemos, MI, USA). The sessions were carried out at the participants' home or at the health care practice. Each test lasted for about 90 minutes.

One year after the initial use of the Web application, emails were sent to those patients (n = 20) who were not actively using the application by that time. Through an open-ended question patients were asked to report their reason for discontinueing use. We received six responses.

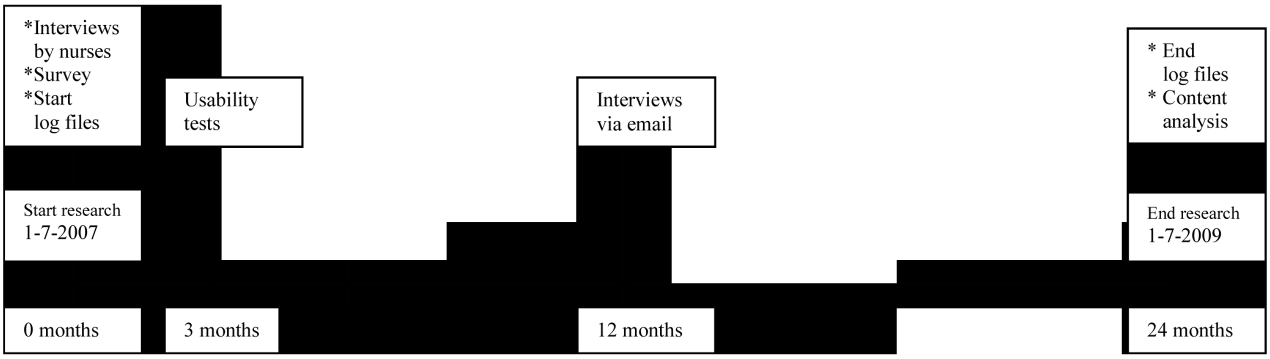
Table 1. Research instruments and study characteristics

Research instruments	n	Purpose	Participants
Interviews by nurses	226	Reasons for nonuse of the Web application	Nonenrollees[a]
Survey	50	Who uses the Web application?	Enrollees[b]
Log files	50	What features of the Web application are used? Long-term usage pattern (24 months) Profiles of continuous and discontinued users	Enrollees[b]
Usability tests	20	Reasons for use of the Web application Reasons for the decline in usage	Enrollees[b]
Email interviews	6	Reasons for the decline in usage	Enrollees[b]
Content analysis	50	What sort of information is communicated via the emails?	Enrollees[b]

[a] Primary care patients who chose *not* to participate in the DiabetesCoach project (n = 300).

[b] Primary care patients who chose to participate in the DiabetesCoach project (n = 50).

Figure 1. Chronology of the data collection process.



Data Analysis

Statistical Analyses Survey

We performed statistical analyses using SPSS version 16.0 (IBM Corporation, Somers, NY, USA). Standard descriptive statistics were performed, and chi-square tests (Fisher exact test) and *F* tests were used to identify significant differences between the different participant groups—highly active versus low/inactive users—in demographics.

Analysis of Interviews by Nurses

The researcher categorized collected responses of patients (n = 226, nonenrollees). Percentages of the answer categories were computed by multiple response analysis.

Content Analysis of Email Messages

The coding process of the patient–nurse email messages was based on the grounding theory [29], and the codes of 10 content categories (see Multimedia Appendix 1) that emerged were discussed and classified by two coders (NN, JvG). The unit of analysis contained a single statement reflecting a complete thought or idea [30,31]; this may be expressed as a simple sentence, a sentence clause, a sentence fragment, or a single word. Statements with the same meaning within the same message were coded only once. All email messages were coded independently by NN and JvG. There was 85.7% agreement across categories, with the few instances of disagreement discussed and reconciled. Duplicate messages (n = 12, patient messages) were removed.

Analysis of Usability Tests

The usability test data were analyzed using deductive analysis. NN used standard approaches for qualitative data and took detailed notes during the sessions. Notes included the problems experienced during use of the Web application such as poor navigation structures, lack of triggers to use the system, technical errors, and problems with logging on to the system [32,33]. In total, the researcher noted 166 unique problems among 20 patients. Each patient mentioned more than one problem. The coding for problem categories was derived from a conceptual framework developed earlier for the identification of usability problems with eHealth technologies [34].

Analysis of User Profiles

To identify the hardcore users we measured the actual use of the Web application by patients (n = 50) during the study period (24 months). Our measure of user activity was defined by three measures: (1) *activity pattern (continuous vs discontinued)*; measures how regularly patients have actually used the Web application until the end of the total study period (Multimedia Appendix 2, Multimedia Appendix 3); (2) *activity degree (high vs low)*; measures for how many months patients have actually used the Web application during the total study period (Multimedia Appendix 3); and (3) *number of log-ins* (Multimedia Appendix 4).

To set the norm for *discontinuity*, we looked at the activity pattern of patients (measure 1). We found that after a period of 7 months of no activity at all, patients began using DiabetesCoach again (see, for example, patient 38 in Multimedia Appendix 2 and Multimedia Appendix 3), but none of the patients did this after 8 months of no activity. In this study we therefore chose to set the norm for discontinuity at 8 months or more of no consecutive activity (Multimedia Appendix 3: search within the activity pattern (nonactive) for the number (8) or higher).

To set the norm for *high activity* (measure 2), we looked for the most active discontinued user (patient 45, Multimedia Appendix 3) and used this users' activity degree as a cut-off (67%). Continuous users with an activity degree of ≤67% were characterized as *low active users*. Continuous users with an activity degree of >67% were characterized as *highly active (hardcore) users*.

All categories of the user profiles from highly active (7 patients) and low active (10 patients), to inactive users (3 patients) were represented in the usability tests.

Results

Nonuse of the Web Application

Only 14% (n = 50) of the 350 patients responded positively to the invitation to use the Web application. Nurses interviewed 226 nonenrollees to gain insight into the barriers that inhibited their enrollment. The reasons given (n = 226) were lack of Internet (146/226, 65%), use will not have any added value (25/226, 11%), not in the mood to spend much time on the computer (23/226, 10%), not in the mood to be occupied with the disease (10/226, 4%), lack of skills to use the Internet (10/226, 4%), too busy or no time (4/226, 2%), or other, such as "patient is about to move to another town" (8/226, 4%). Obviously, patients experienced more external barriers to access (not having the equipment and lacking the right skills: 156/226, 69%) than internal motivational barriers (not willing to use it, no added value, too busy: 62/226, 27%).

Use of the Web Application

Who Uses the Web Application and Why?

The enrollees (n = 50) were aged between 43 and 80 (mean 61) years. Most were male (n = 37), of Dutch origin (40/43, 93%), with a high or medium level of education (Table 2), and treated with a diet and tablets such as metformin. Treatment satisfaction was already high before use of the Web application (40/42, 95%).

Patients mentioned three main reasons for using the Web application:

- *Increased possibilities for self-care*: the system's features stimulated patients to play a more active role in self-managing their diabetes.
- *More continuously received feedback from the nurse*: patients experienced the feeling of being better looked after by their nurse. The email feature enabled intensified contact between patient and nurse (also in-between the regular trimonthly visits).
- *Improved access to care*: email was convenient for the patients because the nurse was hard to reach by phone.

Table 2. Enrollee characteristics

Characteristic	n	%
Education (n = 43)		
Low	5	12
Medium	22	51
High	16	37
Health status (n = 43)		
Excellent	0	0
Very good	6	14
Good	25	58
Fair	12	28
Poor	0	0
Diabetes treatment (n = 43)		
None	2	5
Diet	4	9
Diet and tablets	37	86
Diet, tablets, and insulin	0	0
Diabetes duration (years) (n = 42)		
0–2	12	29
3–6	16	38
>7	14	33

What Features of the Web Application Are Used?

The log files revealed that the Web application was predominantly used for online monitoring (2216/6289, 35%; total hits of the core features of the Web application by patients during the study period: n = 6289), personal data (1648/6289, 26%) and patient–nurse email contact (1458/6289, 23%), and to a lesser extent for online education (473/6289, 8%), calendar (334/6289, 5%), personal lifestyle coach (160/6289, 3%), and the printing feature (108/6289, 2%). Patients were particularly interested in *online monitoring* for creating measurement overviews (graphs) of their blood sugar levels, weight, and blood pressure (see Multimedia Appendix 5). The *email feature* was used to supplement the online monitoring to provide explanations for their monitored values. The nurse provided weekly feedback to patients and responded to changes in metabolic values and adjusted the treatment regimen (medication) when necessary.

The *personal data* feature was used together with the online monitoring feature to track medication use to see whether a drug had been effective for improving health. The *calendar* was used to a lesser extent. Instead, email was used to communicate about appointments.

What Sort of Information is Communicated Via Emails?

In total, 323 email messages were sent during the study period. In the qualitative content analysis of the email messages, a total of 10 content categories were distinguished (see Table 3 and Multimedia Appendix 1). Certain contrasts were noticed in the content of the patient–nurse email exchange. It turned out that the nurse, more so than the patients, communicated about administrative issues and treatment plans. Communication about treatment plans referred to medication use, with a particular emphasis on medication adjustments. For the nurse the Web application functioned primarily as a means of coordinating care for more efficient communication (time savings, eg, through online appointment scheduling).

Patients, on the other hand, communicated more than nurses about their state of health and how they were feeling. For example, they let their nurse know that they were doing well, as a confirmation or ratification of the treatment regimen. As such, email was primarily used to ensure the nurse was aware of what was going on. Nurses, for their part, responded by giving affective feedback such as expressions of empathy and compliments.

Table 3. Email message content by content category quantified by statement[a]

Content categories	Total messages (n = 323)		Patients' messages (n = 130)		Nurses' messages (n = 193)	
	n	%	n	%	n	%
Measurements[b]	104	32.2	42	32.3	64	33.2
Administrative communication[c]	101	31.3	25	19.2	77	39.9
Affective communication[d]	99	30.7	38	29.2	63	32.6
DiabetesCoach remarks[d]	49	15.2	28	21.5	21	10.9
Medication use[f]	42	13.0	12	9.2	31	16.1
Physical symptoms[g]	29	9.0	19	14.6	10	5.2
Use of DiabetesCoach functionalities[h]	24	7.4	3	2.3	21	10.9
Lifestyle support[i]	20	6.2	14	10.8	8	4.1
Current events[j]	18	5.6	6	4.6	12	6.2
Other[k]	20	6.2	10	7.7	10	5.2

[a] Statement = a thematic unit (a unit of meaning within a message); one single message can contain one or more statements.

[b] Communication about clinical values such as blood sugar, blood pressure, weight, and cholesterol.

[c] Communication about referrals, appointment scheduling, etc.

[d] Expression of emotions such as compliments, relief, and worries, as well as social talk (warm wishes and thanks).

[e] Communication about (technical) problems with the use of the Web application.

[f] Communication about medication use.

[g] Communication about physical symptoms/health problems.

[i] Communication about DiabetesCoach functionalities, other than online monitoring, such as use of the lifestyle coach.

[j] Communication about new diabetes-related websites and courses.

[k] Communication not related to the use of the Web application.

Decline in Usage Over Time

Long-Term Usage Pattern

Over the total study period (24 months) each patient visited the Web application on average 49 times (2464 hits/50 patients; mean number of log-ins). See Multimedia Appendix 4 for a more detailed insight of the number of log-ins. A decline in usage over time can be observed in all three practices (Figure 2). Practice 3 had a relatively higher overall usage, probably because most technical problems had been solved by the time practice 3 started to use the application (3 months later).

The features personal data, online monitoring, and email contact were all used regularly during the study period (Figure 3).

Figure 2. Long-term use of the web application by patients per practice.

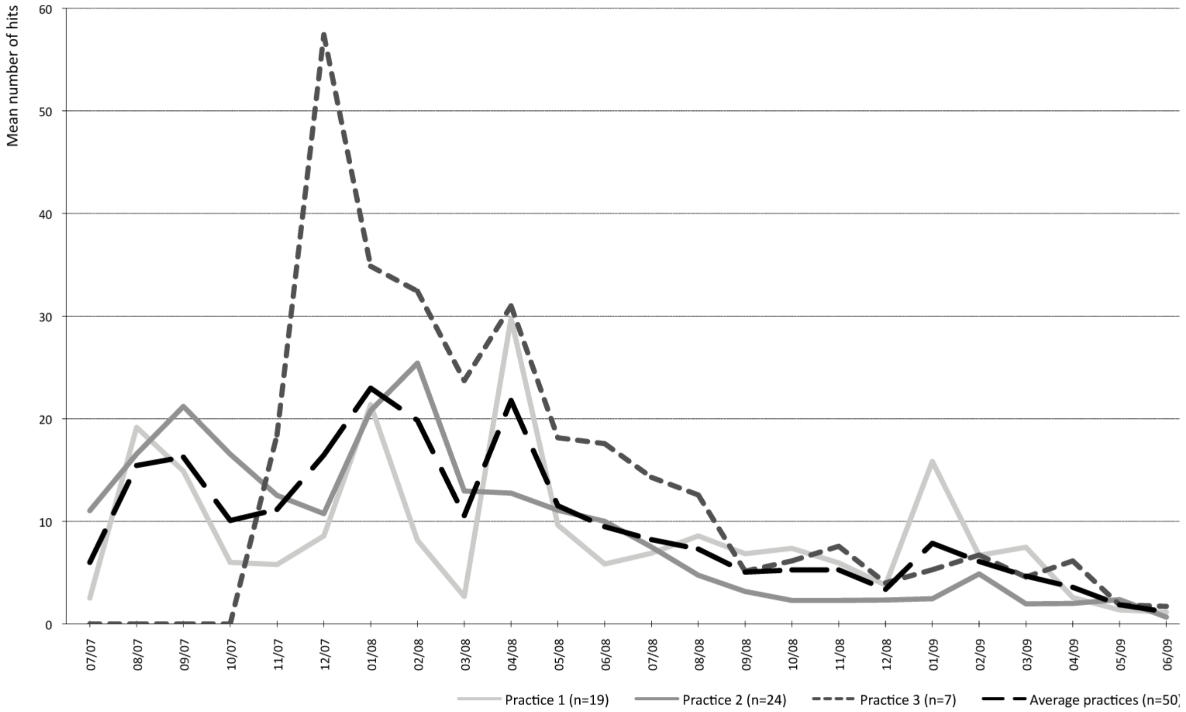
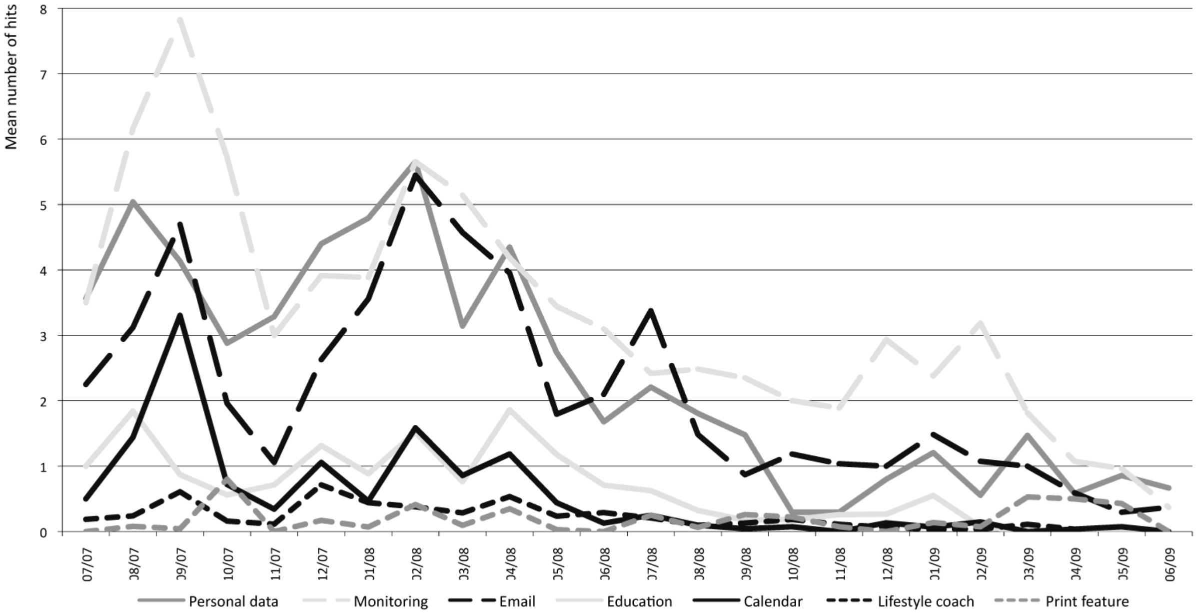


Figure 3. Long-term use of the core features of the web application by patients.



Reasons for the Decline in Usage

Reasons for the decline in usage could be attributed to a *ceiling effect* and *poor user-friendlinessof the Web application*. The results from the email interviews showed that patients forgot to use the Web application because of the absence of a reminder feature (lack of push factors). The usability tests showed that patients wished to get reminders to use the Web application, preferably through their regular (daily) email program. Patients wished to get notifications on newly posted messages by their nurse and on new and updated information on the site. Integrating the Web application with existing traditional “offline” care could also serve as a push factor. For example, patients with type 2 diabetes can be asked to use the technology for discussing online monitoring during their visit to their general practitioner or nurse.

> *Perhaps if my diabetes nurse would provide some more help or pay some more attention to it, it might result in more interest.* [Patient 17]
>
> *I wouldn't mind it being a bit more interactive; that you would get a signal to at least enter something every week and then to get some reply.* [Patient 1]

The most remarkable observation during the usability test was that the patients were unaware of the possibilities of the system, caused by uncommon navigation structures. In particular, the email feature was undiscovered, which could explain why the message overview was used more extensively than the actual sending of messages (Multimedia Appendix 5). Moreover, the calendar could be more interactive; patients wished to schedule their own appointments via the log book. However, the current log book settings only allowed the nurse to do this.

Furthermore, the email interviews revealed a ceiling effect; for some, using the application no longer had any added value. Patients with their blood sugar level under control had a less pronounced need to use a Web application for self-care support.

> *Medical checkups have been reduced to twice a year by mutual consultation with my general practitioner. A good result for me personally, but as a result there is very little for me to report.* [Patient 46]

Profiles of Highly Active Versus Low/Inactive Users

Multimedia Appendix 2 and Multimedia Appendix 3 present an overview of the monthly use of the Web application during the study period (24 months). It can be seen that use of the Web application fluctuated over time. There was no fixed regimen; each patient used the DiabetesCoach whenever it suited them (free use).

Three groups of users could be distinguished:

(1) *Continuous users who are highly active; hardcore users* (n = 16):

- Activity pattern: period of no activity <8 months (Multimedia Appendix 3)
- Activity degree: 68%–100% (17–24 months use, Multimedia Appendix 3)
- Number of log-ins: 45–191 (Multimedia Appendix 4).

(b) *Continuous users, but with lower levels of activity* (n = 18):

- Activity pattern: period of no activity <8 months
- Activity degree: 29%–67% (7–16 months use)
- Number of log-ins: 10–96.

(c) *Discontinued (inactive) users* (n = 16):

- Activity pattern: period of no activity ≥8 months
- Activity degree: 0%–67% (0–16 months use)
- Number of log-ins: 0–56.

Figure 4 presents user activity over a sustained period of time. About 66% of the enrollees continued using the Web application. Of those regular visitors, 30% can be defined as hardcore users; patients who are highly active in using the Web application.

Multimedia Appendix 2 and Multimedia Appendix 3 show that all patients from practice 1 were continuous users, whereas patients from practice 2 were more likely to be discontinued users. One possible reason for this is the closer contact between the patients and their nurse; the nurse of practice 1 was more actively involved in email contact (interactive feedback) with her patients than the nurses of practice 2 and 3 (respectively 4.5, 3.8, and 2.4 messages sent per patient on average).

When taking into account patient characteristics, the discontinued users did not differ substantially from the continuous users, although more of the discontinued users tended not to be taking medication (11/12, 92%).

We believe that more engagement in system use (being highly active) might result in better adherence to self-care activities. This is why we compared highly active users versus low/inactive users with respect to their characteristics and preferences.

We expected that patients with a greater need for care, such as the elderly, people on medication, and patients who had diabetes for a longer time, would benefit most from the technology and would therefore be more inclined to use the Web application. The results displayed in Table 4 show that highly active users were significantly more often medication users than low/inactive users were (2-sided Fisher exact test, P = .005) and had a significantly longer diabetes duration (1-sided analysis of variance, $F_{1,41}$ = 5.0, P = .03).

Table 4. Patient characteristics related to user activity

Characteristic	Highly active (n = 16)		Low/inactive (n = 34)		*P* value
	n	%	n	%	
Gender (n = 50)					.60
Male	12	75	25	73	
Female	4	25	9	26	
Age (years) (n = 50)					.28
43–56	6	37	11	32	
57–64	7	44	9	26	
65–80	3	19	14	41	
Education (n = 43)					.94
Low	2	13	3	11	
Medium	7	47	15	54	
High	6	40	10	36	
Health status (n = 43)					.59
Very good	3	20	3	11	
Good	8	53	17	61	
Fair	4	27	8	29	
Medication use (n = 43)[a]					.005
Yes (tablets)	6	40	1	4	
No	9	60	27	96	
Diabetes duration (years) (n = 42)[a]					.03
0–2	2	13	10	37	
3–6	5	33	11	41	
>7	8	53	6	22	

[a] $P < .05$

Table 5 presents an overview of the core features and ranks them according to the features that were used most:

- Ranking highly active group: (1) online monitoring, (2) email, (3) personal data.
- Ranking low/inactive group: (1) personal data, (2) online monitoring, (3) email.

Highly active users seemed to have other goals than low/inactive users. Highly active users had a higher need for online monitoring, probably because they were more likely to be frequent medication users who regularly had to pass on their clinical values to their nurse. Particularly for these patients, online monitoring would be convenient (increased access). Low/inactive users, on the other hand, appreciated the ability to document personal details such as treatment plans and medication use (comparable with a personal health record).

The features online monitoring, email, and personal data appealed to both groups, yet the highly active users used all of the features more often, spread over a longer period of time (see Table 6). In particular, they used the interactive features of online monitoring and email more extensively.

Table 5. User activity related to the use of system features: ranking of the features

	Personal data[a]	Monitoring	Email	Education	Calendar	Lifestyle coach
Highly active (n = 16)						
Total hits (2 years)	781	1601	908	240	244	96
Ranking	20%	41%	24%	6%	6%	3%
Low/inactive (n = 34)						
Total hits (2 years)	867	615	550	233	120	64
Ranking	35%	25%	23%	10%	5%	3%

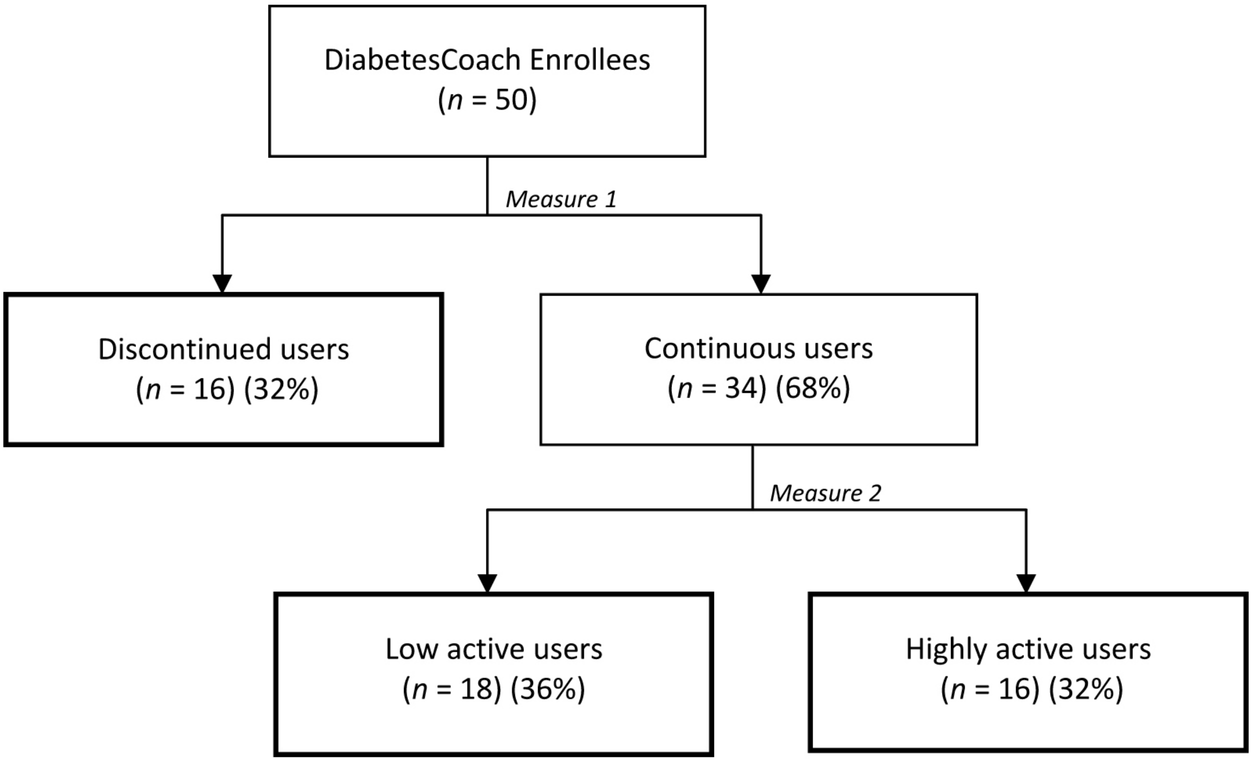
[a] Ranking: 20.2% = 781 (total hits personal data)/3870 (total hits of all core features) × 100.

Table 6. User activity related to the use of system features: mean number of hits

	Personal data	Monitoring	Email	Education	Calendar	Lifestyle coach
Highly active (n = 16)						
Total hits (2 years)	781	1601	908	240	244	96
Mean hits per patient[a]	49	100	57	15	15	6
Low/inactive (n = 34)						
Total hits (2 years)	867	615	550	233	120	64
Mean hits per patient[a]	26	18	16	7	4	2

[a] Mean hits per patient: 49 = 781 (total hits personal data)/16 (number of highly active patients).

Figure 4. User activity of DiabetesCoach enrollees.



Discussion

Main Findings

The aim of this study was to explore the factors that influenced the use of a Web-based application for supporting the self-care of patients with type 2 diabetes. The major advantages of using the Web application were improved access to care and enhanced patient–nurse communication. The features that appealed to the patients most, and with which they were often engaged, were online monitoring in combination with personal feedback through email and documentation of medication usage. These personalized and interactive features stimulated active participation by both the patient and the nurse. Patients felt better monitored by means of the continuously received feedback and were more motivated to take a more active role in self-managing their diabetes.

Unexpectedly, there was a high preference for the documentation of personal data referring to medication and treatment plans. The documentation feature is not interactive; no communication takes place. However, it is comparable, in a certain way, with a personal health record [35-37], as it includes relevant data over the course of the individual's lifetime. Patients liked to track how medication use affects their health. The personalized aspect means a lot to the patients.

Factors That Hindered Long-Term Usage of the Web Application

Selection Bias

A great concern among eHealth technologies in general, and behavioral intervention programs in particular, is that they may reach those who need them the least (ceiling effect), or they fail to reach the ones with the greatest need for care, such as patients with chronic conditions (inverse care law) [38,39]. Although in the Netherlands the e-patient is taking shape [40], this study's results still demonstrate a digital divide; the most prominent barrier to enrollment was the lack of Internet access in the patient's home. Moreover, we found a selective enrollment of relative healthy people: most diabetes patients were well regulated and thus were not the ones who could benefit most from the system. Goldberg et al [5] found similar results in their study; patients felt unengaged because they had already achieved adequate glycemic control. The use of convenience samples should be avoided because it encourages selective enrollment. It attracts patients who are already motivated and who are often the ones who least need the technology.

A Ceiling Effect

In the present study a ceiling effect ("I am doing well, so I do not need the technology") caused attrition. According to Wangberg et al [23], attrition as such is not necessarily a bad thing—in this case it can also be seen as an indicator of success, since the intervention is no longer needed. However, the ceiling effect can have another side to it: because patients do not always have a good insight into their health conditions they might wrongly think that the technology is no longer needed (overestimators). Such a ceiling effect should be avoided. Technology should therefore have persuasive elements such as feedback mechanisms and triggers (eg, email messages) to stimulate users to persist in such cases.

Poor User-Friendliness (Lack of "Push" Factors)

The results also illustrate the importance of providing automated reminders, a simple user interface, and personalized content by anticipating the needs of the individual patient. If the patient is not in need of education, then the other features should encourage the patient to use the system. The provision of features with various purposes would be more encouraging to use for a wider audience. Some users asked for the integration of monitoring, recording personal data, and logistics such as scheduling appointments. However, most of the features were presented as stand-alone applications.

Implications

To foster the widespread use of eHealth technologies like the DiabetesCoach, Internet use should be encouraged among the 65+ age range of the population; it is among the elderly that we have the largest growth potential [41,42]. To do so, the primary health care practices could consider providing training in computer and Internet skills and an opportunity to use a computer with Internet access in the practice itself for those who cannot afford the technology.

Furthermore, we believe that the less motivated or relatively unhealthy patients could benefit the most from the use of eHealth technologies because of their greater need for care and their greater challenge for health improvement. Verheijden et al [39] found that patients with deteriorating health conditions, who are thus more health care dependent, benefitted more from system use and were therefore more inclined to persistently use the Web application. In this study we found evidence, albeit very tentative, for our assumption that use of medication and the duration of the diabetes contributed to technology engagement; highly active users of DiabetesCoach were significantly more often medication users and had significantly longer duration of diabetes. Our results correspond to the findings of Wu et al [43], who found in their study of patients with chronic heart failure that those who used the system had more symptoms. The findings suggest that patients with worse disease conditions are most likely to benefit from eHealth applications. It is therefore expected that the Web application could be most useful for patients who use insulin and have a recent diagnosis of diabetes. Future studies should focus on the encouragement of eHealth among patient populations who can get the most out of it, such as those populations with high rates of behavioral risk factors and multiple chronic conditions [19,44].

In order to understand and overcome technical flaws, users should be able to give feedback during usage so that the system can be fine-tuned to their needs and user profiles. Preferably, users should actively participate in the development of the content (health 2.0) [45]. Patient-centered and -participatory design methods should be used when developing eHealth applications in order to ensure high-quality, user-informed products of demonstrated effectiveness [6,8,46-49]. Through such design approaches we are better able to customize the technology to individual preferences and user profiles. This means that the design of eHealth should start with a careful analysis of individual needs and accompanying system requirements to explore which technology is best suited for whom.

To increase adherence, technology should have push factors for persistence such as feedback mechanisms and triggers [50]. As such, it is relevant to know what kind of technology features or cues would trigger users, such as through words, images, or sounds. Reminders or triggers for use could be sent via text messages [51] and to the patients' regular email [52]. Mobile phone technology is gaining ground as a simple interface for the health consumer, given the increasing ubiquity of this technology worldwide, and will therefore be especially useful for patients who seldom use their computer.

Personalized feedback appeared to be one of the most promising features for long-term usage. In fact, two types of personalized feedback via email messages can be distinguished: personalized feedback from a caregiver via secure email and personalized feedback via automated messages and prompts. From the results of this study and the findings of Mohr et al [53] and Fry and Neff [54], we can assume that the use of personalized feedback from a real person is more persuasive than automated tailored feedback. Future research should focus on establishing which type of personalized feedback works best for whom (patients with short-term care needs ie, prevention/cure, versus patients with long-term care needs, ie, chronic disease management) and in which situation (purpose of the communication: task focused versus affective).

Moreover, integrating the technology with existing clinical care could serve as a push factor. Stevens et al [55] found that higher levels of engagement can be reached when technology requires users to log in, for example once a month. Therefore, it is expected that the effects of technology use will be stronger on patients who log in every month (fixed regimen) than on patients who log in only once in a while. By integrating eHealth technology into existing traditional offline care (visits), patients will be triggered to log in within the framework of a fixed regimen.

Besides, education should be provided in a more interactive way, for example through Web 2.0 tools that are built around user-generated or user-manipulated content, such as wikis, blogs, podcasts, and social networking sites [45,56-59].

Limitations

The limitations of this study include the small and select sample of participants. Users were self-selected, as they were motivated to use the Web application. The patients and nurses who chose to participate in the project may possibly differ from other patient groups. Further research should be conducted, preferably with larger sample groups and among nonenrollees, to gain more thorough insights into the technology preferences of the different patient groups. Nevertheless, we believe that our results provide insights beyond the current literature into patients' engagement in Web-based disease management programs. The use of a mixed-methods design [24,25] has contributed positively to this. By using interviews and usability tests we were able to explain the actual usage, and the survey provided insights into who used the technology. All of the results combined provided an insight into the usage pattern and preferences of individual users for specific technology features. Log files enabled us to assess the actual and long-term usage of the technology features.

In this study, attrition was not measured with the usual measures, such as Kaplan-Meier [22,60,61]. Most attrition measures analyze survival. However, we could not use these measures in our study because they provide insights only into the drop in usage, and not in the pattern of usage. Such survival curves are perhaps more useful for eHealth interventions with a fixed pattern of use—for example, e-therapy interventions. In our study, the pattern of usage was not fixed. Therefore, we searched for activity patterns in measuring continuity of use and we measured the degree of activity to distinguish between the infrequent users and the highly active users.

Conclusions

Our findings confirm the need for further research into usage patterns and user profiles [22]. Strategies that engage users with technology are important for addressing the low take-up of eHealth technologies. This study has set out three key strategies for increasing the initial and long-term use of eHealth technologies: (1) avoiding selective enrollment, (2) making use of participatory design methods, and (3) developing push factors for persistence. Innovations in health care will diffuse more rapidly when technology is employed that both is simple to use and has applicable components for interactivity in order to foresee the patients' need for continuous and personalized feedback, in particular for patients with a greater need for care. More longitudinal research on the use of eHealth technologies such as this study and recently published studies on attrition and adherence factors [52,53,62-70] is needed to provide insights into the way usage fluctuates over time. Through the present study we gained an insight into the differences between highly active users and nonusage dropouts, which can be seen as a first step toward decreasing attrition. The next step could be found when examining the opportunities technology has to offer. Future research should therefore focus on what kinds of system features can increase the use of eHealth technologies.

Acknowledgments

We would like to thank the primary healthcare foundation Stichting Gezondheidscentrum Eindhoven (SGE) and the home care organization Zuidzorg for their cooperation in this study.

Conflicts of Interest

None declared

Authors' Contributions

NN conceived the study and was the primary designer, although all of the authors contributed to the design. NN coordinated the collection and analysis of the data and wrote the original draft of the article. JvG, BB, SK and ES made substantial revisions and an analysis of the data. All of the co-authors contributed to and approved the final version of the manuscript.

Multimedia Appendix 1

Email message content categories.

[PDF File (Adobe PDF File), 48KB - jmir_v13i3e71_app1.pdf]

Multimedia Appendix 2

Activity pattern of patients (in months).

[PDF File (Adobe PDF File), 243KB - jmir_v13i3e71_app2.pdf]

Multimedia Appendix 3

User activity (based on activity pattern and activity degree).

[PDF File (Adobe PDF File), 74KB - jmir_v13i3e71_app3.pdf]

Multimedia Appendix 4

Number of log-ins and number of hits per feature (per patient).

[PDF File (Adobe PDF File), 138KB - jmir_v13i3e71_app4.pdf]

Multimedia Appendix 5

Number of hits on specific features by patients.

[PDF File (Adobe PDF File), 46KB - jmir_v13i3e71_app5.pdf]

References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract 2010 Jan;87(1):4-14. [doi: 10.1016/j.diabres.2009.10.007] [Medline: 19896746]
2. Task Force on Community Preventive Services. Recommendations for healthcare system and self-management education interventions to reduce morbidity and mortality from diabetes. Am J Prev Med 2002 May;22(4 Suppl):10-14. [Medline: 11985932]
3. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
4. Blanson Henkemans OA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technol Health Care 2009;17(3):253-267. [doi: 10.3233/THC-2009-0545] [Medline: 19641261]
5. 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]
6. Verhoeven F, Tanja-Dijkstra K, Nijland N, Eysenbach G, van Gemert-Pijnen L. Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. J Diabetes Sci Technol 2010 May;4(3):666-684 [FREE Full text] [Medline: 20513335]
7. Verhoeven F, van Gemert-Pijnen L, Dijkstra K, Nijland N, Seydel E, Steehouder M. The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review. J Med Internet Res 2007;9(5):e37 [FREE Full text] [doi: 10.2196/jmir.9.5.e37] [Medline: 18093904]
8. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
9. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
10. 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]
11. Izquierdo RE, Knudson PE, Meyer S, Kearns J, Ploutz-Snyder R, Weinstock RS. A comparison of diabetes education administered through telemedicine versus in person. Diabetes Care 2003 Apr;26(4):1002-1007 [FREE Full text] [Medline: 12663564]

12. Moore TJ, Alsabeeh N, Apovian CM, Murphy MC, Coffman GA, Cullum-Dugan D, et al. Weight, blood pressure, and dietary benefits after 12 months of a Web-based Nutrition Education Program (DASH for health): longitudinal observational study. J Med Internet Res 2008;10(4):e52 [FREE Full text] [doi: 10.2196/jmir.1114] [Medline: 19073541]
13. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
14. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
15. Flynn D, Gregory P, Makki H, Gabbay M. Expectations and experiences of eHealth in primary care: a qualitative practice-based investigation. Int J Med Inform 2009 Sep;78(9):588-604. [doi: 10.1016/j.ijmedinf.2009.03.008] [Medline: 19482542]
16. Tanriverdi H, Iacono CS. Diffusion of telemedicine: a knowledge barrier perspective. Telemed J 1999;5(3):223-244. [doi: 10.1089/107830299311989] [Medline: 10908437]
17. Grant RW, Cagliero E, Chueh HC, Meigs JB. Internet use among primary care patients with type 2 diabetes: the generation and education gap. J Gen Intern Med 2005 May;20(5):470-473. [doi: 10.1111/j.1525-1497.2005.04239.x] [Medline: 15963175]
18. Nijland N, van Gemert-Pijnen JE, Boer H, Steehouder MF, Seydel ER. Increasing the use of e-consultation in primary care: results of an online survey among non-users of e-consultation. Int J Med Inform 2009 Oct;78(10):688-703. [doi: 10.1016/j.ijmedinf.2009.06.002] [Medline: 19625210]
19. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95. [doi: 10.1197/jamia.M1833] [Medline: 16221943]
20. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
21. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006;8(3):e20; author reply e21 [FREE Full text] [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
22. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
23. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. Patient Prefer Adherence 2008;2:57-65 [FREE Full text] [Medline: 19920945]
24. Creswell JW, Plano Clark VL, Guttman M, Hanson W. Advanced mixed methods research designs. In: Tashakkori A, Teddlie C, editors. Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, CA: Sage Publications; 2003:209-240.
25. Doyle L, Brady AM, Byrne G. An overview of mixed methods research. J Res Nurs 2009;14(2):175-185. [doi: 10.1177/1744987108093962]
26. Ericsson KA, Simon HA. Protocol Analysis: Verbal Reports as Data. Cambridge, MA: MIT Press; 1993.
27. Flanagan JC. The critical incident technique. Psychol Bull 1954 Jul;51(4):327-358. [Medline: 13177800]
28. Gremler DD. The critical incident technique in service research. J Serv Res US 2004;7(1):65-89. [doi: 10.1177/1094670504266138]
29. Glaser BG. Basics of Grounded Theory Analysis: Emergence Vs. Forcing. Mill Valley, CA: Sociology Press; 1992.
30. Roter DL, Larson S, Sands DZ, Ford DE, Houston T. Can e-mail messages between patients and physicians be patient-centered? Health Commun 2008;23(1):80-86. [doi: 10.1080/10410230701807295] [Medline: 18443995]
31. Strijbos J, Martens R, Prins F, Jochems W. Content analysis: what are they talking about? Comput Educ 2006;46(1):29-48. [doi: 10.1016/j.compedu.2005.04.002]
32. Verhoeven F, van Gemert-Pijnen J. Discount user-centered e-health design: a quick-but-not-dirty method. In: Leitner G, editor. HCI in Work and Learning, Life and Leisure: 6th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering, USAB 2010, ... / Programming and Software Engineering). Berlin, Heidelberg: Springer; 2010:101-123.
33. Nielsen J. Usability Engineering. San Francisco, CA: Morgan Kaufmann; 1993.
34. Nijland N, van Gemert-Pijnen J, Boer H, Steehouder MF, Seydel ER. Evaluation of internet-based technology for supporting self-care: problems encountered by patients and caregivers when using self-care applications. J Med Internet Res 2008;10(2):e13 [FREE Full text] [doi: 10.2196/jmir.957] [Medline: 18487137]
35. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736. [doi: 10.1197/jamia.M2547] [Medline: 18756002]
36. 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] [doi: 10.1136/bmj.328.7449.1159] [Medline: 15142919]
37. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126. [doi: 10.1197/jamia.M2025] [Medline: 16357345]

38. Eysenbach G. Poverty, human development, and the role of eHealth. J Med Internet Res 2007;9(4):e34 [FREE Full text] [doi: 10.2196/jmir.9.4.e34] [Medline: 17951217]
39. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
40. Eurostat newsrelease. European Commission. 2010. Internet access and use in 2009: one person in two in the EU27 uses the internet daily while for young people the proportion is three quarters URL: http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/4-08122009-BP/EN/4-08122009-BP-EN.PDF [accessed 2010-06-23] [WebCite Cache ID 5qhZjXKTr]
41. Fox S, Jones S. Pew Internet. Washington, DC: Pew Internet & American Life Project; 2009 Jun. The social life of health information URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2010-06-23] [WebCite Cache ID 5qha39NwA]
42. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
43. Wu RC, Delgado D, Costigan J, Maciver J, Ross H. Pilot study of an Internet patient-physician communication tool for heart failure disease management. J Med Internet Res 2005;7(1):e8 [FREE Full text] [doi: 10.2196/jmir.7.1.e8] [Medline: 15829480]
44. Curry SJ. eHealth research and healthcare delivery beyond intervention effectiveness. Am J Prev Med 2007 May;32(5 Suppl):S127-S130. [doi: 10.1016/j.amepre.2007.01.026] [Medline: 17466817]
45. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
46. Atienza AA, Hesse BW, Baker TB, Abrams DB, Rimer BK, Croyle RT, et al. Critical issues in eHealth research. Am J Prev Med 2007 May;32(5 Suppl):S71-S74. [doi: 10.1016/j.amepre.2007.02.013] [Medline: 17466821]
47. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
48. Demiris G, Afrin LB, Speedie S, Courtney KL, Sondhi M, Vimarlund V, et al. Patient-centered applications: use of information technology to promote disease management and wellness. A white paper by the AMIA knowledge in motion working group. J Am Med Inform Assoc 2008;15(1):8-13. [doi: 10.1197/jamia.M2492] [Medline: 17947617]
49. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
50. Fogg B. A behavior model for persuasive design. In: Proceedings. 2009 Presented at: Persuasive 2009, the 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, CA, USA. [doi: 10.1145/1541948.1541999]
51. Franklin V, Waller A, Pagliari C, Greene S. "Sweet Talk": text messaging support for intensive insulin therapy for young people with diabetes. Diabetes Technol Ther 2003;5(6):991-996. [doi: 10.1089/152091503322641042] [Medline: 14709202]
52. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. J Med Internet Res 2008;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]
53. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
54. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
55. Stevens VJ, Funk KL, Brantley PJ, Erlinger TP, Myers VH, Champagne CM, et al. Design and implementation of an interactive website to support long-term maintenance of weight loss. J Med Internet Res 2008;10(1):e1 [FREE Full text] [doi: 10.2196/jmir.931] [Medline: 18244892]
56. Boulos MN, Maramba I, Wheeler S. Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. BMC Med Educ 2006;6:41 [FREE Full text] [doi: 10.1186/1472-6920-6-41] [Medline: 16911779]
57. Kamel Boulos MN, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. Health Info Libr J 2007 Mar;24(1):2-23. [doi: 10.1111/j.1471-1842.2007.00701.x] [Medline: 17331140]
58. McGee JB, Begg M. What medical educators need to know about "Web 2.0". Med Teach 2008;30(2):164-169. [doi: 10.1080/01421590701881673] [Medline: 18464141]
59. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 Diabetes Portal--the management tool, the generator, and the gatekeeper: qualitative study. J Med Internet Res 2010;12(2):e17 [FREE Full text] [doi: 10.2196/jmir.1267] [Medline: 20511179]
60. Singer JD, Willett JB. Modeling the days of our lives: using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. Psychol Bull 1991;110(2):268-290. [doi: 10.1037/0033-2909.110.2.268]

61. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford, UK: Oxford University Press; 2003.
62. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. J Med Internet Res 2010;12(1):e9 [FREE Full text] [doi: 10.2196/jmir.1347] [Medline: 20350926]
63. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
64. Chiu TM, Eysenbach G. Stages of use: consideration, initiation, utilization, and outcomes of an internet-mediated intervention. BMC Med Inform Decis Mak 2010;10:73 [FREE Full text] [doi: 10.1186/1472-6947-10-73] [Medline: 21092275]
65. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
66. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
67. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
68. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]
69. Van 't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: a longitudinal study. J Med Internet Res 2010;12(3):e37 [FREE Full text] [doi: 10.2196/jmir.1345] [Medline: 20813716]
70. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. J Med Internet Res 2010;12(1):e3 [FREE Full text] [doi: 10.2196/jmir.1361] [Medline: 20147006]

Edited by G Eysenbach; submitted 05.07.10; peer-reviewed by W Brouwer, M Wanner, M Couper; comments to author 06.11.10; revised version received 11.01.11; accepted 19.05.11; published 30.09.11.

Please cite as:
Nijland N, van Gemert-Pijnen JEWC, Kelders SM, Brandenburg BJ, Seydel ER
Factors Influencing the Use of a Web-Based Application for Supporting the Self-Care of Patients with Type 2 Diabetes: A Longitudinal Study
J Med Internet Res 2011;13(3):e71
URL: http://www.jmir.org/2011/3/e71/
doi:10.2196/jmir.1603
PMID:21959968

Original Paper

Real-Time Social Support Through a Mobile Virtual Community to Improve Healthy Behavior in Overweight and Sedentary Adults: A Focus Group Analysis

Yoshimi Fukuoka[1], PhD, RN; Emiko Kamitani[2], MS; Kemberlee Bonnet[3], MA; Teri Lindgren[4], PhD, MPH, RN

[1]Institute for Health and Aging, Department of Social and Behavioral Sciences, University of California, San Francisco, San Francisco, CA, United States

[2]Department of Community Health System, School of Nursing, University of California, San Francisco, San Francisco, CA, United States

[3]Institute for Health and Aging, University of California, San Francisco, San Francisco, CA, United States

[4]College of Nursing, Rutgers,, the State University of New Jersey, Newark, NJ, United States

Corresponding Author:
Yoshimi Fukuoka, PhD, RN
Institute for Health and Aging
Department of Social and Behavioral Sciences
University of California, San Francisco
3333 California Street
Suite340
San Francisco, CA, 94118
United States
Phone: 1 415 476 8419
Fax: 1 415 476 3915
Email: Yoshimi.Fukuoka@ucsf.edu

Abstract

Background: The onset of type 2 diabetes mellitus can be prevented or delayed by lifestyle changes. Communication technologies such as a mobile phone can be used as a means of delivering these lifestyle changes.

Objectives: The purposes of this analysis were to explore applicability of potential components of a mobile phone-based healthy lifestyle program and to understand motivators and barriers to continued engagement in a mobile phone healthy lifestyle program.

Methods: We conducted 6 focus groups (4 female and 2 male groups) in May and June 2010 with 35 focus group participants. The qualitative data were analyzed by 3 researchers using a qualitative description method in an ATLAS.ti software program. Inclusion criteria for enrollment in a focus group were as follows: (1) being aged from 30 to 69 years, (2) speaking and reading English, (3) having a sedentary lifestyle at work or during leisure time (screened by the Brief Physical Activity Survey questionnaire), and (4) having a body mass index (BMI) >25 kg/m^2 (Asian >23 kg/m^2) based on self-reported weight and height or 5) having a self-reported prediabetic condition.

Results: The mean age was 51 (SD 10.6) years; 54% (n = 19) were white; 71% (n = 25) used a mobile phone at least once a week during the last month prior to the study enrollment; and mean BMI was 32.5 (SD 6.5) kg/m^2. In the qualitative analyses, the following 4 major themes and their subthemes emerged: (1) real-time social support (real-time peer support from participants who are similarly engaged in a diet or physical activity program, and professional support from health care providers or a researcher), (2) tailoring of mobile phone programs (3) self-monitoring and motivation, and (4) potential barriers and sustainability of the program (fear of failing, age and mobile technologies, and loss of interest over time).

Conclusions: Participants from a wide range of age and racial groups expressed interest in a mobile phone-based lifestyle program. Such a program that incorporates the themes that we identified may be able to help motivate participants to increase their physical activity and to improve their diet.

(J Med Internet Res 2011;13(3):e49) doi:10.2196/jmir.1770

KEYWORDS

Mobile phone; physical activity; weight loss; diet; overweight; sedentary lifestyle; focus group; social support; diabetes mellitus

Introduction

The rapidly growing incidence of diabetes is a serious worldwide public health concern. Approximately 7.8% of the population has diabetes and 18% is prediabetic in the United States [1]. The number of people with diabetes is expected to increase in the future. Intensive face-to-face lifestyle modification programs such as changing diets, increasing the level of physical activity, and losing a moderate amount of weight can prevent or delay the onset of type 2 diabetes [2,3]; however, these programs have been expensive to implement and sustain over time. Communication technologies are available to deliver and disseminate such lifestyle modification programs to much broader audiences, making them more cost effective, but we need to understand how to better apply these technologies [4].

The mobile phone has become one of the most personalized technologies in everyday life and is omnipresent across the world today. In the United States, 87% of adults own a mobile phone, and adoption rates are even higher in other global locations [5]. Mobile phone services and accessories include short message service (SMS), multimedia messaging service, internet access, Bluetooth technologies, and global positioning systems. These technologies and services can be incorporated into lifestyle modification programs. In addition, given the portability, affordability, availability, and feasibility of mobile phones, they can be used as a means to promote a healthy lifestyle (physical activity, diet, weight loss, etc) and to prevent or manage chronic illnesses.

Qualitative methodologies, including focus groups, have been used to evaluate feasibility and acceptability of existing mobile phone-based disease management or healthy lifestyle programs [6-8]. However, few studies have explored the individual needs and expectations of a mobile phone-based program to motivate and promote healthy lifestyle modifications in sedentary and overweight or obese adults. Understanding the user's need and expectations is the first step toward designing an effective mobile phone-based healthy lifestyle program for individuals who are at risk of developing diabetes, as research has shown that a user-centered design is the best way to develop an effective lifestyle program [9]. Thus, we conducted 6 focus groups to explore the knowledge, attitudes, and behaviors of sedentary adults toward diabetes, diet, physical activity, weight loss, and the use of mobile phone-based healthy lifestyle programs in a diverse sample of sedentary overweight or obese women and men. The purposes of this analysis were to explore applicability of potential components of a mobile phone-based healthy lifestyle program and to understand motivators and barriers to continued engagement in a mobile phone healthy lifestyle program.

Methods

Study Design and Sample

Six focus groups (4 female and 2 male groups) were conducted in May and June 2010. Potential participants were recruited from urban communities through newspapers and flyers in hospitals. We screened 63 potential participants over the telephone, of whom 44 met all inclusion criteria. Of those who met all inclusion criteria, 35 participated in the focus group and 9 either did not show up for their scheduled focus group or had scheduling conflicts with potential focus group dates. The study was approved by the University of California San Francisco Institutional Review Board. All participants signed a written informed consent prior to the focus group study.

Inclusion and Exclusion Criteria

The study inclusion criteria were (1) being aged from 30 to 69 years, (2) speaking and reading English, (3) having a sedentary lifestyle at work or during leisure time (screened by the Brief Physical Activity Survey questionnaire) [10], and (4) having a body mass index (BMI) >25 kg/m^2 (Asian >23 kg/m^2) based on self-reported weight and height or (5) having a self-reported prediabetic condition. Exclusion criteria were (1) known medical conditions or other physical problems requiring special attention in an exercise program, (2) severe hearing or speech problems, (3) current participation in a lifestyle modification program or research study, and (4) known eating disorder. We did not exclude people who had never used mobile phones or were not current mobile phone users, since a previous study indicated that both nonusers and users of mobile phones were able to effectively use mobile phone applications [10].

Procedure

Individuals who were interested in participating in the focus group were screened over the telephone for inclusion and exclusion criteria. After the telephone screening, the potential participants were scheduled for a focus group. Focus groups were created based on gender and not on whether they owned or used mobile phones. Upon arrival for the focus group, participants gave written informed consent. Participants were assigned a focus group name that was associated with an alphanumeric identifier to ensure participant confidentiality; they then provided sociodemographic and medical information. The research nurse then measured participants' weight, height, waist and hip circumferences, resting blood pressure, and hemoglobin A1c in a private office. Each focus group session began with brief introductions to promote comfort and sharing of ideas within the group. To initiate the discussion, prescripted, open-ended questions were used to explore knowledge, beliefs, and attitudes about diabetes, physical activity, diet, and weight control focusing on barriers and motivators in patients who are at a risk of developing diabetes. We then proceeded from a broad discussion of diabetes, physical activity, and diet to a more narrowly focused discussion of how a mobile phone program might be used by participants (Multimedia Appendix 1). Open-ended, "what-if" questions were used to elicit participants' views on developing a useful and effective mobile phone-driven lifestyle change program. For example, we asked: "If you were going to use a mobile phone to motivate people to be physically active or eat a healthy diet, how would you use the phone?" Within these discussions, participants were asked for examples of possible applications, which led to more structured, investigator-driven prompts intended to further explore specific preconceived ideas of potential ways to use mobile phones to promote lifestyle behavior changes. For example, when participants stated that messages could be used to motivate them to continue the program, we asked: "If you were going to receive motivational messages via the mobile phone, what kind of messages would you like to receive?" This paper focuses on the participants' ideas of how to use a mobile phone and their views of investigator-designed examples of a mobile phone-driven lifestyle change program. Participants were paid $20 at the end of the focus group. The focus group interviews were digitally recorded and transcribed verbatim by a professional transcriptionist. To ensure the fidelity of the transcription, we conducted spot checks of interviews, comparing the transcript with the digital recording.

Data Management and Analysis

Descriptive statistics were used to describe the sample characteristics and medical history. These data were analyzed using SPSS 18.0 (IBM Corporation, Somers, NY, USA). Transcripts were imported into ATLAS.ti for Windows (version 6.0; ATLAS.ti GmbH, Berlin, Germany), a data management program designed to facilitate retrieval and organization of qualitative data for coding and analysis. Interview data were analyzed employing a qualitative description approach. Qualitative description was used to reduce and thematically analyze the semistructured interview data. Qualitative description involves a type of analysis that is "low inference" and less interpretive, not requiring a "highly conceptual or abstract rendering of data," making it useful for researchers who wish to obtain "unadorned" or "minimally theorized" answers to specific questions [4]. Two investigators reviewed the transcripts independently and identified multiple codes based on significant dialogues raised from interview questions. These elements were clustered into categories by observed similarity. Data were then coded using hermeneutic methods to assign meaning to each category. As the final stage of analysis, larger meaningful themes were extracted from these categories, and then the relationship between themes was articulated.

Results

Sample Characteristics

The demographic characteristics of the participants are summarized in Table 1. The mean age was 51 (SD 10.6) years, ranging from 30 to 69 years; 46% (n = 16) of participants were minorities; 43% (n = 15) were never married; and 26% (n = 9) had a part-time or full-time job. During the last month prior to study enrollment, 71% (n = 25) of participants used a mobile phone at least once a week. The mean BMI was 32.5 (SD 6.5) kg/m^2. The mean waist and hip circumferences were 106 (SD 14) cm and 117 (SD 16) cm, respectively.

Table 1. Sample demographics (N = 35)

Demographic characteristic	Mean (SD) or %(n)
Mean (SD)/(range) age (years)	51.0 (10.6)/(30–69)
Female gender	57% (20)
Race/ethnic group	
White	54% (19)
African American	31% (11)
Asian	9% (3)
Other/more than one race	6% (2)
Education: completed college or graduate school	40% (14)
Marital status	
Never married	43% (15)
Divorced/widowed	31% (11)
Married/cohabiting	26% (9)
Annual household income	
<$20,000	49% (17)
$20,000–75,000	29% (10)
>$75,000	11% (4)
Don't know/declined to answer	12% (4)
Part- or full-time employment	26% (9)
Used a mobile phone at least once a week during the last month prior to study enrollment	71% (25)
Current smoker	34% (12)
BMI kg/m^2	32.5 (SD 6.5)
Resting blood pressure (mmHg)	126 (SD 14)
Total cholesterol >200 mg/dL (self-report)	
Yes	43% (15)
No	46% (16)
Don't know	12% (4)
High density lipoprotein<40 mg/dL (self-report)	
Yes	29% (10)
No	37% (13)
Don't know	34% (12)
Waist circumference (cm)	106 (SD 14)
Hip circumference (cm)	117 (SD 16)
Mean (SD) hemoglobin A1c (%)/range	5.7 (0.06) (5.1–7.4)

Qualitative Data

Four major themes emerged from the focus group qualitative data analyses: (1) real-time peer social and professional support, (2) tailoring of timing, frequency, and content of messages in mobile phone programs, (3) combination of motivations, self-monitoring, and goal setting, and (4) potential barriers and sustainability of mobile phone programs. Each major theme consisted of specific subthemes. Details of the major themes and subthemes are described.

Real-Time Peer Social and Professional Support

One of the advantages of a mobile phone-based healthy lifestyle program was the ability to create a virtual environment to share experience and information, and to provide support to study participants. The desire for social support from other participants was the most frequent theme reported. In this theme, participants indicated that they wanted real-time support via the mobile phone. Two types of real-time support were identified in the data: (1) peer support from participants who are similarly engaged in a diet or physical activity program, and (2) professional support from health care providers or researchers. In the focus groups, female participants were more likely than male participants to report the importance of a peer support. The participants tended to regard a mobile virtual community as a peer social network or as means to receive real-time professional support. Each subtheme is described below.

Peer social support

The study participants reported that a mobile phone program could provide a home base for a real-time peer social support environment. Because individuals can access their mobile phones at their convenience, it is much easier to connect to each other and share their experiences, feelings, and obstacles. They discussed ways in which the phone could act as a virtual companion, allowing them to reach out to others who are also struggling with their diet and exercise challenges, in effect creating a peer social network. The views below exemplify this virtual environment for peer support and social networking:

> *Because it comes in as the companion-type thing. It's like people who understand what you're going through or you can share and you don't necessarily have to say,"Hey, I'm ___."You know what I'm saying?* [30-year-old woman]
>
> *...you also have a forum where you can share with other people who have similar struggles, likes/dislikes, you know, it gives you an outlet so that you can express yourself or what you're going through, you know?* [30-year-old woman]
>
> *...its easier to do with somebody, or whatever, you know and having a group of people that you are friends with, meeting people, you know, that are in the same situation and having that support network—calling each other up, you know,"How are you doing? Are you still dieting? What did you eat today? I ate this."* [31-year-old woman]

However, participants also seemed to value traditional face-to-face social support combined with a mobile or virtual peer social network. For example, as one woman participant, 58 years old, explained:

> *This would work really well for me if about once a month I was in a group like this with other women who got together to talk about their experience of what it was like to have this come over their cell phones. I think if it was just done in a vacuum it wouldn't—it wouldn't work as well.* [59-year-old woman]

Interestingly, no participant talked about accessing support through existing open social networking platforms, such as Facebook or Twitter. Their desire focused on interacting with people who are in the same healthy lifestyle program.

Professional support

Some participants indicated that they also wanted the mobile phone program to act as a connector for professional advice, or as a sponsorship program, such as Alcoholics Anonymous (AA). This subtheme differs from the previous subtheme of peer support and social networking because in this case, the participants tended to be seeking support and encouragement from health care professionals or individuals explicitly defined as "sponsors." The participants reported that a mobile phone program that included real-time automated messages and feedback could serve as an access point to research staff or health care providers for private virtual consultations about their own progress. The following views exemplify their perspectives:

> *Having a cell phone and being monitored, kind of have a sponsor. Like a little AA.*
>
> *It would be great if—you know they have 1-800-NoButts for smokers who want to quit, I would love to get encouragement over the phone without dealing with a big group of people where you get all embarrassed that you have to lose weight and they know how much weight you've lost. You just talk to someone over the phone and it's totally confidential and they're like, "Good job, you did great." This is—try to do this tomorrow. It would be really cool.* [37-year-old woman]

Moreover, in order to enhance this professional support or sponsorship component, feedback and 2-way communication are essential components of the program.

> *I could see getting a text message that said,"Look at your pedometer now. How many steps have you taken?"and you respond and tell them how many steps you've taken. And if it comes at the same time every day, then you know somebody's paying attention, and that would be a motivation, but it isn't nagging'cause it's a question. It's just saying,"How many steps have you taken?"* [42-year-old man]

Tailoring of Timing, Frequency, and Content of Messages in Mobile Phone Programs

The mobile phone program should be modified to participants' own needs. Tailoring of the program was another of the most frequently reported needs among the participants. A mobile phone healthy lifestyle program needed to be flexible enough for them to adapt their own lifestyle and routine, and not dictate their daily activities.

Individual lifestyles vary among people; thus, a mobile phone program should be adaptable to individual needs. In particular, it is important that participants receive both scheduled messages and just-in-time messages timed to meet specific needs. Just-in-time messages can be the most effective message for a participant to change their eating and physical habits. One participant stated that a just-in-time message of "put that candy bar down" would be helpful to her at the time of day when her will power often weakens. Below are examples of participants discussing the importance of tailored, scheduled messages to promote a routine:

> *...you have your timeframe and your availability time. So, you know, if the phone rings and it's not between 8:00 and 10:00 or something like that on that availability time and then (if it isn't in the designated time frame) you ignore it.* [37-year-old woman]
>
> *Staying on the routine. At a certain time, you know, I can expect to get this text at ten o'clock on Monday*

> *every Monday. I know they're going to text me ten o'clock Monday and you know that is happening.* [37-year-old woman]

Frequency of message delivery was also an issue that participants wanted to tailor. Here is one participant's suggestion for such message tailoring:

> *And at one set time a day, or so many times a week the person can—like whenever they sign up, okay, I want my reminders or my texts every day. You know, different people like different stages so they should be able to have that if they need it more or less.* [37-year-old woman]

In terms of message content, participants suggested that messages address individual weaknesses. Messages that are tailored to the individual by addressing that individual's weaknesses are more useful than those that are generalized. As one participant noted:

> *Every aspect of our lives reflects our attitudes toward food and without taking that into consideration, just saying, you know, "Eat this" or "Don't eat that," that's just not going to work.* [60-year-old woman]

Combination of Motivations, Self-Monitoring, and Goal Setting

Participants talked about 4 main components that would motivate them to engage in dietary changes and increasing physical activity. These motives are primarily internal, as they were based in their own personality. The 4 motives identified were competition (with self and/or others in the program), rewards (internal and external), feeling good about themselves, and facing life-threatening illness.

> *But you know, it's sort of like competing with yourself. I started taking my blood pressure a couple of years ago and, um, I find myself wanting to get better blood pressure. I think it's real low-key. It's real low-key. But when I get a good blood pressure, I go, "That's alright!" I sort of pat myself on the back, and when I get high blood pressure, I look for excuses.* [57-year-old man]

> *Yeah, that would be, like you said, a good idea for us to come back and show our growth and whatever we did and how...—and that's where the reward comes—see our group who's the best, who's done the best, you know what I mean?* [50-year-old woman]

> *I think, one, if it was like really fun or competitive. I think that would motivate me. And I think the second thing that would motivate me is if I had a like life-threatening illness where I had to exercise or I would decline or whatever. But I think those are the only two instances. It has to be really fun or—or you're going to die unless you do it.* [49-year-old woman]

> *I was thinking more along the lines of, you know, monthly (goals). So like you can have your daily goal set...You know, like one day, okay—well, I didn't get the star three days this week. But if you got the star the rest of the month, you see what I'm saying you would get some type of gift card or, you know, something like that.* [37-year-old woman]

The combination of self-monitoring and realistic goal setting also acted as a means of motivation, and participants responded to the notion of having a diary as a way to see their progress or problems. These diaries need to be easy to use or, as one participant stated, not "arduous," but they liked the idea of being able to complete it in real time and to consult it as needed. This self-monitoring tool could then assist participants in setting goals and making decisions about diet and exercise.

> *Just keep it with you and consult it before you pig out for dinner and say, "Well I can't have that—can't have that! you've already reached my— I can only have 500 calories for this meal."* [55-year-old woman]

These goals also need to reflect the ways in which individuals are motivated to reach for goals.

Realistic goals that are set for individuals are necessary for them to stay motivated and to remain in the program. As one participant noted:

> *It'll help that you know, like I say, at the end of the week you can look at your input and you can, you know, give yourself a personal goal. Okay, well, it's right here, I want it to be up here next week. And you can see where you actually need to go and what you need to do to get there.* [36-year-old man]

Potential Barriers and Sustainability of Mobile Phone Programs

Participants discussed potential barriers to both initially engaging with the program and sustaining this engagement over time. Additionally, they provided 3 subthemes for potential barriers and sustainability of the mobile phone program: (1) fear of failing, (2) age and mobile technologies, and (3) loss of interest over time.

Fear of failing

One of the most frequently reported potential barriers in the program was that some individuals may not participate in the program or may drop out from the program due to fear of failing to meet goals. This fear of failing was exemplified in discussions of feeling guilty and feeling like a failure; however, participants also addressed ways to mitigate this fear of failing that included positive messaging and setting multiple goals. The following statements highlight this subtheme:

> *I really think depending on what you say in that [message] would make a huge difference. I mean if it is very positive and if it's reinforcement, positive reinforcement, yes. But if you say, "Did you meet your goal this week?" then I'm going to be full of guilt, you know what I mean.* [62-year-old woman]

> *For me it would be like—I would need more aspects of my life incorporated. Say, for instance, I talk to you this week that next week I want to do such-and-such and such-and-such. And then you call me up and you say, "How about doing such-and-such*

> *and while you're at it—you know, While you're doing it you're getting this much exercise or you could do this instead of taking the bus." Incorporate it in my life so if I fail in one area I haven't necessarily failed in another area and I am kind of in a way feeling like a failure.I think that's why I could never do a diet.* [60-year-old woman]

To minimize participant's fear of failing and to motivate participants to change their lifestyle, the program should be positive, and feedback and messages should be encouraging and nonjudgmental. As one participant suggested:

> *It has to be a really warm human person that's not going to sit there and judge you.* [62-year-old woman]

Age and Mobile Technologies

Some participants reported that they were not comfortable using mobile technologies. Most of these participants were older and had limited experience in using mobile technologies. They indicated that, due to this lack of experience, older participants might view the program as non-user-friendly or too complicated to use. Two women in their 60s stated:

> *There might be an issue here too with the age.I mean you people really—they have these machines down, you know. They do it in their sleep, you know, text. But there might be a hurdle for people who are older and there might be some fear around—I mean I still can't text. I mean I'm lucky when I can text correctly. And I work for software companies.* [62-year-old woman]
>
> *I'm not tech savvy, so, I'm from the "old school" and I hate the cell phones my children give me.* [69-year-old woman]

In addition, some participants did not own a mobile phone or preferred not to use one because of financial reasons or because of its intrusiveness in their lives. Several participants commented on the cost of the cell phone within their limited budget, one stating: "I'm not paying $40 a month for telephone service." However, others indicated that they felt that cell phones disrupted their lives, as the 2 quotes below exemplify:

> *I don't own a cell phone. I have an aversion to being on call.* [65-year-old woman]
>
> *Could it be the same thing? See, you're on call, you've got to answer to this damn phone. I don't know, I'm probably in the minority.* [65-year-old woman]

Loss of Interest Over Time

Another identified potential barrier was that participants who were initially motivated to be involved in the program could lose interest over time. The participants reported that they might delete a message without reading it, or ignore a message if they were no longer excited about the program or no longer valued the messages that they received.

> *I get something called N___. At first when I got it—it's a download every day—I'd look at it and be interested.Now I mean I've got to delete, I have about five hundred that I haven't read. I know that I get like breaking news. When I first got it I read every one of them and now it's—I mean now it has to be"President dies"or something before I'm interested.* [59-year-old woman]
>
> *...they immediately text message you and I got so many text messages I just stopped reading, I just start deleting.* [59-year-old man]

To avoid this barrier, messages and feedback should not be dull or predictable in pattern. Also, tailored timing, components, and frequency of messages may help to maintain participants' interest in the program.

Discussion

Only a few focus group studies have been performed to understand participants' needs and expectations for a mobile phone program as a process of program development for the prevention of chronic illnesses [8,11,12]. Participants in this study provided valuable data to aid in the development of specific components of a mobile phone program that would motivate individuals with prediabetic conditions to increase their physical activity and improve diet-related behaviors.

The themes we identified were (1) having real-time peer social and professional support, (2) tailoring timing, frequency, and content of messages in mobile phone programs, (3) combining motivations, self-monitoring, and goal setting, and (4) identifying potential barriers and sustainability of mobile phone programs. These findings suggest that creating a mobile virtual community could be effective in engaging and sustaining participants in lifestyle change programs. This virtual community acts as a closed social network that can provide both peer and professional support to participants. It is flexible enough to provide messages and self-monitoring that are tailored to individual participants and address barriers to participation (see Multimedia Appendix 2 for the components of a mobile virtual community).

Participants indicated that the primary component of a successful mobile phone-based healthy lifestyle program was real-time peer social and professional support that would create a mobile virtual community. Through this mobile virtual community, participants could create a peer social network that would support their efforts to change their diet and physical activity and receive immediate feedback from health care providers, a designated sponsor, or a peer [13,14]. Participants indicated that they would be comfortable sharing their experiences and struggles with others in their virtual community, but they would not be comfortable doing so with people outside the created virtual community. Contrary to Greene et al., participants in our study indicated that they wanted a closed social network rather than an open one [15]. Although Greene et al. reported that Facebook enabled participants to share experiences and to receive answers to their questions or to receive direct feedback from individuals with diabetes, none of our participants mentioned an interest in open social networks such as Facebook [15]. In addition, social support strongly relates to motivations for lifestyle change. Peer social support is well known to be a crucial motivational tool to promote physical activity [16], particularly in women. Since a mobile phone program is very portable and participants can bring it with them wherever they go, they can receive continuous support and be connected to their peers at all times. Thus, participants in this mobile phone program may perceive the virtual mobile community as a beneficial component of the program.

Tailoring of the program is another key component of a successful mobile phone program [17-21]. Frequency, timing, and content of messages, as well as goal reminders, should be tailored to an individual's needs. Just-in-time messages or feedback at the moment of decision or action play an important role in keeping participants motivated [22]. To monitor participants' behaviors and send just-in-time messages, the program should have the ability of not only sending messages, but also of receiving reports from participants about their current activities and of providing feedback. A program of this capacity can send real-time, context-aware feedback and messages, and can get feedback in real time. The real-time feedback component of the program that elicits messages before patients take counterproductive actions is important to motivate then to make healthy choices [23]. The program, therefore, must incorporate 2-way messaging. Previous programs using mobile phone technology have focused on using 1-way SMS (text messages), but this may not be sufficient for all participants. Importantly, feedback and goal setting need to be tailored to an individual based on their strength, ability, and barriers, as well as achievement of the previous goal [23,24].

This study reinforces the need for self-monitoring; however, participants also connected this idea with broader motivations for enacting behavioral changes. Participants identified the value of competition to keep one engaged and the importance of some form of reward mechanism. The challenge for a mobile phone program is to incorporate different ways to motivate participants. The program would have to be able to engender a sense of competition, at least within participants that are so motivated, but it would also need to determine effective rewards that can be potentially tailored, both individually and over time.

Participants also identified potential barriers to initiating and sustaining engagement with the program. Fear of failing was one of the major potential barriers addressed by participants. A previous study found that women who used a mobile phone physical activity program perceived the program as a virtual coach to increase their motivation for a healthier lifestyle. Coaching is an effective approach to behavior change, but pressure from a coach or another's expectation can create fear of failure in the participant [25,26]. To minimize participants' fears, nonjudgmental approaches, positive reinforcement, and establishing more than 1 short-term goal should be factored into the mobile phone program to mitigate fears of failing.

Age and the potential intrusiveness of mobile phones in people's lives could also be a barrier to engagement with the program. Older participants expressed concerns about their ability to use the mobile phone technologies, especially texting, and that their lack of technical ability might limit their willingness to participate in the program. Over 25% of the participants had not used a mobile phone in the last month. For some, this was due more to lack of funds than to a dislike of mobile phones. However, a few expressed a rather strong antagonism toward the impact of mobile phones on their way of life. Given this expressed antipathy to mobile phones and the number of nonusers in the focus groups, conducting further research with only nonusers could be warranted to more fully understand this barrier.

In terms of sustaining participation in this type of program, there is limited research that has explored issues around long-term engagement in these types of programs. Participants in this study indicated that they needed to continue to value the program. Suggestions for how to maintain interest over time included keeping the program simple and user friendly. This component may be particularly important to accommodate older adults, people with less education, and people who are unfamiliar with mobile phones.

Additionally, the messages need to be short and varied over time. In addition, a system of rewards needs to immediately reflect their positive behavior change but may also need to change over time. A randomized variety of messages may help to avoid the boredom, and a program tailored to the individual's needs could help to keep participants motivated. However, more research is needed to understand how these rewards can be incorporated into mobile phone technologies and how to modify them over time. Finally, resistance to using mobile technologies among older individuals is also a potential barrier [27] that needs to be more fully explored. The participants' suggestions and ideas based on their own experiences are great sources to develop the most effective components of a mobile phone-based healthy lifestyle program.

Limitations of the Study

The study results should be interpreted within the following limitations. All participants in this focus group were recruited from only 1 geographic location, San Francisco, approximately 74% of the participants were unemployed, and 49% had incomes less than $20,000. Thus, the findings in the study may not be representative of more affluent or more fully-employed populations in the United States.

Conclusions

Participants from a wide range of age and racial groups expressed interest in a mobile phone-based lifestyle program. A mobile phone-based lifestyle program that incorporates these identified themes may be able to help motivate participants to increase their physical activity and to improve their diet. This study can help researchers to further understand individual needs and expectations of a mobile phone program to motivate and promote healthy lifestyle modifications in sedentary and overweight or obese adults.

Acknowledgments

This study was supported by NIH/NCRR UCSF-CTSI Grant Number UL1 RR024131 and NINR K23NR011454. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.”

Conflicts of Interest

None declared

Multimedia Appendix 1

Questions for Focus Groups

[PDF file (Adobe PDF), 45 KB - jmir_v13i3e49_app1.pdf]

Multimedia Appendix 2

Summary of Focus Group Themes

[PDF file (Adobe PDF), 53 KB - jmir_v13i3e49_app2.pdf]

References

1. American Diabetes Association. American Diabetes Association. 2011. Diabetes Statistics: Data from the 2011 National Diabetes Fact Sheet URL: http://www.diabetes.org/diabetes-basics/diabetes-statistics/ [accessed 2011-07-01] [WebCite Cache ID 5zrUW13KN]
2. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
3. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001 May 3;344(18):1343-1350 [FREE Full text] [doi: 10.1056/NEJM200105033441801] [Medline: 11333990]
4. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [Medline: 10940958]
5. Centers for Disease Control and Prevention. CDC. 2009. Mobile eHealth Data Brief URL: http://www.cdc.gov/socialmedia/Data/Briefs/mobileresearch.pdf [accessed 2011-07-01] [WebCite Cache ID 5zrWioDxO]
6. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. J Med Internet Res 2004 Dec 2;6(4):e42 [FREE Full text] [doi: 10.2196/jmir.6.4.e42] [Medline: 15631966]
7. Arsand E, Tatara N, Østengen G, Hartvigsen G. Mobile phone-based self-management tools for type 2 diabetes: the few touch application. J Diabetes Sci Technol 2010 Mar;4(2):328-336 [FREE Full text] [Medline: 20307393]
8. Pinnock H, Slack R, Sheikh A. Misconnecting for health: (lack of) advice for professionals on the safe use of mobile phone technology. Qual Saf Health Care 2007 Jun;16(3):162-163. [doi: 10.1136/qshc.2006.021345] [Medline: 17545339]
9. Holtzblatt K. Customer-centered design for mobile applications. Personal and Ubiquitous Computing 2005;9(4):227-237. [doi: 10.1007/s00779-004-0324-5]
10. Fukuoka Y, Vittinghoff E, Jong SS, Haskell W. Innovation to motivation: pilot study of a mobile phone intervention to increase physical activity among sedentary women. Prev Med 2010;51(3-4):287-289. [doi: 10.1016/j.ypmed.2010.06.006] [Medline: 20600263]
11. Boushey CJ, Kerr DA, Wright J, Lutes KD, Ebert DS, Delp EJ. Use of technology in children's dietary assessment. Eur J Clin Nutr 2009 Feb;63(Suppl 1):S50-S57. [doi: 10.1038/ejcn.2008.65] [Medline: 19190645]
12. Halifax NV, Cafazzo JA, Irvine MJ, Hamill M, Rizo CA, McIssac WJ, et al. Telemanagement of hypertension: a qualitative assessment of patient and physician preferences. Can J Cardiol 2007 May 15;23(7):591-594. [Medline: 17534469]
13. Eagle N, Pentland AS, Lazer D. Inferring friendship network structure by using mobile phone data. Proc Natl Acad Sci U S A 2009 Sep 8;106(36):15274-15278 [FREE Full text] [doi: 10.1073/pnas.0900282106] [Medline: 19706491]
14. Jen WY. The adoption of mobile weight management services in a virtual community: the perspective of college students. Telemed J E Health 2010 May;16(4):490-497. [doi: 10.1089/tmj.2009.0126] [Medline: 20507202]
15. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292. [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
16. Salvy SJ, Roemmich JN, Bowker JC, Romero ND, Stadler PJ, Epstein LH. Effect of peers and friends on youth physical activity and motivation to be physically active. J Pediatr Psychol 2009 Mar;34(2):217-225. [doi: 10.1093/jpepsy/jsn071] [Medline: 18617572]

17. Bull FC, Holt CL, Kreuter MW, Clark EM, Scharff D. Understanding the effects of printed health education materials: which features lead to which outcomes? J Health Commun 2001;6(3):265-279. [doi: 10.1080/108107301752384442] [Medline: 11550593]
18. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. Patient Educ Couns 1999 Feb;36(2):193-203. [Medline: 10223023]
19. Ryan P, Lauver DR. The efficacy of tailored interventions. J Nurs Scholarsh 2002;34(4):331-337. [Medline: 12501736]
20. Suggs LS. A 10-year retrospective of research in new technologies for health communication. J Health Commun 2006;11(1):61-74. [doi: 10.1080/10810730500461083] [Medline: 16546919]
21. Trevena LJ, Davey HM, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. J Eval Clin Pract 2006 Feb;12(1):13-23. [doi: 10.1111/j.1365-2753.2005.00596.x] [Medline: 16422776]
22. Intille SS. A new research challenge: persuasive technology to motivate healthy aging. IEEE Trans Inf Technol Biomed 2004 Sep;8(3):235-237. [Medline: 15484427]
23. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
24. Tufano JT, Karras BT. Mobile eHealth interventions for obesity: a timely opportunity to leverage convergence trends. J Med Internet Res 2005;7(5):e58 [FREE Full text] [doi: 10.2196/jmir.7.5.e58] [Medline: 16403722]
25. Elder JP, Ayala GX, Harris S. Theories and intervention approaches to health-behavior change in primary care. Am J Prev Med 1999 Nov;17(4):275-284. [Medline: 10606196]
26. Sagar SS, Stoeber J. Perfectionism, fear of failure, and affective responses to success and failure: the central role of fear of experiencing shame and embarrassment. J Sport Exerc Psychol 2009 Oct;31(5):602-627. [Medline: 20016111]
27. Charness N, Boot WR. Aging and information technology use: potential and barriers. Curr Dir Psychol Sci 2009;18(5):253-258. [doi: 10.1111/j.1467-8721.2009.01647.x]

Abbreviations

AA: Alcoholics Anonymous
BMI: body mass index
SMS: short message service

Edited by G Eysenbach; submitted 11.02.11; peer-reviewed by A Elyer, J Tufano; comments to author 10.03.11; revised version received 09.05.11; accepted 09.05.11; published 14.07.11.

Please cite as:
Fukuoka Y, Kamitani E, Bonnet K, Lindgren T
Real-Time Social Support Through a Mobile Virtual Community to Improve Healthy Behavior in Overweight and Sedentary Adults: A Focus Group Analysis
J Med Internet Res 2011;13(3):e49
URL: http://www.jmir.org/2011/3/e49/
doi:10.2196/jmir.1770
PMID:21752785

Original Paper

An Evaluation of the Use of Smartphones to Communicate Between Clinicians: A Mixed-Methods Study

Robert Wu[1,2], MD FRCPC MSc; Peter Rossos[3], MD FRCPC MBA; Sherman Quan[2], BSC; Scott Reeves[4,5,6,7], PhD; Vivian Lo[2], MASc; Brian Wong[8], MD FRCPC; Mark Cheung[8], MD FRCPC; Dante Morra[1,2,9], MD FRCPC MBA

[1]University Health Network, Division of General Internal Medicine, Toronto, ON, Canada
[2]Centre for Innovation in Complex Care, University Health Network, Toronto, ON, Canada
[3]University Health Network, Department of Medicine, Toronto, ON, Canada
[4]Centre for Faculty Development, Li Ka Shing International Health Care Education Centre, St Michael's Hospital, Toronto, ON, Canada
[5]Keenan Research Centre, Li Ka Shing Knowledge Institute of St Michael's Hospital, Toronto, ON, Canada
[6]Wilson Centre for Research in Education, University Health Network, Toronto, ON, Canada
[7]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[8]Sunnybrook Health Sciences Centre, Department of Medicine, Toronto, ON, Canada
[9]Centre for Interprofessional Education, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Robert Wu, MD FRCPC MSc
University Health Network
Division of General Internal Medicine
200 Elizabeth Street 14EN-222
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4567
Fax: 1 416 595 5826
Email: robert.wu@uhn.on.ca

Abstract

Background: Communication between clinicians is critical to providing quality patient care but is often hampered by limitations of current systems. Smartphones such as BlackBerrys may improve communication, but studies of these technologies have been limited to date.

Objective: Our objectives were to describe how smartphones were adopted for clinical communication within general internal medical wards and determine their impact on team effectiveness and communication.

Methods: This was a mixed-methods study that gathered data from the frequency of smartphone calls and email messages, clinicians' interviews, and ethnographic observations of clinical communication interactions. Triangulation of qualitative and quantitative data was undertaken to develop common themes that encompass comprehensive and representative insights across different methods.

Results: Findings from our study indicated that over a 24-hour period, nurses sent on average 22.3 emails to the physicians mostly through the “team smartphone,” the designated primary point of contact for a specific medical team. Physicians carrying the team smartphone received on average 21.9 emails and 6.4 telephone calls while sending out 6.9 emails and initiating 8.3 telephone calls over the 24-hour period. Our analyses identified both positive and negative outcomes associated with the use of smartphones for clinical communication. There was a perceived improvement in efficiency over the use of pagers for clinical communication for physicians, nurses, and allied health professionals. In particular, residents found that the use of smartphones helped to increase their mobility and multitasking abilities. Negative outcomes included frequent interruptions and discordance between what doctors and nurses considered urgent. Nurses perceived a worsening of the interprofessional relationships due to overreliance on messaging by text with a resulting decrease in verbal communication. Unprofessional behaviors were observed in the use of smartphones by residents.

Conclusions: Routine adoption of smartphones by residents appeared to improve efficiency over the use of pagers for physicians, nurses, and allied health professionals. This was balanced by negative communication issues of increased interruptions, a gap in perceived urgency, weakened interprofessional relationships, and unprofessional behavior. Further communication interventions are required that balance efficiency and interruptions while maintaining or even improving interprofessional relationships and professionalism.

(J Med Internet Res 2011;13(3):e59) doi:10.2196/jmir.1655

KEYWORDS

Email; cellular phone; interdisciplinary communication; hospital communication systems

Introduction

Effective communication between clinicians to coordinate patient care is critical for providing quality health care to patients [1]. Frequent interruptions through paging is a major communication issue [2-4], and poor communication can result in inefficiencies and errors [5-7]. The burden from inefficient communication has been well documented in multiple areas of hospital care, and a systematic review has linked interruptions to medical errors [8-12].

The use of smartphones such as BlackBerrys may improve hospital communication as they provide multiple communication modalities [3]. For example, with urgent issues, direct calling can eliminate the need to wait for a page to be answered. In contrast, for nonurgent issues, asynchronous communication through email can be used, which could reduce disruptions. While many clinicians own cellular phones and smartphones, their use for clinical communication is variable [13,14]. Other than two studies that reported a perceived improvement in workflow efficiency from clinicians' surveys, the study of smartphones used in hospital communication has been limited [15,16].

To understand the impact of the use of smartphones on the delivery of hospital care, we conducted a mixed-methods study to describe how smartphones were used, to identify advantages and disadvantages associated with their use, and to determine how their use can be improved.

Methods

Intervention

Beginning in March 2008, each resident on the general internal medicine units received an individual BlackBerry smartphone to use for clinical communication typically within or among the medical teams [15]. BlackBerry devices were selected because of the secure email functionality and because these were the standard smartphones used by hospital administration. In addition, each team also had a "team BlackBerry" that was designated as the primary point of contact for nurses and allied health professionals to communicate with the teams. The team BlackBerry was typically carried during the day by the senior resident and then given to the covering junior resident during sign over. Specifically, nurses would contact residents by sending emails to the team BlackBerry with the following structures and information: (1) the patient's name, (2) the nurse's name, (3) the issue and purpose of contact, and (4) their preferred response (callback, email, or no response required). For urgent patient issues, nurses and other clinicians were asked to call the team BlackBerry directly. For nonurgent issues, clinicians were asked to contact the team BlackBerry through email. This recommendation was based on previous findings in which a high number of direct calls were found to be very disruptive for residents [15].

Design

A mixed-methods approach was adopted to obtain a variety of data sources on the communication processes occurring on the wards. The following three methods were used: (1) quantitative measures assessing the frequency and use of smartphone calls and email messaging, (2) semistructured interviews with clinicians, and (3) ethnographic observations of clinical communication interactions were conducted using two techniques: (1) nonparticipatory "work shadowing" (defined below) and (2) observations at the nursing stations.

The settings were four general internal medicine wards at two large urban teaching hospitals in Canada. At each site, there were four medical teams each consisting of an attending physician, a senior resident, junior residents, and medical students. The study was conducted from January 5, 2009, until May 28, 2010.

Data Collection

Between January 5, 2009 and May 28, 2010, quantitative and qualitative data were collected on communication patterns from in-depth interviews and from observations of communications.

Communication Patterns

Quantitative data on the usage of the smartphones to receive and place telephone calls and emails were gathered by accessing email accounts and phone records of consenting residents. A total of 12,936 emails and 13,717 phone calls were analyzed from 34 residents (Table 1).

In-depth Interviews

Semistructured interviews were conducted to explore clinicians' perceptions of their experiences using the smartphones. To conduct qualitative comparisons and to ensure that a variety of clinicians' perspectives were represented, we adopted a purposive sampling strategy where different groups of health care professionals with differing views on the use of smartphones for clinical communications were interviewed (Table 1). Each interview lasted between 15 and 40 minutes and was carried out at a mutually convenient location within the two hospital sites. The interviews were conducted and audio taped by an independent research associate and then professionally transcribed before analysis. The interview protocol consisted of a series of open-ended questions with appropriate follow-up probes that focused on users' perspectives and their experiences using the smartphone technology.

Observations of Communications

Ethnographic observational methods were employed to explore and understand the in-depth communication processes and behaviors around smartphone use [17]. Field notes included recording the usage on the types of communication tools used as well as communication interactions and incidents between clinicians. Two types of observational methods were used. First, a nonparticipatory "work-shadowing" approach was employed whereby a researcher followed medical residents for 2- to 5-hour periods during the day and evening shifts starting from 10 am to 11 pm. Both incoming and outgoing communications were recorded during these shifts. Second, observations were also conducted for 2-hour periods at the general internal medicine nursing stations, the hubs of interprofessional communication starting from 10 am and lasting until 10 pm (Table 1).

The study was approved by the Research Ethics Board, University Health Network, Toronto, Ontario.

Table 1. Breakdown on the data methods by data collection

Data Methods	Site 1	Site 2	Total
Blackberry usage			
Residents, n	17 residents	17 residents	34 residents
Emails, n	3946 emails	8990 emails	12,936 emails
Phone calls, n	5714 calls	8003 calls	13,717 calls
Semi structured interviews			
By attending physicians, n	4	0	4
By medical residents, n	3	1	4
By nurses, n	8	7	15
By allied health professionals (pharmacists, social workers, occupational therapists), n	7	1	8
Work shadowing			
Senior residents, (n) time observed (h:min)	(5) 24 h:44 min	(2) 9 h:33 min	(7) 34 h:17 min
Junior residents, (n), time observed (h:min)	(5) 24 h:11 min	(2) 8 h:13 min	(7) 32 h:24 min
Total residents, (n), time observed (h:min)	(10) 48 h:55 min	(4) 17 h:46 min	(14) 66 h: 41 min
Day shifts (10 am to 6 pm): total hours observed (h:min)	34 h:36 min	12 h:48 min	47 h:24 min
Evening shifts (5 pm to 11:30 pm): total hours observed (h:min)	14 h:19 min	4 h:58 min	19 h:17 min
Ward observations at nursing stations			
Number of nursing wards sampled	2	2	4
Number of observation sessions conducted	21	15	36
Total hours observed (h:min)	42 h:28 min	29 h:51 min	72 h:19 min
Hours observed weekdays, daytime (10 am to 6 pm) (h:min)	24 h:23 min	17 h:51 min	42 h:14 min
Hours observed,weekdays, evening (6 pm to 10 pm) (h:min)	6 h:3 min	8 h:0 min	14 h:3 min
Hours observed weekend, daytime (10 am to 6 pm) (h:min)	12 h:2 min	4 h:0 min	16 h:2 min

Analysis

To determine communication volume using smartphones, descriptive statistics of calls and emails per day were calculated from email and call logs. Incoming and outgoing communications recorded during work-shadowing sessions were also identified and descriptive statistics of the different communication methods per hour were calculated. All emails sent by the consenting residents were analyzed to determine the frequency of emails that were received by others such as nurses, attending physicians, or allied health providers. Similarly, all emails received by consenting residents were analyzed to determine the frequency of emails received from different types of senders.

Interviews were transcribed, and inductive thematic analysis was performed using the qualitative data analysis software NVivo 8 (QSR international, Doncaster, Victoria, Australia). The transcripts were coded by three members of the research team (authors RW, SR, and VL) to derive and identify a number of common perceptions and broad themes.

For the ethnographic methods, the number and types of communication events occurring during the work-shadowing sessions were tabulated. Field notes from both the work-shadowing and ward observations were reviewed to identify common themes.

Triangulation of qualitative and quantitative data was undertaken to develop themes that encompass comprehensive and representative insights that are common across multiple methods.

Results

Key results are presented in two main sections. First quantitative data reporting smartphone communication volume is described. Second, qualitative data describing perceptions and ethnographic fieldwork is presented.

Description of Communication Volume

The usage of smartphones for telephone calls is shown in Table 2. Outgoing calls were placed to the hospital 41.2% of the time, to another BlackBerry 25.3% of the time, and to external numbers 33.4% of the time. From the 12,936 emails that were sent from or received from the smartphones, the daily frequencies of who communicates with whom were calculated (Figure 1).

The different incoming and outgoing communications observed during work shadowing of residents are listed in Table 2. Note that these numbers are different from those calculated from device usage as work shadowing typically occurred at busier times during the day and evening.

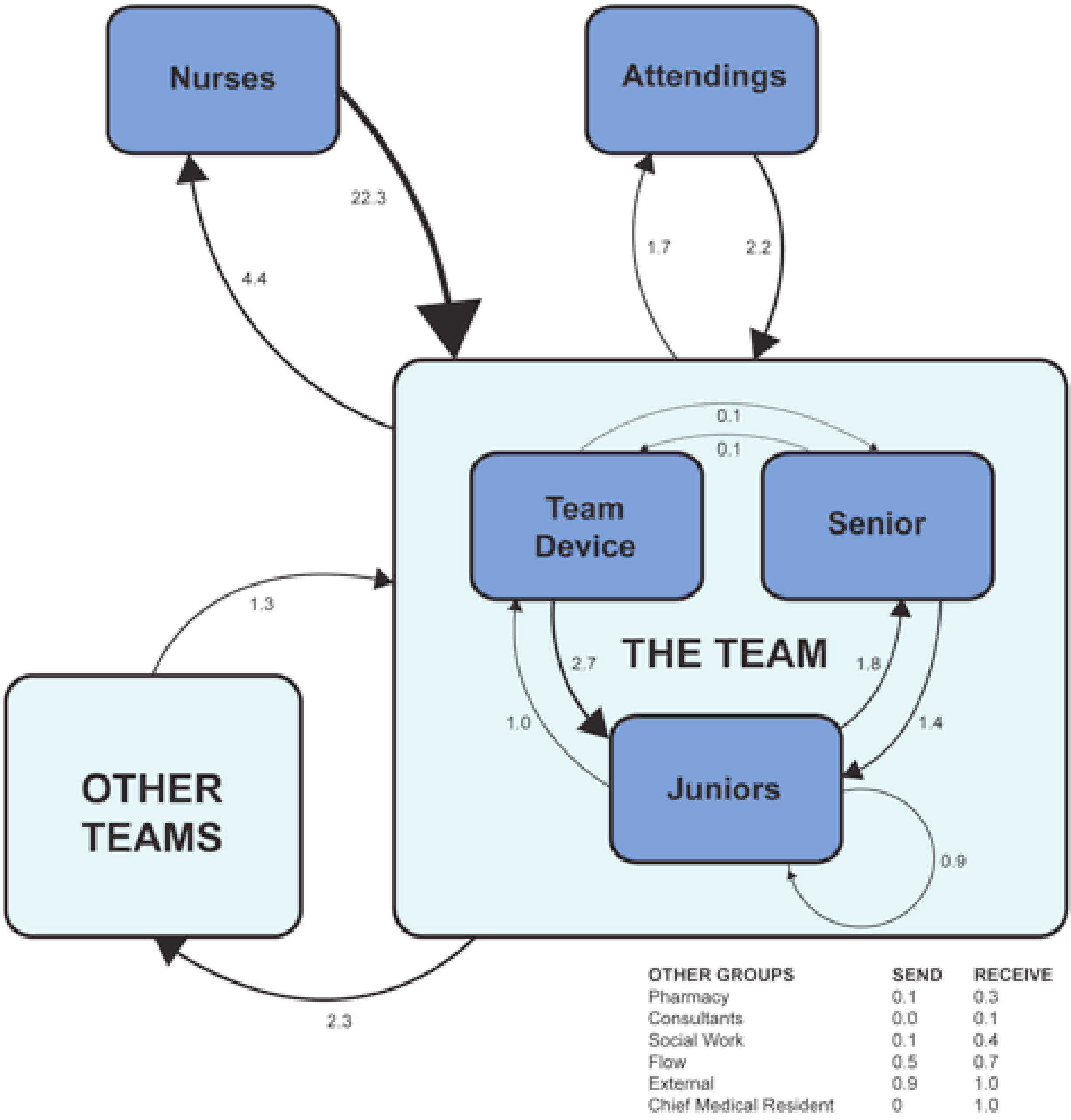
Table 2. Communications to and from smartphones based on usage and work-shadowing data

	Modes of Communication						
	Telephone Calls Mean (SD)		Emails[1] Mean (SD)		Pages[2] Mean (SD)	Face-to-Face Conversations Mean (SD)	
Communication Processes	Received	Initiated	Received	Initiated	Initiated	Received	Initiated
From analysis of device usage							
Team smartphones per 24-hour period	6.4/24h (5.3)	8.3/24h (6.4)	21.9/24h (10.1)	6.9/24h (4.8)	Not Applicable	Not Applicable	Not Applicable
Senior smartphones per 24-hour period	5.2/24h (3.9)	5.3/24h (5.0)	3.3/24h (2.8)	2.1/24h (2.2)	Not Applicable	Not Applicable	Not Applicable
Junior smartphones per 24-hour period	3.9/24h (3.2)	5.4/24h (5.1)	3.6/24h (3.9)	2.4/24h (3.5)	Not Applicable	Not Applicable	Not Applicable
From analysis of work- shadowing observation							
Resident communications per hour	1.1/h (1.2)	1.4/h (0.8)	1.7/h (1.8)	1.0/h (1.1)	0.7/h (0.8)	1.8/h (0.8)	2.3/hr (1.4)

[1] Emails include regular emails as well as short messaging services messages.

[2] Since only residents on general medicine were given smartphones, paging was typically used to contact other services or medical students who did not have smartphones.

Figure 1. Email communication groups and frequency (average emails/day)



OTHER GROUPS	SEND	RECEIVE
Pharmacy	0.1	0.3
Consultants	0.0	0.1
Social Work	0.1	0.4
Flow	0.5	0.7
External	0.9	1.0
Chief Medical Resident	0	1.0

Themes

The analysis of qualitative data (interviews, work shadowing, and ward observations) generated five major themes: efficiency, interruptions, interprofessional relations, gaps in perceived urgency, and professionalism (Table 3).

Table 3. Themes with number of participants and number of occurrences by data collection method

	Data Collection Method		
	Interviews	Work Shadowing Residents	Ward Observations
Key themes			
Theme 1: Efficiency	n participants (n quotes)	n participants (n incidents)	n participants (n incidents)
Improved efficiency	8 physicians (38) 11 nurses (36) 5 allied health (23)	12 residents (39)	23 clinicians (58)
Reduced efficiency	3 physicians (8) 9 nurses (16) 1 allied health (3)	3 residents (4)	3 clinicians (4)
Theme 2: interruptions			
Increased Interruptions	7 physicians (34) 4 nurses (7) 1 allied health (3)	13 residents (46)	7 clinicians (8)
Theme 3: interprofessional relationships			
Improved interprofessional collaboration	5 physicians (9) 3 nurses (3) 2 allied health (2)	0 residents (0)	0 clinicians (0)
Reduced interprofessional collaboration	3 physicians (4) 9 nurses (27) 1 allied health (2)	3 residents (3)	8 clinicians (15)
Theme 4: gaps in perceived urgency			
Differing standards for when emails or direct calls should be used	6 physicians (20) 12 nurses (22) 2 allied health (3)	2 residents (2)	15 clinicians (25)
Theme 5: professionalism			
Perceived lack of etiquette for answering calls or texting on smartphone devices	2 physicians (9) 2 nurses (3) 3 allied health (10)	8 residents (16)	1 clinicians (1)

Efficiency

Nurses and Allied Health Professionals

With the smartphone system, nurses no longer felt the need to wait for a telephone reply, which typically was required with paging, and this resulted in less “phone tag.” Some nurses and allied health professionals perceived a faster response and increased accessibility to physicians (Textbox 1, data extract 1). They found the use of emails helped to convey their patient’s status quickly and efficiently to doctors (Textbox 1, data extract 2). Nurses also reported that since their emails were sent to the team BlackBerry, less time was spent trying to locate a specific resident (Textbox 1, data extract 3). Some of the nurses interviewed appreciated that the smartphone system allowed direct and immediate communication by phone with physicians for urgent issues. Direct calls were observed to be from nurses or other clinicians who previously were unsuccessful at getting a response from a resident through email or paging (Textbox 1, data extract 4).

A number of the nurses interviewed perceived that the new system added a barrier, with increased difficulty reaching doctors. Instead of being able to resolve complex issues quickly with a page coupled with a brief discussion over a telephone call as done in the old system, nurses found it unproductive to have multiple emails being sent back and forth (Textbox 1, data extract 5). Instead, nurses found that telephone or face-to-face conversations allowed more detailed discussions compared with the short text of an email.

Residents

Residents also perceived significant efficiency with the use of smartphones. Since 42% of emails from nurses were informational items, no follow up response was required for these types of communication (Textbox 1, data extract 6). Residents also used the smartphones to reduce the inefficiencies of having to page other services to a ward telephone. By paging other physicians to their smartphone, they no longer needed to wait at a ward telephone for a callback and were able to perform other tasks while waiting for the return call (Textbox 1, data extract 7). To the residents, smartphones made it easier to coordinate activities within the teams through email or telephone calls (Textbox 1, data extract 8). The devices were also used to communicate with other physicians from other teams and services to increase communication around patients (Textbox 1, data extract 9). Residents appeared to incorporate the phone and messaging functions of the smartphone in their clinical work, using them to call a patient's family member, to communicate rapidly with other team members, and to respond to urgent situations (Textbox 1, data extract 10).

Textbox 1. Efficiency: extracts from interviews and field notes

1. I think the message is getting across better. The communication has been opened a lot so you don't have to always sit there on the phone and then they call back and you miss the phone call and then you have to call again. The information is directly there so they know exactly what it is if they want to respond to it immediately or when they have time. [Interview, nurse 13]

2. At 10:01:01, a nurse types [an email]: "Team 1: Call back requested--Message: pt continues to experience chest pain. Last 0.4mg nitro spray given at 8:30. Patient rating pain as 5 out of 10. Will order ECG. Please call unit." [Field notes, ward observations, Nov 27th]

3. A nurse would come to the nursing station and say, "Mrs. Jones' potassium is 2.7. Can you page the resident that's looking after this patient?" Now it makes it easier for us because we no longer have to search on the whiteboard or through the chart to see who is the covering physician because we just send out a general email to team (number) and the person who is carrying the BlackBerry will deal with this issue. They'll either forward it to the right person or they'll deal with it themselves. [Interview, nurse 1]

4. At 2:43, nurse 2 came to the ward clerk. She mentioned that no one was responding after 2 [emails] and so she is going to page. She first paged and then she called another person (name M) to leave a message that patient X's ECG abnormal results are back and want doctor 1 to know. After hanging up, she realized she could call the BlackBerry. "Oh…forgot about it. There are many ways to reach the doctor." She called doctor 1 on the BlackBerry. It turns out the doctor was at the stairway on his way up. Doctor 1 arrived at nursing station and said, "Sorry that I didn't call you back." Discussion for patient's case commences between nurse 2 and doctor 1. [Field notes, ward observations, Nov 2nd]

5. ...because maybe I just find sometimes some additional problems come up again and you have to be able to go through the whole system.... But I find that like if I read through the emails, they've emailed each other back several times. So if you're just able to pick up the phone and call them. I personally think that I've always been pro talking to someone [rather] than just technology. [Interview, nurse 1]

6. I really liked using the BlackBerrys because it very quickly communicated what the nurse was trying to contact... [Interview, resident 4]

7. Especially when you are trying to get in touch with specialists. You don't have to stick around the telephone to wait for a phone call back. You can do your things and the specialist can call you at his convenient time. [Interview, resident 2]

8. When your team needed to get together to discuss something or to run the list or whatever you'd send out five pages, which is ridiculous. You'd have to sit at the phone and wait for five people to call you back. This way you send one message to five people that says "Meeting on 13 in five minutes." And everyone just comes. " [Interview, resident 1]

9. Oftentimes I'll be consulting with another physician on a patient and I'll say, "This is my BlackBerry. Call me back after you've seen the patient or call me back when you have a plan."…So that's extremely valuable, which we never had with pages, and no one would ever page you for that because it was too much of a pain. [Interview, resident 1]

10. At 10:23, team BlackBerry goes off. Senior picks up and talks. She hangs up at 10:24 and goes to the elevator. She looks at the team BlackBerry and starts typing. She then calls using her senior's BlackBerry to a junior resident about going to the 7th floor. At 10:25, she makes another call on the senior's BlackBerry. She hangs up at 10:25 and goes to see a patient on the 7th floor who looks like he is choking/having a seizure. At 10:26, senior asks a nurse for a Yonker for suction and attends to the patient. Junior resident arrives at the patient's room. (Looks like an emergency situation: about 7 staff nurses and doctors were in the patient's room.) [Field notes, work shadowing physician 4]

Interruptions

Despite residents acknowledging that the use of smartphones had reduced the time spent on responding to informational messages, residents also perceived an increase in the overall number of messages and calls received. The substantial volume of interruptions was reflected in the number of phone calls and emails received as determined by both smartphone usage as well as by work shadowing residents (Table 2). From analyzing the frequency of calls and emails, a senior resident who usually carried both the team BlackBerry (point of contact and communication for nursing and allied health staff) and the senior BlackBerry (typically used for communication within the medical teams) would receive on average 11.6 calls and 25 emails within a 24-hour period. The volume of calls and emails could be higher during evenings and weekends when residents were covering for other teams and thus carrying multiple smartphones. (Textbox 2, data extract 1).

Interviews with the staff physicians also highlighted similar observations where physicians noticed the disruptions that the smartphone brought upon their teaching and patient care rounds (Textbox 2, data extract 2). From work-shadowing sessions, there were 4.6 average interruptions per hour for a resident when considering all direct calls, emails, and face-to-face communications (Table 2). These interruptions, however, were sometimes even more frequent. For example, within a 40-minute team meeting with the attending physician, 7 interruptions were observed. This included 5 direct phone calls to residents, where residents ended up taking the calls and leaving the room while the meeting was in progress (Textbox 2, data extract 3). In particular, numerous patient interactions were noted to be interrupted by direct calls. Often residents were observed answering phone calls on their smartphones, exiting the room, and then resuming patient interaction once the call was completed (Textbox 2, data extract 4). Residents did note that this had a negative impact on communication, especially during family meetings.

Considering these high volumes of communications, it is not surprising that residents commented that they felt overwhelmed by the constant interruptions, which may be a result of the increased availability of multiple communication options (Textbox 2, data extracts 5-7). Similar sentiments were echoed by the staff physicians, who observed that these interruptions could have detrimental impacts such as reduced downtime and impede residents’ abilities to provide patient care (Textbox 2, data extract 8).

Textbox 2. Interruptions: extracts from interviews and field notes

1. *In July…and I carried six of [the BlackBerrys] and then there were no rules at all so nurses were just calling them and four would be going off at once and you couldn’t get anything done.* [Interview, resident 3]
2. *So that intrusiveness I definitely find when we’re at the bedside, when we’re teaching, or even when we interact around cases and they go “‘Oh, I’m sorry. I better take this.”…So it’s a two-edged sword.* [Interview, attending physician 3]
3. *At 3:04, attending for the team walks into the meeting room. At 3:07, junior resident’s BlackBerry rings. She picks up the smartphone and walks out to talk (sounds like a patient’s issue). She returns to the room at 3:10. At 3:13, team members go through patients’ cases with the social worker, pharmacist, and the attending physician. At 3:14, the team BlackBerry rings. Senior resident picks up. She looks at the BlackBerry and then starts calling back. As she walks out of the room, she says, “Hi, it is team (number). Who paged?” At 3:14, another junior resident’s BlackBerry goes off. He leaves the room but returns quickly.* [Field notes, work shadowing physician 2]
4. *Senior resident returns to the patient’s room and continues examining her. While in the patient’s room, I (observer) could hear the resident talking on the BlackBerry. I asked her later what calls she had while in the room. It turns out she had 3 phone calls and 2 texts. Two of the calls were from the radiation oncologists and 1 call from the pathologist. She also received 1 text on the team BlackBerry and 1 text on the senior’s BlackBerry from the pharmacist.* [Field notes, work shadowing physician 11]
5. *The only negative I can think of is just the incredible number of communications that you get, you know, text messages and emails and everything else. So just the number can sometimes be overwhelming.* [Interview, resident 1]
6. *At 8:40, resident #1 talks to resident #2. Resident #1 complains that he got about 1000 pages after he had to take over.* [Field notes, ward observations, March 18th]
7. *There was no choice [before]. Now there’s a choice to page; there’s a choice to text. You can ask for no response, email response, call back response, or call. So there’s six choices, right? There’s probably more but that’s sort of the ones that I’ve been using...* [Interview, resident 3]
8. *I recognize that it comes at a significant cost…because other people are interrupting them and it probably comes at a personal cost to them in terms of the need for increased vigilance and attention and less downtime for them, like there are even more intrusions in their lives.* [Interview, attending physician 3]

Interprofessional Relationships

A strong theme that emerged throughout interviews was the impact of this new technology on interprofessional relationships. Nurses commented that the new system reduced opportunities for face-to-face interactions, which many valued. Nurses reported that they found it more difficult to build interprofessional relationships through the new technology. Specifically, the smartphones and use of email messaging reduced verbal conversations, which nurses felt prevented them from getting to know the residents, discouraged interest in their work, and reduced opportunities for nurses to have direct educational experiences with the residents (Textbox 3, data extracts 1 and 2). Additionally, nurses found the process depersonalizing to have the team BlackBerry as the primary point of contact instead of having direct interactions with the specific physician (Textbox 3, data extracts 2 and 3).

In contrast, physicians perceived no major negative changes with this new technology. Since nurses were required to type in their names to send emails, physicians felt that they learned nurses’ names better and thus perceived their interprofessional relationships had actually improved (Textbox 3, data extracts 4 and 5).

Textbox 3. Interprofessional relationships: extracts from interviews and field notes

> *1. With [general internal medicine] it's really hard enough that the residents change every month or every 4 weeks so it's hard to build that relationship with them in terms of what knowledge they have and even them knowing who we are. And then on top of that with the BlackBerry system…it is convenient, but in terms of building the team dynamics, because we are focused on interdisciplinary care, it's hard to build that when a lot of it is through technology.* [Interview, nurse 9]
>
> *2. I know some of the nurses sort of have complaints or have concerns about the [new] paging [system], you lose the face-to-face communication, you lose getting to know the residents, really. Right now they just know everybody as a team. "'Oh, I've sent the team this." They miss that, especially the older nurses who are used to communicating face-to-face and getting to know the physicians more on a personal level.* [Interview, nurse 2]
>
> *3. Nurse types an email, "To team (number): Email response requested. Message: Thank you for replying back. Can I get your name so that I can write a verbal order in chart and how often to check [blood capillary glucose]? Patient said he doesn't take insulin just the metformin and gliclazide. Thank you."* [Field Notes, ward observations, February 20th]
>
> *4. What's great about the email system is that you have the nurse's name cause it's really hard because we work with a lot of nurses. And sometimes— I always introduce myself to the nurses, but they never give me their names back, so it's nice to have that (BlackBerry) in front of you, and then if you forget the nurse you can just check the name again, so I find it much easier to work with people if you know their names."* [Interview, resident 3]
>
> *5. You know, I'd actually say it's maybe helped interaction with nurses only because when they [email] and they put their name. Like in the past, like it's easy to not know nurses' names because there's so many of them, but when they're emailing and they say like, you know, "Please reply to" and then it has the name like Joan, then you can go to the ward and say, "Hi Joan, I got your message. Thanks for sending it." And I actually liked it. I got to know nurses' names actually better through it…"* [Interview, resident 4]

Gap in Perceived Urgency

From our data analyses, we identified a gap in what physicians and nurses perceived as urgent patient issues. If a physician did not perceive the issue communicated to be urgent, often there would not be a response despite a request to respond having been made by the nurse, or the resident would reply with an email when the nurse requested a telephone call or otherwise.

Nurses

Nurses perceived both a lack of acknowledgment of messages as well as not receiving the requested response (Textbox 4, data extracts 1-3). This perception was confirmed by analyzing responses to nurses' emails in which nurses actually only received an email response 50% of the time requested. They also felt many physicians ignore emails, similar to their ignoring pages previously. Nurses found it frustrating and felt belittled when physicians ignored their communications (Textbox 4, extract 4). With a lack of acknowledgment, nurses often felt the need to resend messages. Nurses also reported the need to have clear specifications of when and how to inform physicians such as for abnormal vital signs or laboratory results. It was observed that nurses would consult each other to see if physicians should be informed (Textbox 4, data extract 5).

Residents

Physicians also commented that there were too many direct calls for low importance items and a high number of emails that were of trivial importance (Textbox 4, data extract 6).

Much of this gap of perceived urgency may be attributed to the numerous methods that are now available to contact physicians. Though the smartphone system provides various options to contact residents, the array of choices often created confusion and a mismatch of responses among clinicians (Textbox 4, data extract 7).

Textbox 4. Gap in perceived urgency: extracts from interviews and field notes

> *1. But, for example, if I need to get an order for a medication I would do an email, but if it's something like the patient or the family wants to speak to the doctor and they really need a time right away or something urgent like that, I would ask them to call back. But then I find the doctors don't always call back, anyways. They just use the email system.* [Interview, nurse 9]
>
> *2. At 11:38, nurse 4 typing an email on computer #1: "Team (number), call back requested: I just received your order on (dosage) medication X. Would you like to discontinue the IV medication X? "Page sent at 11:39am 30 seconds. At 12:10pm, nurse 4 approached me (observer) and told me her webpage for a callback was responded with an email from the doctor. She said she did not check her inbox. She was informed by another nurse who saw the email (because the message that nurse 4 sent had her name on it) and that's how the other nurse knows to tell nurse 4.* [Field notes, ward observations, Nov 25th]

3. Yeah, it makes me wonder that maybe I shouldn't be sending that information or [laughs] if the patient wants something, for example, they want to see the doctor to discuss test results or something, and I don't hear back, like even just a simple acknowledgment like, uh, "Okay, we'll try to see them sometime this afternoon" or something would be nice instead of just sending something out into the void. [Interview, nurse 3]

4. There's nothing worse than, like, if I'm sending a message and I think it's important…like it might not be that important to the physician 'cause they know the case more in depth, but I might think it's quite important, or it's something that's really important to the patient. But it's sometimes hard, it's hard to convey that urgency through a written message sometimes, and I don't like to keep sending another page over and over again. It feels like I'm annoying the physician probably [laughs], but if I don't get that response, it makes me second guess myself, like okay, I guess it doesn't deserve a response or… [Interview, nurse 3]

5. Nurse E: "I got the hemoglobin results, it is X; should I email the team about it?" Nurse D says she should not and explains why. [Field notes, ward observations, February 20th]

6. …I've been having some calls for bowel movement problems or for sleeping pills or something; very minor stuff that I think could at least be dealt with, with an email. [Interview, resident 2]

7. ...you know, there's 500 ways to contact a person. You get their email, you get their phone number, you get their pager number, you get face-to-face. So I think to establish the best way to contact somebody, why not just face-to-face or BlackBerry, you know. You can use the phone on the BlackBerry, or you can email on the BlackBerry. I think there's confusion in the unit on how to communicate or the best way to find certain people, and that's when things take longer to happen…I think we're given so many ways to communicate with people sometimes that we don't know which one is the best one or the most efficient to get in touch with them. [Interview, allied health 6]

Professionalism

This theme focused on physicians' behavior around the use of smartphones and how nurses, allied health professionals, and attending physicians perceived the manner in which residents handled interruptions by phone call or email message that could be regarded as unprofessional.

While having a conversation with a resident on the ward, nurses and other allied health professionals occasionally reported that the resident would answer a call on his or her smartphone, thus interrupting their conversation. Nursing and allied health professionals also found it disruptive when residents answered calls during interprofessional rounds (Textbox 5, data extracts 1 and 2).

Similarly, during patient rounds or educational rounds, attending physicians mentioned that they also found residents' behavior with smartphones to be unprofessional at times (Textbox 5, data extracts 3 and 4). One attending physician observed that while smartphones increased the availability of the residents to other clinicians, it reduced the local availability of residents due to the constant distractions and interruptions from the device. In effect, the residents were made more global but less local (Textbox 5, data extracts 5 and 6). As noted by some, the continued calls and checking of messages often took away the quality time which residents spent interacting with the attending physician and their interprofessional colleagues.

Finally, this behavior was also observed during patients' interactions, where residents would pick up phone calls, check, or type messages while talking with patients or supervising a procedure. Although patients' perceptions were not obtained, other clinicians commented that such behavior could create negative perceptions among patients (Textbox 5, data extracts 7 and 8).

Textbox 5. Professionalism: extracts from interviews and field notes

1. I did have a couple of doctors from other teams that were just constantly chatting on the phone while communicating with nursing staff. I found that to be kind of unprofessional. [Interview, nurse 14]

2. In rounds or orientating new residents, I find that, yes, the BlackBerry does go off, and it rings, and it can interrupt face-to-face communication sometimes or that residents are checking emails as they're talking to people face-to-face. So that's also where you're not sure if they're really listening to what you're saying or not. [Interview, allied health 6]

3. And sometimes it may indicate they're being quite unsubtle about indicating that they're not interested in being present all the time. That they're bored. But I think it's quite rude, and I think it's undisciplined. [Interview, attending physician 3]

4. It's probably more annoying than anything. This is true of BlackBerrys, anyone using BlackBerrys. Anyone who has a BlackBerry will talk to you and look you in the eye and then kind of look down towards their BlackBerry. [Interview, attending physician 1]

5. There's definitely a convenience around, for example, their ability to be able to page other services, so if they need to get a hold of one of the consulting services they can page them and walk around and be able to pick it up instantly anywhere and not have to go back to the desk and things like that. But I think it also may limit the depth of their ability *to interact with anybody else around them because they're always sort of being distracted and it happens every single hour if not more. It happens constantly.* [Interview, attending physician 3]

6. And I think almost to some extent it's an implicit permission that gets granted to the house staff to disrupt their own teaching experience and disrupt others around them because everybody is doing it, because everybody is being "BlackBerried." [Interview, attending physician 3]

7. I think most of the time, the BlackBerry is seen as a nuisance and a disruptive factor. And I think that most patients would not be terribly impressed no matter how much time you spent explaining why you need it. [Interview, attending physician 3]

8. Senior walks out of the patient's room while typing on the BlackBerry. She finishes typing and returns to the room at 5:36. Senior looks at her BlackBerry and starts typing inside the room in front of the patient. She pauses to look at the patient and the resident doing the procedure [paracentesis]. She resumes texting again and walks out of the room at 5:38. Another resident walks out, and senior speaks with the resident. Senior returns to the room and speaks with the patient. She asks the patient if he has ever gotten a successful tap before. Senior looks at her BlackBerry and starts typing. [Field notes, work shadowing physician 2]

Discussion

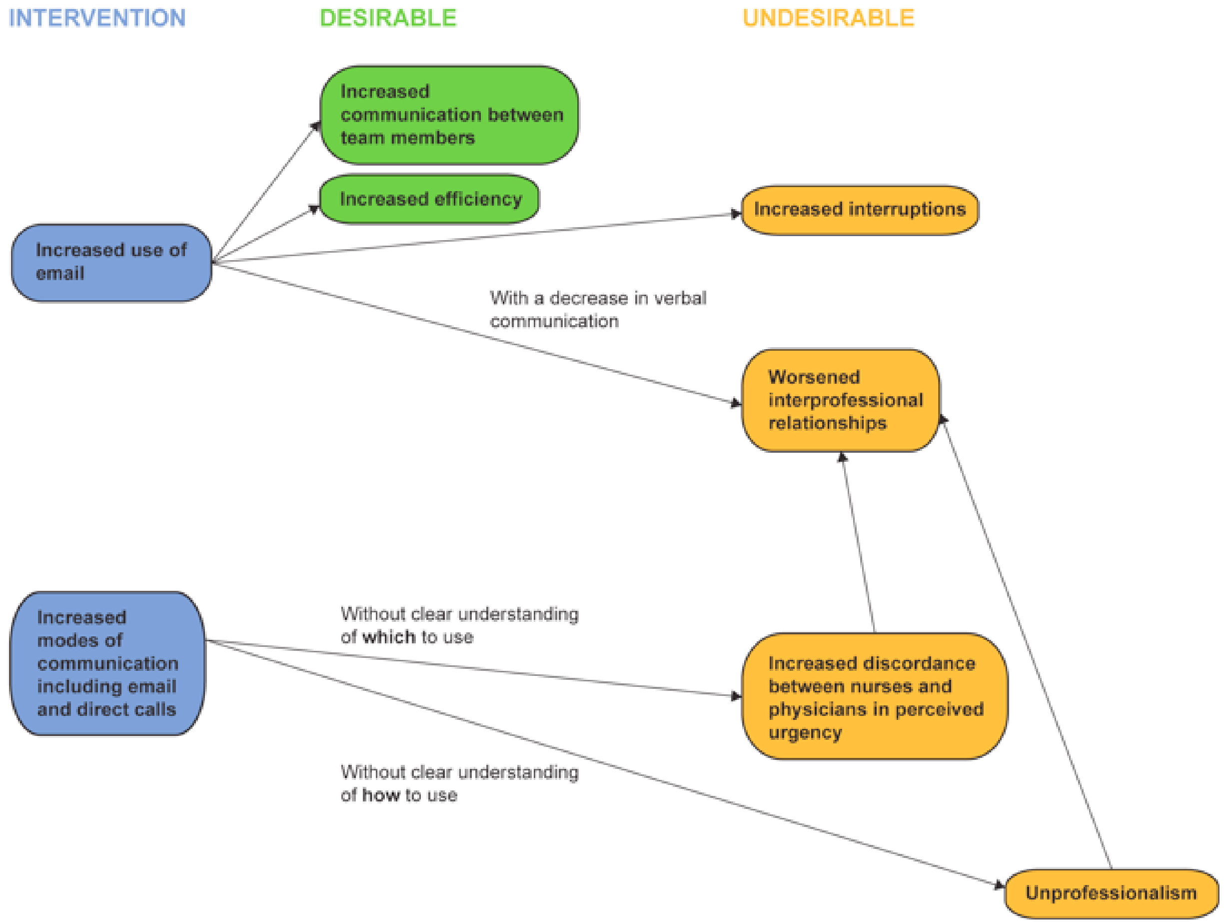
Principal Results

In this study, we described the nature of communications and the perceived impact of a new smartphone system for residents on busy general medical services. Smartphones were used frequently by residents to communicate about patients. We found that residents, nurses, and other clinicians perceived improvements in efficiency over the traditional paging system likely because smartphones appear to address many of the known issues with numeric paging such as inability to triage information or having to wait by a phone for a return call. The overall number of interruptions, however, was perceived to have increased likely because it was now easier for nurses to initiate communications. With more communications occurring over written modes such as emails and text messages, nurses perceived a negative impact to relationships with physicians because of decreased verbal conversations. Smartphone use in this study also highlighted the discordance between what nurses and physicians perceive as urgent. Finally, residents' behavior when communicating with smartphones during patient care and education learning activities was perceived at times as unprofessional.

While the gap between what nurses and physicians perceive to be urgent has been previously identified [18], we found that the gap was likely made greater by increasing the number of methods for clinical communication. Email communication may also decrease verbal interactions—deemed by many clinicians as the most effective and optimal method of communication—which can result in the impediments of interprofessional collaboration and relationships [19]. Finally, the issue of digital professionalism is part of a larger issue of both knowing and educating medical trainees in the proper use of new media [20,21]. While there are established definitions of medical professionalism, these have not necessarily kept pace with the rapidly evolving new digital media [22-25]. Without being given formal guidance in using this new technology, residents appear to handle interruptions by trying to be efficient and minimize response times but at the expense of interprofessional and patient relationships.

Our findings highlight the important aspects to consider when implementing systems to improve clinical communication. While increasing overall team communication and efficiency are important goals, it is also equally critical to consider the range of possible and potential impacts as unintended consequences can occur (Figure 2). Considering these themes (efficiency, interruptions, interprofessional relationships, common understanding of urgency, and professionalism) as a framework could aid in the design and evaluation of communication systems. This may facilitate the development of communication systems in which the appropriate information is sent through the appropriate medium with the appropriate intrusiveness, while incorporating processes that foster interprofessional relationships and promote digital professionalism.

Figure 2. Potential positive and negative effects of changes to communication system in the use of smartphones



As shown in Figure 2, our findings suggest that email or text conversations appear to create the desirable and undesirable effects as listed above. The increased modes of communication when moving from pagers to smartphones include both text messages as well as direct phone calls. Our findings support that increased modes of communication appear to highlight the gap in perceived urgency and highlights professionalism issues.

Comparison With Other Work

The issues with numeric paging have been well documented. These include a high number of interruptions with an average of 9 pages per hour on some clinical services [26]. While all pages interrupt, only 30% have been found to require urgent attention, and the majority do not require a response within an hour [27]. As well, one study found that 14% of pages are sent to the wrong physician, with 47% of those pages requiring urgent attention [7]. The response rate for pages in one observational study was 90% [4].

To our knowledge, this is the first comprehensive mixed-methods evaluation of the use of smartphones on hospital communications. Studies using surveys have found perceived improvements with the use of smartphones among clinicians [15,16]. Through the use of ethnography observations and interviews, we were able to determine *how* smartphones are used and their impacts on different domains including unintended effects.

Limitations

There were limitations to this study. Smartphones were provided only to residents, and the behavior and perceptions would likely be different if other professions were provided with smartphones. The study was also conducted at two hospital sites; thus, generalizing to other institutions with different hospital cultures may yield different results. However, our intervention used standard components of smartphones and email with minimal customization, and other academic hospitals may be able to learn from this experience. Future studies could look at other measurements, outcomes, and impacts on quality of patient care such as time to resolution of urgent items or the effect of interruptions on patient care.

Conclusions

Routine adoption of smartphones by residents appeared to improve efficiency over the use of pagers for physicians, nurses, and allied health professionals. This was balanced by negative communication issues of increased interruptions, a gap in perceived urgency, weakened interprofessional relationships, and unprofessional behavior. Further communication interventions are required that balance efficiency and interruptions while maintaining or even improving interprofessional relationships and professionalism.

Acknowledgments

We would like to thank the Ontario Ministry of Health and Long Term Care Alternate Funding Proposal Innovation Grant for providing funding. We would like to thank Christine Plaza for help with designing the figures.

Conflicts of Interest

None declared

References

1. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
2. Rivera-Rodriguez AJ, Karsh BT. Interruptions and distractions in healthcare: review and reappraisal. Qual Saf Health Care 2010 Aug;19(4):304-312. [doi: 10.1136/qshc.2009.033282] [Medline: 20378621]
3. Coiera E. When conversation is better than computation. J Am Med Inform Assoc 2000;7(3):277-286. [Medline: 10833164]
4. Harvey R, Jarrett PG, Peltekian KM. Patterns of paging medical interns during night calls at two teaching hospitals. CMAJ 1994 Aug 1;151(3):307-311. [Medline: 8039084]
5. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991 Feb 7;324(6):370-376 [FREE Full text] [doi: 10.1056/NEJM199102073240604] [Medline: 1987460]
6. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Hamilton JD. Quality in Australian Health Care Study. Med J Aust 1996 Jun 17;164(12):754. [Medline: 8668084]
7. Wong BM, Quan S, Cheung CM, Morra D, Rossos PG, Sivjee K, et al. Frequency and clinical importance of pages sent to the wrong physician. Arch Intern Med 2009 Jun 8;169(11):1072-1073. [doi: 10.1001/archinternmed.2009.117] [Medline: 19506178]
8. Collins S, Currie L, Bakken S, Cimino JJ. Interruptions during the use of a CPOE system for MICU rounds. AMIA Annu Symp Proc 2006:895. [Medline: 17238514]
9. Collins S, Currie L, Patel V, Bakken S, Cimino JJ. Multitasking by clinicians in the context of CPOE and CIS use. Stud Health Technol Inform 2007;129(Pt 2):958-962. [Medline: 17911857]
10. Brixey JJ, Tang Z, Robinson DJ, Johnson CW, Johnson TR, Turley JP, et al. Interruptions in a level one trauma center: a case study. Int J Med Inform 2008 Apr;77(4):235-241. [doi: 10.1016/j.ijmedinf.2007.04.006] [Medline: 17569576]
11. Laxmisan A, Hakimzada F, Sayan OR, Green RA, Zhang J, Patel VL. The multitasking clinician: decision-making and cognitive demand during and after team handoffs in emergency care. Int J Med Inform 2007;76(11-12):801-811. [doi: 10.1016/j.ijmedinf.2006.09.019] [Medline: 17059892]
12. Hickam DH, Severance S, Feldstein A, Ray L, Gorman P, Schuldheis S, et al. The effect of health care working conditions on patient safety. Evid Rep Technol Assess (Summ) 2003 Mar(74):1-3. [Medline: 12723164]
13. Soto RG, Chu LF, Goldman JM, Rampil IJ, Ruskin KJ. Communication in critical care environments: mobile telephones improve patient care. Anesth Analg 2006 Feb;102(2):535-541 [FREE Full text] [doi: 10.1213/01.ane.0000194506.79408.79] [Medline: 16428557]
14. Derbyshire SW, Burgess A. Use of mobile phones in hospitals. BMJ 2006 Oct 14;333(7572):767-768. [doi: 10.1136/bmj.38995.599769.80] [Medline: 17038714]
15. Wu RC, Morra D, Quan S, Lai S, Zanjani S, Abrams H, et al. The use of smartphones for clinical communication on internal medicine wards. J Hosp Med 2010;5(9):553-559. [doi: 10.1002/jhm.775] [Medline: 20690190]
16. O'Connor C, Friedrich JO, Scales DC, Adhikari NK. The use of wireless e-mail to improve healthcare team communication. J Am Med Inform Assoc 2009;16(5):705-713. [doi: 10.1197/jamia.M2299] [Medline: 19567803]
17. Reeves S, Kuper A, Hodges BD. Qualitative research methodologies: ethnography. BMJ 2008;337:a1020. [Medline: 18687725]
18. Coiera E, Tombs V. Communication behaviours in a hospital setting: an observational study. BMJ 1998 Feb 28;316(7132):673-676 [FREE Full text] [Medline: 9522794]
19. Kock N. The ape that used email: Understanding e-communication behavior through evolution theory. Commun Assoc Inf Syst 2001;5(3):1-29.
20. Parikh SM, Liu E, White CB. Connectivity need not come at the expense of professionalism. Acad Med 2010 Jun;85(6):930; author reply 930. [doi: 10.1097/ACM.0b013e3181dbe54b] [Medline: 20505384]
21. Spiegelman J, Detsky AS. Instant mobile communication, efficiency, and quality of life. JAMA 2008 Mar 12;299(10):1179-1181. [doi: 10.1001/jama.299.10.1179] [Medline: 18334693]
22. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach 2007 Sep;29(7):642-647. [doi: 10.1080/01421590701746983] [Medline: 18236250]
23. Swick HM. Toward a normative definition of medical professionalism. Acad Med 2000 Jun;75(6):612-616. [Medline: 10875505]

24. ABIM Foundation. American Board of Internal Medicine, ACP-ASIM Foundation. American College of Physicians-American Society of Internal Medicine, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. Ann Intern Med 2002;136(3):243-246 [FREE Full text] [Medline: 11827500]
25. Jain SH. Practicing medicine in the age of Facebook. N Engl J Med 2009 Aug 13;361(7):649-651 [FREE Full text] [doi: 10.1056/NEJMp0901277] [Medline: 19675328]
26. Chiu T, Old A, Naden G, Child S. Frequency of calls to "on-call" house officer pagers at Auckland City Hospital, New Zealand. N Z Med J 2006;119(1231):U1913. [Medline: 16582974]
27. Patel R, Reilly K, Old A, Naden G, Child S. Appropriate use of pagers in a New Zealand tertiary hospital. N Z Med J 2006;119(1231):U1912. [Medline: 16582973]

Edited by G Eysenbach; submitted 24.09.10; peer-reviewed by J Ai, P Chhanabhai; comments to author 20.02.11; revised version received 21.07.11; accepted 21.07.11; published 29.08.11.

Please cite as:
Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, Cheung M, Morra D
An Evaluation of the Use of Smartphones to Communicate Between Clinicians: A Mixed-Methods Study
J Med Internet Res 2011;13(3):e59
URL: http://www.jmir.org/2011/3/e59/
doi:10.2196/jmir.1655
PMID:21875849

Original Paper

Harnessing Context Sensing to Develop a Mobile Intervention for Depression

Michelle Nicole Burns[1], PhD; Mark Begale[1], BA; Jennifer Duffecy[1], PhD; Darren Gergle[2], PhD; Chris J Karr[3], MA; Emily Giangrande[1], BS; David C Mohr[1], PhD

[1]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States
[2]Department of Communication Studies, Northwestern University, Evanston, IL, United States
[3]Audacious Software, Chicago, IL, United States

Corresponding Author:
David C Mohr, PhD
Department of Preventive Medicine
Feinberg School of Medicine
Northwestern University
680 N Lakeshore DR #1220
Chicago, IL, 60611
United States
Phone: 1 312 503 1403
Fax: 1 312 503 0982
Email: d-mohr@northwestern.edu

Abstract

Background: Mobile phone sensors can be used to develop context-aware systems that automatically detect when patients require assistance. Mobile phones can also provide ecological momentary interventions that deliver tailored assistance during problematic situations. However, such approaches have not yet been used to treat major depressive disorder.

Objective: The purpose of this study was to investigate the technical feasibility, functional reliability, and patient satisfaction with Mobilyze!, a mobile phone- and Internet-based intervention including ecological momentary intervention and context sensing.

Methods: We developed a mobile phone application and supporting architecture, in which machine learning models (ie, learners) predicted patients' mood, emotions, cognitive/motivational states, activities, environmental context, and social context based on at least 38 concurrent phone sensor values (eg, global positioning system, ambient light, recent calls). The website included feedback graphs illustrating correlations between patients' self-reported states, as well as didactics and tools teaching patients behavioral activation concepts. Brief telephone calls and emails with a clinician were used to promote adherence. We enrolled 8 adults with major depressive disorder in a single-arm pilot study to receive Mobilyze! and complete clinical assessments for 8 weeks.

Results: Promising accuracy rates (60% to 91%) were achieved by learners predicting categorical contextual states (eg, location). For states rated on scales (eg, mood), predictive capability was poor. Participants were satisfied with the phone application and improved significantly on self-reported depressive symptoms ($\text{beta}_{week} = -.82$, $P < .001$, per-protocol Cohen $d = 3.43$) and interview measures of depressive symptoms ($\text{beta}_{week} = -.81$, $P < .001$, per-protocol Cohen $d = 3.55$). Participants also became less likely to meet criteria for major depressive disorder diagnosis ($b_{week} = -.65$, $P = .03$, per-protocol remission rate = 85.71%). Comorbid anxiety symptoms also decreased ($\text{beta}_{week} = -.71$, $P < .001$, per-protocol Cohen $d = 2.58$).

Conclusions: Mobilyze! is a scalable, feasible intervention with preliminary evidence of efficacy. To our knowledge, it is the first ecological momentary intervention for unipolar depression, as well as one of the first attempts to use context sensing to identify mental health-related states. Several lessons learned regarding technical functionality, data mining, and software development process are discussed.

Trial Registration: Clinicaltrials.gov NCT01107041; http://clinicaltrials.gov/ct2/show/NCT01107041 (Archived by WebCite at http://www.webcitation.org/60CVjPH0n)

(J Med Internet Res 2011;13(3):e55) doi:10.2196/jmir.1838

KEYWORDS

Depression; behavior therapy; telemedicine; mobile health; mobile phone; cellular phone; sensors; data mining; artificial intelligence; context-aware systems

Introduction

Major depressive disorder affects nearly 7% of the population annually [1] and is the leading cause of disease burden in middle- and high-income countries worldwide [2]. Individuals with major depressive disorder have higher medical costs [3], exacerbated medical conditions [4], and mortality almost twice that of nondepressed people [5]. Thus, major depressive disorder poses an extraordinary public health problem in terms of prevalence, cost, morbidity, and mortality.

While psychological treatments for depression can be effective [6], they are often plagued by access barriers and high rates of attrition [7,8]. Internet interventions have been touted as an antidote to access barriers, but they appear to produce more modest outcomes [9], in part also due to high attrition [10]. Mobile phones, however, have penetrated nearly all strata of society [11,12] and hold promise as a ubiquitous treatment platform through which the connection between patient and intervention is continuous and reciprocal.

At least two broad classes of treatment components can be delivered via mobile phone. First, these platforms offer the opportunity to deliver interactive tools to patients in their environment. These tools, sometimes referred to as ecological momentary intervention [13], can prompt patients to input information about their situation or internal states, and provide in-the-moment responses personalized to a patient's immediate needs. While such interventions have been explored with regard to health behaviors and severe mental illness (eg, bipolar disorder [14]), recent reviews reveal no such work in unipolar depression [15,16].

Second, mobile systems also have the potential to apply machine learning techniques that can monitor and learn to recognize a patient's circumstances and state. Smartphones contain numerous sensors (eg, global positioning system [GPS], Bluetooth) that could provide clues to patient states and contexts. Smartphones also have the ability to conduct ecological momentary assessment and allow patients to report, or "label," their current states. Machine learning, or data mining techniques, can be used to automatically learn the relationship between these two sources of data. This relationship is captured in what is known as a "learner" that can then be used to develop individualized predictions of patient states solely from low-level sensor data. Once trained using ecological momentary assessment data, the learner could potentially identify patient states continuously and passively, with little effort from the patient. Sensor-based awareness of patient states would enable context-appropriate clinical responses [18,19] (eg, delivering timely feedback, providing guidance during distress or problematic situations, or intervening on nonadherence to homework) without relying on the patient to initiate such therapeutic interactions.

A few such context-aware systems have been developed and tested in mHealth interventions. The Intel Mobile Heart Health prototype uses data from sources such as mobile phone-based ecological momentary assessment and a small electrocardiograph sensor with accelerometer to detect changes in heart rate variability, activity, and mood. If individualized threshold values are reached, the mobile phone delivers cognitive behavioral and mindfulness techniques designed to reduce stress [20]. Although used for social networking rather than health care purposes, the CenceMe mobile phone application has achieved promising accuracy rates within a small sample in predicting socially relevant states such as whether the user was conversing [21]. CenceMe also used such data to generate higher-level descriptions of users' recent behavioral or lifestyle patterns (eg, "party animal," which was determined through presence at parties and number of social interactions). Learners detected users' basic activity level (eg, sitting, running). In the e-SENSE [22] project, body sensors allowed collection of a rich set of physiological data used to infer mood. Predicted mood was then integrated into users' instant messaging chats. However, the required electrodes and belt did not approach the convenience of a context-sensing platform that relies solely on mobile phone sensors. These projects are in very early stages of development, and we are aware of no trials that formally evaluated clinical outcomes.

The current study aimed to maximize the potential of mobile interventions to target depression. We developed a mobile phone intervention, called Mobilyze!, that includes the capacity to deliver ecological momentary intervention. We also developed and piloted a context-aware system to identify participants' location, activity, social context, mood, emotions, and cognitive/motivational states. The mobile intervention was supported by a website, which allowed access to lessons, tools, and graphical feedback on participants' states. The intervention was supported by a manualized telephone coaching protocol [23,24] to enhance adherence.

The aims of this field trial were to evaluate the technical feasibility and reliability, functional reliability, and acceptability of this system. Secondarily, we measured depression outcomes, as well as changes in anxiety symptoms due to their frequent co-occurrence with depression.

Methods

Study Design

This was a single-arm field trial of Mobilyze!, an 8-week multimodal intervention for depression, that included 1) mobile phone sensing and ecological momentary intervention, 2) an interactive website for behavioral skills training, and 3) email and telephone support from a coach assigned to each participant. The trial was approved by the institutional review board at Northwestern University (Chicago, IL, USA).

Participants

Participants were recruited through online advertising venues such as Google AdWords and Craigslist. The advertisements directed individuals to the lab webpage, where interested individuals completed an online screener and provided their contact information. Those who met initial criteria were scheduled for a telephone eligibility interview and emailed a link to a digital consent form. At the beginning of this interview, staff discussed with participants each section of the consent form, which included sections informing users as to how their mobile phone sensor data would be collected, de-identified, used, and stored. Participants' verbal consent was then obtained. Receipt of the electronically signed digital consent form was also required prior to enrollment in the trial.

Inclusion criteria were a diagnosis of major depressive disorder using the telephone-administered Mini-International Neuropsychiatric Interview [25,26], a score of 11 or more (Table 3 in [27]) on the 16-item Quick Inventory of Depression Symptoms–Clinician Rated (QIDS-C) [28], and a score of 10 or more [29] on the 8-item Patient Health Questionnaire (PHQ-8) [30]. All participants spoke and read English, were at least 19 years of age, lived in the United States, reported being within a cellular network range most of the day, and reported having an email account, computer, broadband Internet access, and comfort using the Internet and mobile phones. Exclusion criteria were sensory impairments that would prevent participation, dementia (defined as scoring <25 on the Telephone Interview for Cognitive Status [31]), current participation in psychotherapy, initiation of an antidepressant medication in the past 10 days, or a severe psychiatric condition, as measured by the Mini-International Neuropsychiatric Interview, that rendered the intervention inappropriate (eg, psychotic disorders, bipolar disorders, severe agoraphobia, severe suicidality, and current harmful levels of substance or alcohol use).

Treatment

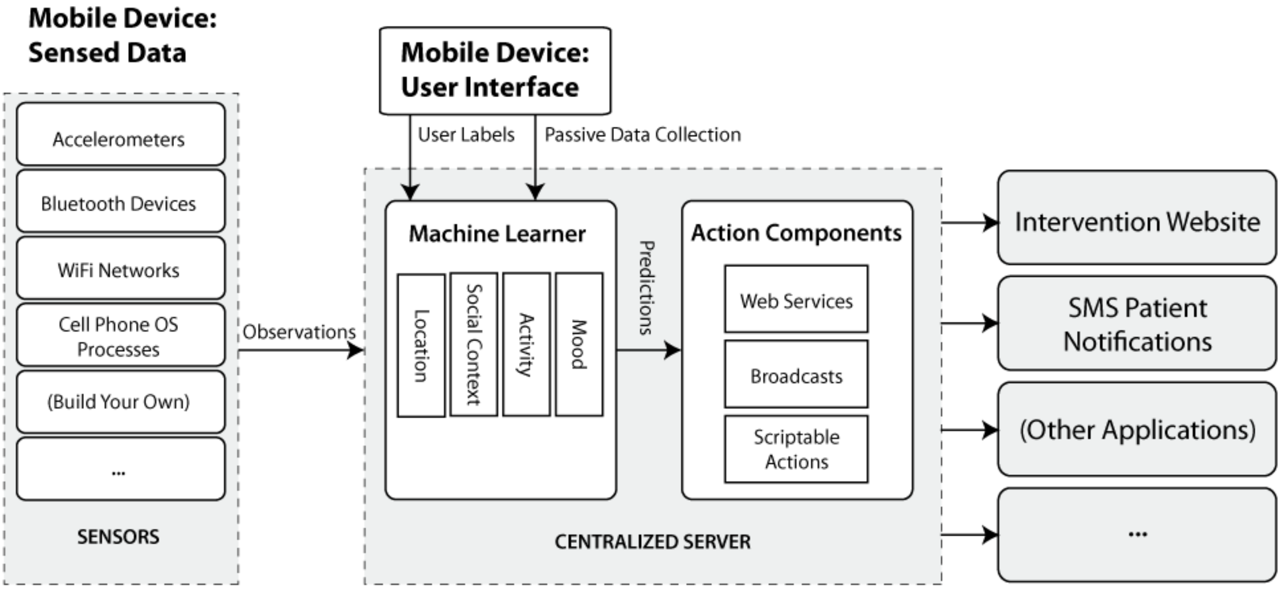
The treatment model was based on a behavioral activation approach [32,33] involving engagement in positive activities, experimentation with behavioral coping strategies, and use of such coping skills to reduce depressogenic avoidance behaviors. The mobile phone was used to translate this therapy into real time and to pilot a context-aware system that aimed to passively (ie, without requiring patient initiation) identify and respond to patient states.

Participants completed 8 weeks of the Mobilyze! intervention. Enrolled participants received at no charge 1) temporary use of a Nokia 5800 XpressMusic mobile phone with accessories, which cost the study US $547.88 per unit, 2) a cellular service plan, including voicemail and unlimited data usage, voice minutes, and text messaging, through T-Mobile at US $50 per month, per participant, and 3) login credentials to the website. Mobile phones were mailed to participants or made available for in-person pick up, according to participant preference. After the intervention concluded, phones were returned by participants via prepaid envelopes.

Context-Aware System

Our context-aware system used an architecture comprising 3 phases [34] (see Figure 1). In phase 1, sensors housed on the mobile phone gathered observations about the participant and their environment. These observations were transmitted to a secure server hosting a learner. In phase 2, the learner used an algorithm to inductively "learn" the relationship between sensor data and the participant's reported social context, activity, location, and internal states [35]. The learner could then predict participant states based on sensor data, and these predictions were passed to the action components in phase 3. The action components provided mechanisms for relaying predictions to other external outreach applications. In the current study, the action component consisted of the phone application itself, which displayed predicted states to the participant. The architecture, however, can be extended in the future to trigger outreach events based on the predictions (eg, sending a short message service [SMS, ie, text] notification, updating a feedback graph on a website, or notifying the coach via email).

Figure 1. A mobile phone-driven context-aware system (OS = operating system, SMS = short message service)



In this project, we used the mobile phone as the primary sensing platform. We also made use of a service-oriented architecture [36] for context-aware computing [37]. While the mobile phone provided the primary sensing platform, it transmitted sensor data securely via encrypted, password-protected tunnels to a variety of server-based components that provided logging, learning, and prediction services. Since these components communicated using extensible messaging and presence protocol (XMPP) [38], a common network protocol, this design allowed us to distribute the individual components among various hardware servers (providing scalability) while providing a common protocol for adding additional services (providing extensibility). We will now discuss the specific implementation of the context-aware system used in the current study.

Phase 1: Data Collection

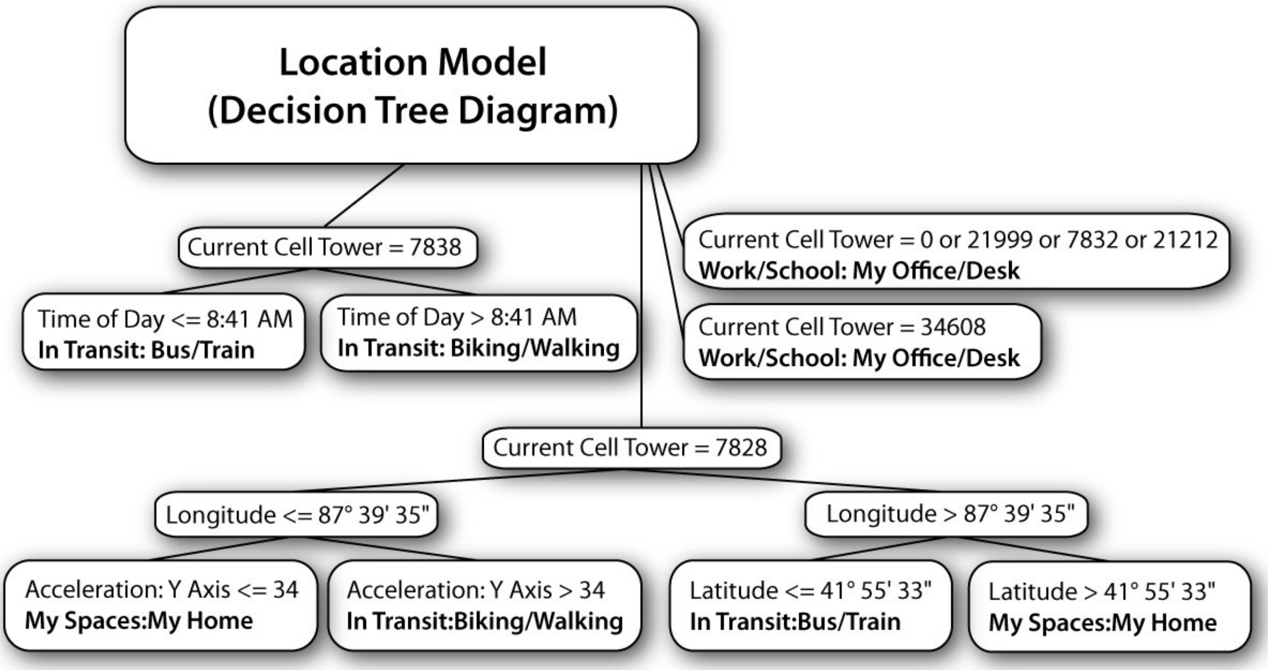
Contextual data were acquired from a collection of 38 sensors (see Multimedia Appendix 1) or more depending on the number of proximal Bluetooth devices or open applications on the mobile phone. Some of these data were collected directly from phone sensors, including GPS, Wi-Fi, Bluetooth detection of other wireless devices (eg, personal computers, some video game consoles), accelerometer, and ambient light. Other contextual data were inferred by the phone application from information available on the phone. This included time/day and activities of the phone's operating system (eg, recent calls, active phone applications).

The phone application transmitted encrypted sensor data to the secure backend server, via the XMPP interface protocol. The backend server logged the information by inserting it into a database. The backend also sent the sensor data to the server housing the learners, again using XMPP to lend the backend greater extensibility.

Phase 2: Learners

Participants were periodically prompted to self-report their states using ecological momentary assessment on the phone. Sensor data acquired at these times were paired with simultaneously labeled state data. Using the labeled data, individualized prediction models were generated to identify specific user states from sensor values, including location, activity, social environment, and internal states (eg, mood, concentration). For every state, a machine learning algorithm generated a participant-specific model to predict that state from sensed data in the future. The machine learning algorithms discarded irrelevant information that did not improve the predictive value of the model, using data partitioning and averaging for continuous states, and information theory measures of information gain and entropy [39] for categorical states. States operationalized by scales (eg, sadness) were predicted using regression trees (ie, pruned Weka [17,40] REPTrees, version 3.6; Machine Learning Group, University of Waikato, Hamilton, New Zealand). Categorical states (eg, location) were predicted using J48 classifiers, which are Weka's adaptation of C4.5 decision trees [41] (see Figure 2). Pruning was also enabled for the location models. For the remaining categorical models, pruning was disabled due to their binary nature (eg, whether the user is having a "task-related conversation" or not).

Figure 2. Decision tree model predicting location from sensor values, generated from a research staff member's state ratings and sensor data (potentially identifying information has been altered)



Phase 3: Action Components

Every 5 minutes, the mobile phone application sent current, unlabeled sensor values to the backend. The learner used these sensor readings to make predictions and infer the participant's state without input from the participant. Because we rely on a service-oriented architecture, predicted states could be used to implement context-aware functionality through multiple devices. As the accuracy of predictions could not be determined prior to the trial, predictions were only used to display the currently predicted states to users on the mobile phone, as opposed to triggering any outreach or intervention.

When participants viewed the state entry ecological momentary assessment forms on the mobile phone, predictions as to their current states prepopulated the responses. A message at the bottom of the screen, in orange text, notified users of the date and time at which the predicted values were generated. If the phone lacked connectivity at that time, the user's last self-reported states were used to prepopulate the responses, and a green message indicated the date and time at which the user labeled these previous states.

Training the Context-Aware System

Development of accurate predictions requires training. Participants were periodically requested to label their states by selecting values for each context category from a drop-down menu on the mobile phone [42]. For example, if a participant is gardening outside while their family is inside the house, she or he might rate physical activity level as "light," location as "my yard/porch," conversation type as "none," and social context as "alone" in their immediate vicinity but "with family member(s)" in the wider environment. The learner would then match the sensed data to these state labels. Every time participants entered their states (this was described to participants as "training the phone"), new models were generated to accommodate the new data.

Specifically, the mobile phone application prompted users to report their current states 5 times per day or more according to participant preference. Participants could also label their states at any time on their own initiative. Using a 7-point Likert scale, participants rated their overall mood, intensity of discrete emotions (ie, happiness, sadness, anger, and anxiety), fatigue, pleasure, sense of accomplishment, concentration and engagement, and perceived control over current activities. Physical exertion was rated on a 4-point scale. Participants selected their location from options including a variety of public and private spaces (eg, their home, a friend's home, their office/desk at work, or a bus/train; see Figure 3). Participants' interactions with others were rated according to the type of conversation in which they were engaged, if any (eg, casual, task-related, disagreement, or none), as well as their relationship to others (eg, friends, family, strangers, or pets) in their immediate vicinity and larger environment. Such prompts occurred at random times between 7 AM and 10 PM, and participants could modify these hours via the website.

Figure 3. Screenshot, ecological momentary assessment of location on the mobile phone

When participants were prompted to enter their states, the application saved the current values of their phone sensors. Once participants had finished reporting their states, their responses were paired with the sensor data and sent to the backend server for processing by the learner.

Ecological Momentary Interventions

Additionally, the mobile phone application used local storage to combine mood ratings entered on the phone with those entered on the website and, consequently, derive a range, average, and standard deviation of mood. Participants received tailored feedback using these values. For example, when a user's self-reported mood was outside her or his typical range, a message would appear on the phone designed to reinforce improvement or suggest using a tool (see Website, below) in the case of deterioration. Other local data were used to provide feedback in such conditions as a substantial change in the participant's mood since their previous rating.

To support adherence to therapeutic activities scheduled via tools on the website, participants could choose to receive reminders on their mobile phones prior to the scheduled start times. Each reminder was followed by a check-in prompt after the activity was scheduled to end. If users indicated they had not completed a scheduled task, they received follow-up questions via a troubleshooting feature. Obstacles to completing the task (eg, lack of motivation) were identified per responses to these follow-up questions, and users were then provided with guidance on overcoming their specific difficulties.

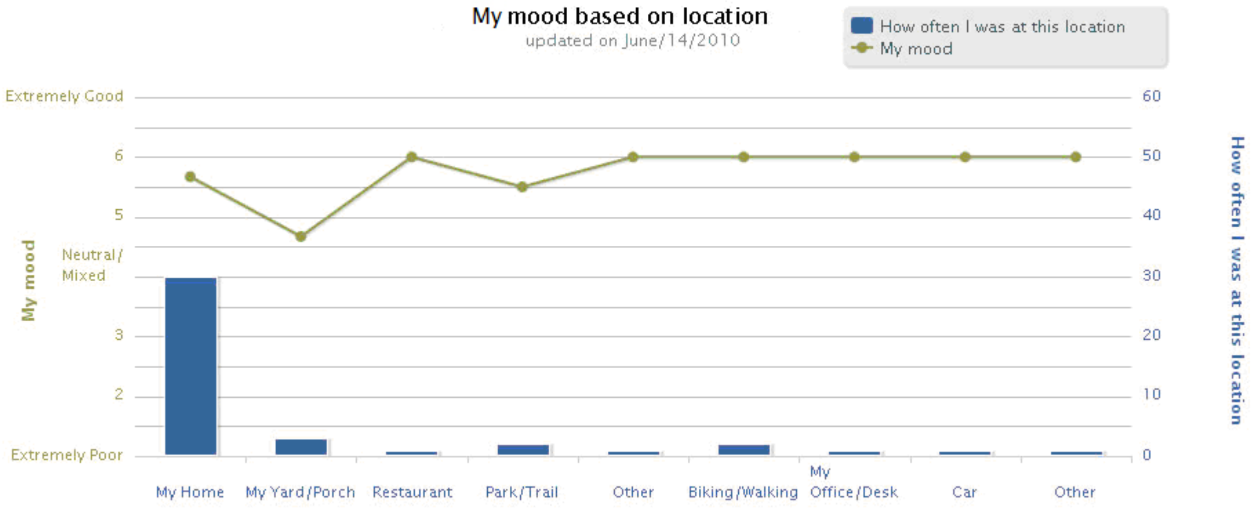
Website

The website served several functions. It provided a medium with which both participants and coaches could visualize state data entered into the mobile phone application. It also contained an interface allowing coaches to view the activity of their patients, including logins, lessons completed, and states labeled on the phone.

On the website, 9 didactic lessons that taught principles of behavioral activation were available. The lessons could be accessed through a computer or the mobile phone. These lessons were released weekly and required approximately 15 minutes to read. The initial lesson included psychoeducation on depression and detailed instructions on training the mobile phone. Subsequent lessons taught participants how to monitor the effect of daily behaviors on mood, schedule positive activities, experiment with new responses to distressing situations, recognize and change avoidance patterns (eg, rumination and procrastination), and maintain gains after treatment. Each lesson was paired with an interactive tool to provide participants with opportunities to apply the treatment concepts discussed in the lesson. Examples include an activity calendar with which participants could monitor and schedule their activities, as well as tools leading participants through each step involved in designing behavioral experiments or formulating tailored plans to replace avoidance behaviors with active coping behaviors. These tools were designed to be completed in just a few minutes, and participants were instructed to use the current week's tool on a daily basis.

Interactive graphical feedback tools displayed data from the backend server, which replicated and reformatted participant-labeled data and concurrent sensor values. These were transmitted to the Web server, which hosted the intervention website as well as its own database. Thus, participants could graph their ratings of emotions and other subjective experiences, and explore how these ratings were related to their reported locations (see Figure 4). The aim was to help participants better understand how they spend their time, and identify behaviors they would like to increase or decrease. Participants could also access a graph of their mood over a selected time period (ranging from the last 7 days to the duration of the study) to view their progress. Furthermore, coaches could create graphs associating any 2 states (eg, mood against level of physical exertion, mood against fatigue) to help participants identify patterns.

Figure 4. Graphical feedback available to users on the website (blue bars denote locations that a participant reported on the mobile phone, and the frequency with which each location was reported; the green line denotes the participant's average reported mood in each location)



Human Support

A manualized telephone coaching protocol [23] was used to enhance adherence via principles of supportive accountability [24]. Supportive accountability is a model of support that involves collaborative goal setting for adherence, as well as monitoring and encouragement in the context of a supportive relationship. This support model has been previously piloted in an Internet intervention [43]. Coaches also elicited usability feedback on Mobilyze! and provided a point of contact for difficulties. The first coaching call lasted approximately 45 minutes with the aim to establish rapport, address participant concerns or questions, and elicit and reinforce the participant's motivations for learning to manage depression. Participants then received a 5- to 10-minute coaching call and at least 1 email weekly. Emails were tailored to participants' mobile phone and website usage patterns, which coaches could monitor online. Participants also received a 30- to 45-minute technical demonstration of the mobile phone (in person or over the telephone) by the lab software engineer or study coordinator. While supportive accountability is designed to be applied by non-mental health providers, the coaches in this trial were PhD-level clinicians given the novelty of the intervention. The author of the coaching manual led weekly group supervision meetings.

Outcome Measures

Outcomes were assessed at baseline, week 4 (midtreatment), and week 8 (posttreatment). Participants were compensated $20 for each of the first 2 assessments and $50 for the final assessment and return of the mobile phone. Interview measures were administered over the telephone, while self-report measures were administered via a secure online questionnaire hosting service.

The telephone-administered Mini-International Neuropsychiatric Interview [25,26] was used to evaluate changes in major depressive disorder diagnosis and to characterize the sample in terms of comorbid anxiety disorders at baseline. The QIDS-C [28] was used to assess evaluator-rated depression symptom severity. The PHQ-9 [44] was used to evaluate self-reported depressive symptom severity. Finally, the Generalized Anxiety Disorder 7-item scale (GAD-7) [45,46] was used as a secondary outcome evaluating general anxiety symptom severity.

A semistructured interview designed by this research team was used to gather participant feedback regarding the intervention components. Feedback was also solicited via online self-reports at each assessment, and coaches recorded feedback received during the coaching sessions. Usage data were transmitted from the mobile phone to the lab's secure server via encrypted, password-protected tunnels, and from the website using secure sockets layer (SSL). The number of times participants logged into the website was calculated by considering all site activity occurring within the same hour as corresponding to a single login.

Analysis

Accuracy of the individual patient learner models was evaluated after the trial using Weka's cross-validation [47] procedure. This procedure estimated the classification accuracy that would result from using models built from the complete labeled sensor datasets. In 10-fold cross-validation, the superset of all self-reported values for a particular state are paired with their concurrent sensor values and divided into 10 subsamples. One of the subsamples is the validation set, and the other 9 are combined to form the training set. A new model is generated using the training set. Next, the new model predicts states from sensor readings in the validation set, and these predictions are compared against the actual self-reported states to calculate the rate of accurate classifications (ie, the number of correctly predicted states divided by the total number of cases in the validation set). For regression trees, a correlation was calculated between predicted and actual values in the validation set. This process is repeated 9 times, such that each of the other 9 subsamples is used once as the validation set. The estimator for the models is the average of these 10 accuracy rates. Next, bias corrected and accelerated 95% confidence intervals were constructed around the average accuracy rates for each state across participants, using 1000 bootstrap samples in SPSS 19 (IBM Corporation, Somers, NY, USA). For states predicted using regression trees, accuracy was calculated by averaging the correlations between predicted and actual values across participants.

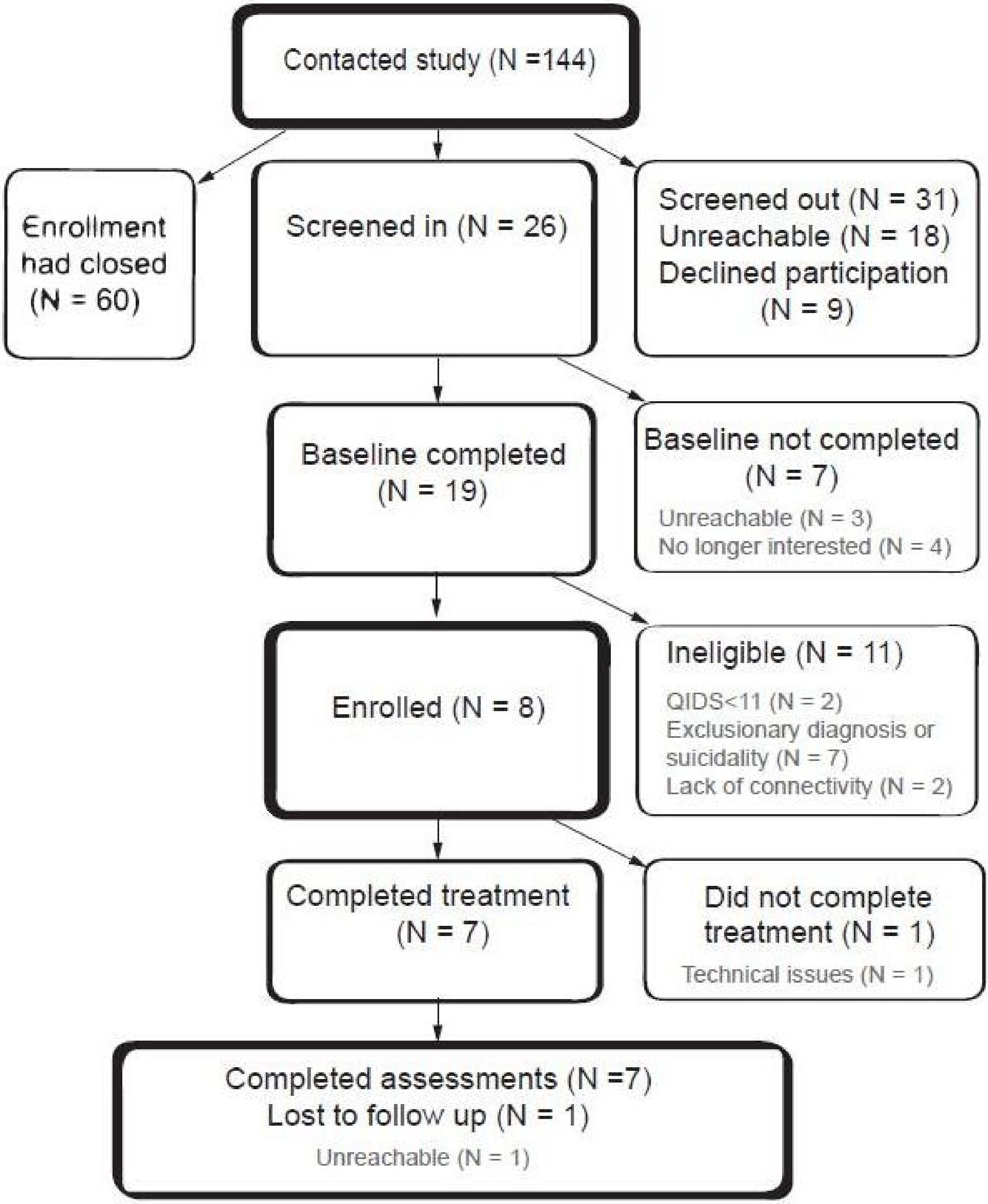
Continuous outcome measures (ie, the PHQ-9, QIDS-C, and GAD-7) were modeled against time on an intent-to-treat basis, using mixed models for repeated measures. These models were chosen because they can handle a certain amount of nonrandom missing data [48]. Adding the slope of time as a random effect did not significantly improve log likelihood values over using only a random intercept; thus, only a random intercept was included to allow for individual variation between participants. Generalized estimating equations logistic regression was used to model the binary repeated outcome measure of presence versus absence of major depressive disorder diagnosis. These analyses were run using SAS 9.2 (SAS Institute, Cary, NC, USA) with restricted maximum likelihood methods.

Results

Participant Characteristics

We enrolled 8 participants (7 females and 1 male) in the trial (see Figure 5). Ages of participants ranged from 19 to 51 years, mean 37.4 (SD 12.2). All participants reported at least a high school education, completing on average 14.3 years of education (SD 2.3). One participant (13%) identified as Hispanic and Caucasian, while the other 7 (88%) identified as non-Hispanic Caucasians; 7 (88%) were married and 1 (13%) was single.

Figure 5. Flow of participants through the trial (QIDS = Quick Inventory of Depression Symptoms-Clinician Rated)



Three of the participants (38%) were diagnosed with comorbid generalized anxiety disorder, but no other anxiety disorders, at baseline. Another participant (13%) was diagnosed with both generalized anxiety disorder and social phobia, and 2 other participants (25%) were diagnosed with generalized anxiety disorder, social phobia, and agoraphobia in the absence of panic disorder.

Adherence

Seven participants completed all 8 weeks of participation, while 1 dropped out in week 3 due to technical problems using the phone and website.

Mobile Phone Training

On average, the 8 participants trained their phone 15.3 times (SD 8.3) during the first week of the intervention, dropping to

9.0 ratings (SD 6.5) during the fourth week, and 4.8 ratings (SD 4.6) during week 8. The participants reported that they would have trained the phone more often, had they received more prompts to label their states. Problems with network connectivity were encountered that reduced the number of prompts from the Web server that reached the phone, and at least 1 of the 7 treatment completers never received ecological momentary intervention. During the trial, we corrected this issue by storing ecological momentary intervention logic and the dates/times for phone prompts locally, on the phone itself. This way, participants did not require connectivity to receive prompts to train the phone or ecological momentary intervention.

Website Use

The 8 participants logged on to the website an average of 7.9 times and completed 4.8 out of 9 lessons over the trial. Website use may have been reduced due to technical problems with some of the tools and in rolling out new lessons to participants in the first half of the trial.

Accuracy of Predictions

Across the 7 treatment completers, the mean rate of accurate classification for location was 60.3%, bootstrap 95% CI 43.2–77.2. The other categorical variables were difficult to compare across participants, as separate binary models were created for each available category. This was due to users' ability to endorse more than 1 option (eg, "friends" and "family" can simultaneously occupy one's immediate environment). As not all users reported ever being in certain categories, the existing models differed between users. The categorical states that were endorsed at least once by all treatment completers are listed in Table 1.

Table 1. Mean accuracy indicators for machine learning models of categorical states common to all treatment completers (N = 7)

Model	Mean accuracy (%)	95% CI[a]
Location	60.3	43.2–77.2
Alone in the immediate vicinity (Y/N[b])	80.1	76.2–84.5
Friends in the immediate vicinity (Y/N)	90.8	84.3–95.7
Alone in the larger environment (Y/N)	72.6	61.0–82.8
Miscellaneous people in the larger environment (Y/N)	90.9	83.8–97.3
Having a casual conversation (Y/N)	66.1	54.0–77.6
Not conversing (Y/N)	64.5	58.4–70.3

[a] Bias corrected and accelerated 95% confidence intervals (CIs) using 1000 bootstrap samples.

[b] yes/no.

The models for the scale-based states were not as encouraging. The cross-validation correlation coefficients were negative for all these states. Many participants' regression trees predicted 1 particular value regardless of sensor data. Thus, at first glance one would expect the correlation coefficients to be near zero across all the participants. However, overall mood ratings were negatively skewed (skewness –0.49, SE 0.10), and this caused the average value gleaned from the training folds to produce outlying, negative deviations between the predicted and labeled values when high-labeled values were present in the validation dataset. The correlation coefficients are therefore not reported, as they are somewhat misleading; rather, the accuracy of the scale-based state models should be conceptualized as no better than chance.

Clinical Outcomes

Per-protocol clinical outcomes are displayed in Table 2. The participant who dropped out failed to complete self-report measures at weeks 4 and 8, as well as interview measures at week 8. Another participant remained in treatment, but failed to complete interview assessments at week 4. However, all participants were included in the following analyses. Intent-to-treat analyses revealed that depressive symptoms self-reported on the PHQ-9 decreased significantly over time, $t_{13} = 7.02$, $\text{beta}_{\text{week}} = –.82$, $P < .001$. Evaluator-rated depressive symptoms on the QIDS-C also improved, $t_{13} = 8.22$, $\text{beta}_{\text{week}} = –.81$, $P < .001$. Over the course of treatment, participants also became less likely to meet diagnostic criteria for Major Depressive Disorder, $Z = 2.15$, $b_{\text{week}} = –.65$, $P = .03$. Scores on the GAD-7 indicated anxiety symptoms also decreased, $t_{13} = 4.59$, $\text{beta}_{\text{week}} = –.71$, $P < .001$.

Table 2. Per-protocol clinical outcomes

Variable	Baseline	Week 4	Week 8	Pre–post Cohen *d*[a]
% with MDD[b]	100% (8/8)	28.57% (2/7)	14.29% (1/7)	–
Mean (SD) PHQ-9[c] score	17.13 (3.80)	9.00 (4.47)	3.57 (4.12)	3.43
Mean (SD) QIDS-C[d] score	13.75 (2.71)	7.14 (3.89)	3.43 (3.10)	3.55
Mean (SD) GAD-7[e] score	15.50 (4.72)	9.43 (4.04)	6.43 (2.30)	2.58

[a] Standardized mean difference divided by the average of the baseline and week 8 SD.

[b] Major depressive disorder diagnosis on the Mini-International Neuropsychiatric Interview [25,26].

[c] Patient Health Questionnaire-9 [44].

[d] Quick Inventory of Depression Symptoms–Clinician Rated [28].

[e] Generalized Anxiety Disorder 7-item scale [45].

Participant Feedback

Posttreatment, the 7 treatment completers rated their satisfaction with the mobile phone in general by agreeing or disagreeing with the statement “I am satisfied with it.” Their average rating was 5.71 (SD 1.38) on a scale of 1 (representing strong disagreement) through 7 (representing strong agreement). The most common problems for which participants sought technical support were loss of connectivity, shortness of battery life, and phone freezing during use. Connectivity difficulties were further evidenced on the feedback measures; several participants indicated they were never prompted to train the phone. Similarly, we received reports of failure to receive mood predictions and scheduled activity reminders.

Despite these difficulties, during coaching calls 6 of the 7 treatment completers (86%) indicated the intervention was helpful in understanding triggers for negative moods and increasing their ability to recognize and modify distressing behaviors and cognitions. Two participants attributed these improvements to ecological momentary assessment on the mobile phone, while 4 of these participants cited ecological momentary assessment as well as the website. On the interview feedback measure, 1 of the participants stated that receiving the mood predictions was the most helpful component of the intervention. Another participant indicated that, because the predictions were helpful, she would have liked to continue using the mobile phone application after the study was over. Participants’ suggestions included lengthening the intervention and adding additional features such as a blog, messaging with coaches, or a recording tool to allow verbal elaboration on states when training the phone.

Discussion

To our knowledge, we have created the first ecological momentary intervention and the first context-sensing system for depression. While we encountered some technological difficulties, patients who used the Mobilyze! intervention showed significant clinical improvement and reported a high level of satisfaction with the intervention. This system holds great promise in offering an entirely new depression treatment option with the ability to maintain continuous contact with the patient in his or her environment, recognize the patient’s behavior and states, and intervene accordingly. While we encountered numerous challenges, there were also many lessons learned, which we will discuss in detail below.

We will discuss the following problems, lessons, and future directions. We encountered problems with connectivity and battery drainage, which can be addressed through more effective management of the mobile application. Context sensing requires adequate user-labeled data, accurate and meaningful sensor data, and appropriate data mining algorithms; we identified potential improvements in each of these areas. Finally, our experience suggests several options for maximizing the clinical utility of model predictions.

Management of Mobile Phone Applications

We encountered numerous technical problems related to battery drainage and connectivity. We followed the XMPP specification requiring a continuous connection between interfacing devices. Thus, the mobile phone continually searched for a connection to the backend server. This exacerbated battery drain by the phone application, particularly when the network connection was weak. A potential fix is to use a different underlying network protocol. For example, we may have been able to implement the same functionality using binary SMS messages instead of XMPP. However, this would be a less extensible solution, as not all platforms support this feature. Rather, we will minimize this problem in the future by equipping the application with a battery management system. This functionality would reduce phone communication with the backend server when low battery power or weak network connectivity is detected. Developing onboard data processing applications that use compression and feature extraction to reduce the quantity of data shipped to backend servers may also improve power efficiency, as can delaying transmission of sensor data that are not immediately needed to times when the phone is charging.

In the future, we will explore additional methods of maximizing battery life. Continuous sensing imposes an “energy tax.” Applications that manage the frequency of data acquisition from energy-expensive sensors such as GPS can improve battery life. Adopting duty cycling techniques, which manage the sleep cycle of sensors to balance the sensing frequency and battery consumption, will promote technical efficiency [49]. In other words, system management must be smart enough to adapt to the local environment within the phone.

Disruption in connectivity sometimes prevented delivery of ecological momentary assessment prompts from the backend server to the user, thereby reducing the frequency with which users trained the phone. On discovery of this issue during the trial, dates and times for prompts to occur were generated and saved locally on the phone; however, participants may have already become accustomed to more infrequent labeling. For functions that do not require complex processing (which can drain the battery if performed on the phone), future applications should strive for as much connectivity independence as possible to ensure more reliable functionality and to reduce battery drain.

Context Sensing

The ultimate aim of the machine learning system is to provide actionable evidence for outreach and other tailored patient interactions. To this end, our subsequent development process will focus on these semisequential stages: 1) increasing the amount of user-labeled data, across a greater variety of contexts, 2) improving accuracy and utility of sensor data, 3) improving accuracy of machine learning algorithms, and 4) maximizing the clinical utility of model predictions. The following section discussed each of these stages in more detail.

Increasing the Quantity and Variability of User-Labeled Data

One of our “known unknowns” is how well machine learning can extract the desired models from the types of sensor data we collected. This is an unknown because there was an insufficient amount of user-labeled data, and this contributed to the negative correlation coefficients on the scale-based states. Reinforcement techniques, such as those used successfully in the video game industry, could increase the frequency with which users label their states and reduce influence of outliers. Chris Bateman, a well-known game designer, described different types of rewards that motivate continued use of games [50]. Many of these would easily translate to Mobilyze!. For example, “currency rewards” could be used to reinforce training the phone by providing patients with points that can be used for shopping (eg, for new music files to play on their phone). “New toys” could also be used to promote adherence to ecological momentary assessment, by offering users new capabilities (eg, more sophisticated graphs of their states, and the ability to create a new, individualized state to label and visualize) in exchange for training the phone. “Rank rewards” can be as simple as informing users they have increased in rank (eg, “novice,” “expert,” or “master”) after they train the phone a set number of times.

Due to the preliminary nature of this study, model accuracy was evaluated retrospectively and through manual requests to the Weka software. This limited our ability to conduct ecological momentary assessment in an adaptive manner that would improve each user’s particular models. Our software will be adapted to interface with Weka, such that accuracy is automatically monitored throughout the trial. Accuracy measurements will be used to guide the ecological momentary assessment, as users can be queried more often on states for which their models are less accurate. Adherence to ecological momentary assessment procedures may be reinforced by communicating accuracy measurements or improvements to the user, such that users can track their progress in training the phone. This kind of reinforcement can be conceptualized as “completeness,” as some users are motivated to strive for perfect scores [50].

In addition to an inadequate quantity of labeled data, these data did not exhibit variability across the entire range of measured states. This may have occurred for several reasons. Users probably did not train the models across their entire set of contexts due to differential levels of convenience or social norms regarding use of mobile phones. Another possibility for low variability is that there actually was little variability in certain states, at least as detectable via the Mobilyze! ecological momentary assessment.

In an informed approach, variability of the user-labeled states used to construct each model can be harnessed to evaluate whether the user experience can be improved. The rate, sensitivity, and quality of user prompts can be modified accordingly. For example, users could be prompted to complete a “scavenger hunt” and train the phone when previously invariant states change (eg, a user who has always reported a sedentary state could be specifically prompted to train the model the next time they are more physically active). State labeling queries could also be recalibrated for greater sensitivity. Such adaptations can be automated to occur in response to individual differences in variability. For example, if an individual almost always reports being at home, the location query could be changed to allow specification of the particular room of the home. In a similar vein, it is likely that there are more variable states unique to a given user (eg, a user may report relatively static overall mood, but experience somatic symptoms of depression that vary quite a bit). In the future, the Mobilyze! system will allow users to create individually tailored states. Indeed, some participants in this trial indicated the Mobilyze! ecological momentary assessment did not query all the states they believed to be important. Finally, automated checks on label skewness can be used to encourage users to label their states more often when they are experiencing underreported values.

Improving Accuracy and Utility of Sensor Data

Accurate context sensing is fundamentally dependent on the quality of the sensor data and the quality of the features derived from the sensor data. Our experience has taught us that sensor data are often inaccurate, or technically accurate but misleading. For example, the ambient light sensor reported values brighter than the maximum meaningful value, and the accelerometer produced seemingly contradictory data when the phone was dropped or swinging as it was carried in a bag. To be functionally predictive, the raw sensor data should first be accurate from a technological perspective and then, if possible, be manipulated to extract features that describe more actionable, meaningful data [35]. An example of removing technological errors would be to discard implausible GPS coordinates. For the accelerometer, a feature could be extracted by using a time series of readings to determine trajectory of the phone and classify it as likely to be “swinging,” “dropping,” or neither.

Our more detailed plans for improving the quality of sensed data can best be illustrated by specific examples, which we will describe here in terms of the GPS data. GPS was anticipated to predict location, and yet the data proved so unreliable that the learning algorithm included GPS coordinates in only 1 of the treatment completers' location models.

Mobilyze! generates 2 sets of GPS data. One set comprises GPS readings sampled every 5 minutes, provided users had adequate telecommunications connectivity to enable GPS detection. This set does not include instances during which users trained the models. The second, much smaller set of GPS data consists of GPS readings at all physical locations where users labeled their location.

Accurate data in both datasets are critical for different reasons. Labeled GPS data are used to generate the models, which are designed to compare accurate data with accurate labeled states. If these data are compromised, the models themselves become unreliable. The second, unlabeled GPS dataset will ultimately be used to launch intervention. If these data are inaccurate, intervention may be triggered at inappropriate times. Given its completeness, the unlabeled dataset can also provide a higher-resolution perspective on the accuracy of the GPS transponder and users' daily GPS trajectories.

The following sections describe techniques we used to clean and understand GPS data post hoc. With appropriate engineering, however, these techniques could be used in real-time scenarios. As no single technique of cleaning and reformatting data is likely to effectively resolve all issues, a strategy combining multiple approaches should be used to improve accuracy. For the purposes of this description, we will use the location data provided by a project staff member due to her ability to confirm her location and travel patterns. Her GPS data were exported into Keyhole Markup Language (a standard method of defining geographic path information) and visualized on Google Earth.

Technological Error Exclusion

On several occasions data reflected visibly impossible scenarios, depicting GPS values that were up to 15 miles away from actual user-reported positions. Values that are impossible based on user-reported geographic position can be excluded. For example, we removed any values that depicted the staff member in a body of water, once she confirmed she had never been on the water. In the future, real-world geographic information system and commercial databases could be used to automate this process. For example, GPS readings could be compared against topographical features, and users could be queried to confirm less probable scenarios (eg, GPS indicates the user is in a remote area with extremely low population density, on a mountain, or in a body of water).

Other future techniques could exclude GPS points that suggest a user has traversed a greater than probable distance within a particular amount of time (eg, >70 miles/hour). Here, unlabeled data may be used to clean labeled data. Unlabeled GPS coordinates obtained before or after labeled coordinates can be used to calculate the implied velocities. Long term, this would be a way to leverage chronological GPS data to assist the learners. This technique could also prevent intervention based on improbable GPS data, regardless of whether users have recently labeled their states.

Feature Extraction

Some data inaccuracy is likely caused by the limited resolution of the GPS transponder, which may result in GPS coordinates that are inaccurate but relatively close to the user's actual location. Based on observations of data reported by the phone platform, GPS data were accurate in some cases to the 8- to 15-m^2 approximate resolution (if edge cases are removed). This range may not always be effectively actionable, and yet uniformly excluding unlikely values may not be the best solution. Feature extraction may help to salvage some of these GPS readings.

For example, users could be asked to predefine geographical places of significance or high frequency ("hot spots") at the beginning of the intervention using freely available geographic information system databases. A set of coordinates would thus be predefined and, in addition to raw GPS data, a "distance-from-hot-spot" feature could be added to enhance the quality of the GPS data. For example, a user may be at home, but the GPS data are fluctuating within a 50-m radius from the "home" hot spot. The feature would consist of the calculated distance from the home hot spot. Further, if this distance is within the phone's estimated resolution range, a binary feature could be defined to indicate that the user is likely to be at home.

Sensors could also define a user's physical trajectory. By storing a set of recently logged GPS positions in an external dataset, a spatial model could describe a best-fit line that estimates users' trajectory. This best-fit line could be used to generate a feature for the "estimated GPS" coordinates. This may be particularly helpful to handle missing data points occurring when the phone cannot obtain GPS readings due to connectivity problems.

Many other features could be envisioned. New features could be constructed based on interactions between sensors or between predictions generated by other models. For an example of the latter, being at home and alone may be depressogenic, while being at home and with friends may predict positive mood. Thus, the interaction between predictions for location and relationship to others in the environment could be inserted into the model predicting mood. There are also algorithms that use combinations of mathematical functions to automatically extract features from raw sensor datasets, and select or prune features based on their impact on learner accuracy [51,52].

Improving Accuracy of Learners

Although accuracy of the categorical states was promising, the accuracy of scale-based state models was poor and merits further discussion. On further examination of the scale-based state models, there appear to be at least 3 ways in which their accuracy might be directly improved. First, transformations can be used to normalize the distribution of user-labeled ratings prior to model generation and accuracy testing. Second, regression tree pruning can be selectively disabled. Regarding this option, we found post hoc that when regression trees were generated without pruning, models may be more likely to include sensor values rather than predicting a constant value and ignoring sensor data. Third, the system should be improved more generally to use ensemble methods. For example, bootstrapping or selective weighting can be applied to the labeled sensor data, generating multiple new training datasets [53,54]. Next, a model is created from each of these datasets using 1 of the learning algorithms (eg, decision trees). For each incoming set of sensor data, a prediction would be calculated by each of these models. The final prediction would be an aggregate of these individual predictions. Aggregation is likely to result in greater accuracy than use of individual models [55].

Improving Clinical Utility of Predictions

As we were unable to anticipate the accuracy with which users' states would be predicted in this field trial, predictions were not used to trigger intervention; rather, they were simply displayed to the user. Once models meet particular accuracy benchmarks, future studies will integrate them more fully into clinical applications. To leverage the potential benefits of a mobile platform, we will iteratively examine the methods by which feedback and outreach are delivered. For example, if outreach is delivered suggesting use of a particular tool, pre- and post-mood ratings could be used to evaluate its success for the user in that context. Experimental designs could also be employed to evaluate the impact of more sophisticated forms of feedback. We have further developed the Mobilyze! website such that it can display the GPS sensor readings on Google Earth. We aim to create a "mood map," where all GPS coordinates at which users labeled their mood will be displayed as colored dots. These dots will be color coded to convey the mood rating at each point, such that users can better visualize the impact of physical location on their mood.

Process Lessons

When we began designing this project in late 2008, Nokia's Symbian operating system and the Android operating system were our only options in terms of multitasking mobile phone platforms. It appeared more prudent to choose Symbian due to its longer history and platform maturity, whereas Android had just been deployed in October 2008. However, by 2010 Android became the most frequently purchased system in the United States [56], and Nokia has announced plans to retire the Symbian operating system in favor of Windows Phone 7 [57]. This speaks to the rapidity of development and change of technology used in mobile applications. In the future, we will use a cross-platform development framework (eg, Nitobi's PhoneGap) capable of deploying applications to several major mobile phone platforms, which should allow greater adaptability to this changing market. This would also allow individuals who already possess compatible smartphones to download the Mobilyze! phone application to their own phones, thus mitigating the costs of providing smartphones to users, as well as the inconvenience for users who would otherwise need to either carry 2 phones or transfer their existing phone settings to a new phone.

While we have learned much from this initial development enterprise, we have also found the numbers of challenges impossible to manage within the scope of a single project. These challenges included the cross-disciplinary expertise needed to process the collected data and the time invested to ensure basic functionality of ecological momentary intervention and website components. Investment of effort into the other intervention components resulted in limited resources to investigate the myriad techniques that could be used to clean and reformat the sensor data or improve the machine learning methods. Current plans are to continue research and development in a staged sequence, beginning with improvement of state predictions in the absence of any clinical intervention or website. Clinical intervention components will be added after the machine learning components function adequately.

Other Limitations

The current trial did not include a control condition. Once Mobilyze! is refined as previously described, a randomized controlled trial will be used to determine whether clinical improvements can be attributed to Mobilyze! as opposed to regression to the mean or other confounding factors. Further, Mobilyze! is a multilevel intervention comprising more traditional Internet approaches to behavioral intervention (ie, didactic lessons, exercises), coaching, ecological momentary assessment, and ecological momentary intervention. Factorial designs will be required to isolate the specific contributions of each of the components of Mobilyze! [58,59].

Conclusion

This trial demonstrated the feasibility, appeal, and potential efficacy of a new paradigm for mobile intervention targeting depression. The website delivered behavioral skills training, while the mobile phone provided self-monitoring with tailored, real-time feedback and intervention. Our initial experience suggested the feasibility and utility of a context-aware system using mobile phone sensing capabilities. We have described the numerous lessons learned and areas for continued development.

Mark Weiser, the father of ubiquitous computing, wrote that "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" [60, p. 94]. A context-aware system that can sense a person's behavior and state, coupled with a treatment platform that can positively reinforce adaptive behaviors and provide support for changing those that contribute to depression, has the potential to provide a fundamentally new model for mental health intervention.

Acknowledgments

This study was funded by National Institute of Mental Health Grant R34 MH078922 to David C. Mohr, Ph.D.

Conflicts of Interest

None declared

Multimedia Appendix 1

List of all collected phone sensor data

[PDF file (Adobe PDF File), 4 KB - jmir_v13i3e55_app1.pdf]

References

1. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
2. World Health Organization. The Global Burden of Disease: 2004 Update. Geneva: World Health Organization; 2008.
3. Simon GE. Social and economic burden of mood disorders. Biol Psychiatry 2003 Aug 1;54(3):208-215. [Medline: 12893097]
4. Katon W, Ciechanowski P. Impact of major depression on chronic medical illness. J Psychosom Res 2002 Oct;53(4):859-863. [Medline: 12377294]
5. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. J Affect Disord 2002 Dec;72(3):227-236. [Medline: 12450639]
6. Cuijpers P, Van Straten A, Warmerdam L, Smits N. Characteristics of effective psychological treatments of depression: a metaregression analysis. Psychother Res 2008 Mar;18(2):225-236. [doi: 10.1080/10503300701442027] [Medline: 18815968]
7. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409. [doi: 10.1002/jclp.20659] [Medline: 20127795]
8. Mohr DC, Hart SL, Howard I, Julian L, Vella L, Catledge C, et al. Barriers to psychotherapy among depressed and nondepressed primary care patients. Ann Behav Med 2006 Dec;32(3):254-258. [doi: 10.1207/s15324796abm3203_12] [Medline: 17107299]
9. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009 Dec;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
10. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
11. comScore, Inc. comScore.com. 2011. The 2010 Mobile Year in Review URL: http://www.comscore.com/Press_Events/Presentations_Whitepapers/2011/2010_Mobile_Year_in_Review [accessed 2011-07-25] [WebCite Cache ID 60RtzEoWK]
12. Lenhart A, Purcell K, Smith A, Zickuhr K. Pew Internet. Washington, DC: Pew Internet & American life Project; 2010 Feb 3. Social media and mobile Internet use among teens and young adults URL: http://www.pewinternet.org/~/media//Files/Reports/2010/PIP_Social_Media_and_Young_Adults_Report_Final_with_toplines.pdf [accessed 2011-04-01] [WebCite Cache ID 5xd2fHPYM]
13. Patrick K, Intille SS, Zabinski MF. An ecological framework for cancer communication: implications for research. J Med Internet Res 2005 Jul 1;7(3):e23 [FREE Full text] [doi: 10.2196/jmir.7.3.e23] [Medline: 15998614]
14. Depp CA, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson TL, et al. Mobile interventions for severe mental illness: design and preliminary data from three approaches. J Nerv Ment Dis 2010 Oct;198(10):715-721. [doi: 10.1097/NMD.0b013e3181f49ea3] [Medline: 20921861]
15. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39. [doi: 10.1348/135910709X466063] [Medline: 19646331]
16. Wenze SJ, Miller IW. Use of ecological momentary assessment in mood disorders research. Clin Psychol Rev 2010 Aug;30(6):794-804. [doi: 10.1016/j.cpr.2010.06.007] [Medline: 20619520]
17. Witten IH, Eibe F. Data Mining: Practical Machine Learning Tools and Techniques. Amsterdam: Morgan Kaufman; 2005.
18. Pharow P, Blobel B, Ruotsalainen P, Petersen F, Hovsto A. Portable devices, sensors and networks: wireless personalized eHealth services. Stud Health Technol Inform 2009;150:1012-1016. [Medline: 19745466]
19. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181. [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
20. Morris M, Guilak F. Mobile Heart Health: Project highlight. IEEE Pervas Comput 2009 Jun;8(2):57-61. [doi: 10.1109/MPRV.2009.31]
21. Miluzzo E, Lane ND, Fodor K, Peterson R, Lu H, Musolesi M, et al. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application. In: SenSys'08: Proceedings of the 6th ACM Conference on Embedded Networked Sensor Systems; 2008 Nov 5-7; Raleigh, NC. New York, NY: ACM; 2008:337-350.
22. Gluhak A, Presser M, Zhu L, Esfandiyari S, Kupschick S. Towards mood based mobile services and applications. In: Kortuem G, Finney J, Lea R, Sundramoorthy V, editors. Smart Sensing and Context: Second European Conference, EuroSSC 2007, Kendal, England, October 23-25, 2007, Proceedings (Lecture Notes in Computer Science ... Networks and Telecommunications). Berlin, Germany: Springer; 2007.
23. Duffecy J, Kinsinger S, Ludman E, Mohr DC. Northwestern University, Department of Preventive Medicine. 2010. Telephone Coaching to Support Adherence to Internet Interventions (TeleCoach) URL: http://www.preventivemedicine.northwestern.edu/researchprojects/pdf/TeleCoachManual.pdf [accessed 2011-04-01] [WebCite Cache ID 5xd6OBKKy]

24. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
25. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:quiz 34-quiz 57. [Medline: 9881538]
26. Esposito E, Wang JL, Williams JV, Patten SB. Mood and anxiety disorders, the association with presenteeism in employed members of a general population sample. Epidemiol Psichiatr Soc 2007;16(3):231-237. [Medline: 18020197]
27. University of Pittsburgh Epidemiology Data Center. ids-qids.org. 2011. Inventory of Depressive Symptomatology (IDS) & Quick Inventory of Depressive Symptomatology (QIDS) URL: http://www.ids-qids.org/index2.html [accessed 2011-04-01] [WebCite Cache ID 5xd6wYPDy]
28. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 1;54(5):573-583. [Medline: 12946886]
29. The John D. & Catherine T. MacArthur Foundation’s Initiative on Depression & Primary Care and 3CM, LLC. The MacArthur Initiative on Depression and Primary Care. 2004. Depression Management Toolkit URL: http://www.depression-primarycare.org/images/pdf/macarthur_toolkit.pdf [accessed 2011-04-01] [WebCite Cache ID 5xd7H7Sxt]
30. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
31. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status. Neuropsychiatry Neuropsychol Behav Neurol 1988;1(2):111-117.
32. Addis ME, Martell CR. Overcoming Depression One Step at a Time: The New Behavioral Activation Approach to Getting Your Life Back (New Harbinger Self-Help Workbook). Oakland, CA: New Harbinger Publications; 2004.
33. Martell CR, Addis ME, Jacobson NS. Depression in Context: Strategies for Guided Action. New York, NY: WW Norton; 2001.
34. Choudhury T, Consolvo S, Harrison B, LaMarca A, LeGrand L, Rahimi A, et al. The mobile sensing platform: An embedded activity recognition system. IEEE Pervas Comput 2008 Jun;7(2):32-41.
35. Fogarty J, Hudson SE. Toolkit support for developing and deploying sensor-based statistical models of human situations. In: CHI '07 Proceedings of the SIGCHI conference on Human Factors in Computing Systems, Apr 28-May 3, San Jose, CA. New York, NY: ACM; 2007:135-144.
36. The Open Group SOA Working Group. The Open Group. 2011. Service Oriented Architecture: What is SOA? URL: http://www.opengroup.org/soa/source-book/soa/soa.htm [accessed 2011-04-01] [WebCite Cache ID 5xd8seEHA]
37. Winograd T. Architectures for context. Hum Comput Interact 2001;16(2-4):401-419. [doi: 10.1207/S15327051HCI16234_18]
38. XMPP Standards Foundation. The Foundation. Protocol Stack URL: http://xmpp.org/xmpp-protocols/xmpp-core/ [accessed 2011-07-25] [WebCite Cache ID 60RspebAr]
39. Mitchell TM. Machine Learning. New York, NY: McGraw-Hill; 1997.
40. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. SIGKDD Explorations 2009;11(1). [doi: 10.1145/1656274.1656278]
41. Quinlan JR. C4. 5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann Publishers; 1993.
42. Bao L, Intille SS. Activity recognition from user-annotated acceleration data. Pervasive Computing 2004;3001:1-17. [doi: 10.1007/978-3-540-24646-6_1]
43. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. J Med Internet Res 2010;12(5):e48 [FREE Full text] [doi: 10.2196/jmir.1370] [Medline: 21169164]
44. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [Medline: 11556941]
45. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097 [FREE Full text] [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
46. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [Medline: 17339617]
47. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Belmont, CA: Chapman & Hall/CRC; 1984.
48. Brown H, Prescott R. Applied Mixed Models in Medicine. Chichester, NY: J. Wiley & Sons; 1999.
49. Wang Y, Lin J, Annavaram M, Jacobsen QA, Hong J, Krishnamachari B, et al. A framework of energy efficient mobile sensing for automatic user state recognition. In: MobiSys '09 Proceedings of the 7th international Conference on Mobile Systems, Applications, and Services, Jun 22-25, Kraków, Poland. New York, NY: ACM; 2009:179-192.
50. Bateman C. Chris Bateman. 2005. Designing Rewards in Games URL: http://onlyagame.typepad.com/only_a_game/2005/08/designing_rewar.html [accessed 2011-04-01] [WebCite Cache ID 5xdB8XRH9]

51. Forster K, Brem P, Roggen D, Troster G. Evolving discriminative features robust to sensor displacement for activity recognition in body area sensor networks. In: Proceedings of the fifth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Dec 7-10, Melbourne, Australia. NA: IEEE Xplore; 2009:43-48.
52. Sherrah JR, Bogner RE, Bouzerdoum AKoza JR, Deb K, Dorigo M, Fogel DB, Garzon M, Iba H, et al. The evolutionary pre-processor: automatic feature extraction for supervised classification using genetic programming. In: Genetic Programming 2nd Conference (Proceedings of the 2nd Annual Conference on Genetic Algorithms), Jul 13-16. San Francisco, CA: Morgan Kaufmann; 1997:304-312.
53. Breiman L. Bagging predictors. Mach Learn 1996 Aug;24(2):123-140. [doi: 10.1007/BF00058655]
54. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 1997 Aug;55(1):119-139. [doi: 10.1006/jcss.1997.1504]
55. Dietterich T. Ensemble methods in machine learning. In: Kittler J, Roli F, editors. Multiple Classifier Systems: First International Workshop, MCS 2000 Cagliari, Italy, June 21-23, 2000 Proceedings (Lecture Notes in Computer Science). Berlin, Germany: Springer; 2000:1-17.
56. Gohring N. Computerworld. 2010 Aug 3. Canalys: Android Takes the Lead in the US URL: http://www.computerworld.com.au/article/355426/canalys_android_takes_lead_us/ [accessed 2011-04-01] [WebCite Cache ID 5xdBIRdi6]
57. Nokia. Nokia Corporation. 2011. Form 20-F 2010 as filed with the Securities and Exchange Commission on March 11, 2011 URL: http://www.nokia.com/NOKIA_COM_1/About_Nokia/Financials/form20-f_10.pdf [accessed 2011-04-01] [WebCite Cache ID 5xdBT1dRi]
58. Collins LM, Murphy SA, Nair VN, Strecher VJ. A strategy for optimizing and evaluating behavioral interventions. Ann Behav Med 2005 Aug;30(1):65-73. [doi: 10.1207/s15324796abm3001_8] [Medline: 16097907]
59. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118. [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
60. Weiser M. The computer for the 21st century. Sci Am 1991;265(3):94-104.

Abbreviations

GAD-7: Generalized Anxiety Disorder 7-item scale
GPS: global positioning system
PHQ: Patient Health Questionnaire
QIDS-C: Quick Inventory of Depression Symptoms–Clinician Rated
SMS: short message service
XMPP: extensible messaging and presence protocol

Edited by G Eysenbach; submitted 26.04.11; peer-reviewed by J Proudfoot, B Dear; comments to author 17.05.11; revised version received 10.06.11; accepted 14.06.11; published 12.08.11.

Please cite as:
Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, Mohr DC
Harnessing Context Sensing to Develop a Mobile Intervention for Depression
J Med Internet Res 2011;13(3):e55
URL: http://www.jmir.org/2011/3/e55/
doi:10.2196/jmir.1838
PMID:21840837

Original Paper

Features of Mobile Diabetes Applications: Review of the Literature and Analysis of Current Applications Compared Against Evidence-Based Guidelines

Taridzo Chomutare[1,2], MSc; Luis Fernandez-Luque[2,3], MSc; Eirik Årsand[1,2], PhD; Gunnar Hartvigsen[1,2], PhD

[1]Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway
[2]Medical Informatics & Telemedicine Group, Department of Computer Science, University of Tromsø, Tromsø, Norway
[3]Northern Research Institute, Tromsø, Norway

Corresponding Author:
Taridzo Chomutare, MSc
Norwegian Centre for Integrated Care and Telemedicine
University Hospital of North Norway
Sykehusvn. 23
Tromsø, 9038
Norway
Phone: 47 47680032
Fax: 47 77754098
Email: taridzo.chomutare@telemed.no

Abstract

Background: Interest in mobile health (mHealth) applications for self-management of diabetes is growing. In July 2009, we found 60 diabetes applications on iTunes for iPhone; by February 2011 the number had increased by more than 400% to 260. Other mobile platforms reflect a similar trend. Despite the growth, research on both the design and the use of diabetes mHealth applications is scarce. Furthermore, the potential influence of social media on diabetes mHealth applications is largely unexplored.

Objective: Our objective was to study the salient features of mobile applications for diabetes care, in contrast to clinical guideline recommendations for diabetes self-management. These clinical guidelines are published by health authorities or associations such as the National Institute for Health and Clinical Excellence in the United Kingdom and the American Diabetes Association.

Methods: We searched online vendor markets (online stores for Apple iPhone, Google Android, BlackBerry, and Nokia Symbian), journal databases, and gray literature related to diabetes mobile applications. We included applications that featured a component for self-monitoring of blood glucose and excluded applications without English-language user interfaces, as well as those intended exclusively for health care professionals. We surveyed the following features: (1) self-monitoring: (1.1) blood glucose, (1.2) weight, (1.3) physical activity, (1.4) diet, (1.5) insulin and medication, and (1.6) blood pressure, (2) education, (3) disease-related alerts and reminders, (4) integration of social media functions, (5) disease-related data export and communication, and (6) synchronization with personal health record (PHR) systems or patient portals. We then contrasted the prevalence of these features with guideline recommendations.

Results: The search resulted in 973 matches, of which 137 met the selection criteria. The four most prevalent features of the applications available on the online markets (n = 101) were (1) insulin and medication recording, 63 (62%), (2) data export and communication, 61 (60%), (3) diet recording, 47 (47%), and (4) weight management, 43 (43%). From the literature search (n = 26), the most prevalent features were (1) PHR or Web server synchronization, 18 (69%), (2) insulin and medication recording, 17 (65%), (3) diet recording, 17 (65%), and (4) data export and communication, 16 (62%). Interestingly, although clinical guidelines widely refer to the importance of education, this is missing from the top functionalities in both cases.

Conclusions: While a wide selection of mobile applications seems to be available for people with diabetes, this study shows there are obvious gaps between the evidence-based recommendations and the functionality used in study interventions or found in online markets. Current results confirm personalized education as an underrepresented feature in diabetes mobile applications. We found no studies evaluating social media concepts in diabetes self-management on mobile devices, and its potential remains largely unexplored.

(J Med Internet Res 2011;13(3):e65) doi:10.2196/jmir.1874

KEYWORDS

Mobile health (mHealth); diabetes mellitus; blood glucose self-monitoring; social networks; personal health records (PHR); personalized education; diabetes self-management; health informatics

Introduction

Social media and user-friendly mobile devices are one of the most significant recent developments in information and communication technology. As well as being commercially successful and popular, social media hold a potential for interesting new use cases in health care. On the other hand, enhanced usability and pervasiveness of mobile devices have resulted in renewed interest in and development of new requirements for health care applications. Research has consistently shown that diabetes management is one application area [1-5] where mobile devices could enhance the quality of life for people living with chronic illnesses [6].

Although there is now a wide body of literature on the use of mobile devices in self-management of diabetes, present knowledge about good practice in designing integrated health applications seems rather limited. An integrated application provides intermodule and interapplication communication interfaces transparent to the user, resulting in a seamless whole. We have not found research focused on the gaps between the functional requirements (evidence-based recommendations in clinical guidelines) and the functionality available in current tools. Some reviews address the design of user interfaces [7,8] and the effect of mobile applications on health outcomes [9-12], but these studies have not treated functional requirements in much detail.

The conclusion in Farmer and colleagues' [13] most influential work, a randomized controlled trial involving 93 patients, was that decision-support features were important to realize benefits from blood glucose monitoring. On the other hand, in a UK and Canada review from 2009, Seto et al [14] found no benefits from monitoring blood glucose, but argued for the inclusion of communication with primary care providers [3,15,16] in the design of interventions. In a study involving 30 patients with type 2 diabetes, Faridi et al [17] found no statistically significant improvement in glycosylated hemoglobin with the use of mobile applications compared with standard therapy. It is likely that conflicting outcomes in these and other similar studies are partially due to differences in the design of the mobile tools as interventions. Well-designed mobile tools with decision-support features [13] such as personalized education have demonstrated potential to enhance self-management outcomes [8].

We present an in-depth analysis of the features of diabetes mobile applications. In addition, we contrast the requirements derived from evidence-based recommendations with the functions available in existing interventions. The rationale is to identify gaps and contribute to improving the tools available to the target group. The aim of the analysis is to answer three questions: (1) what functionality is available on the market for diabetes mobile applications?, (2) what gaps exist in relation to the evidence-based recommendations for this target group? and (3) what new use cases from social media could enhance such applications?

Methods

Our goal was to review as many and as diverse diabetes mobile applications as possible, both in the literature and in commercial markets. Many successful applications do not have any grounding in research, hence our decision to include the online markets and gray literature, where people in general showcase their innovation, often based on personal needs. While the literature typically reflects emerging applications and new trends, the market gives a good indication of mature applications and functionality. Comparing and contrasting the current functionality with recommendations in clinical guidelines constitutes a gap analysis.

Selection Criteria

The main inclusion criterion was that the application had a self-monitoring of blood glucose (SMBG) component. This inclusion criterion had the potential to preclude relevant applications, but in reality none of the excluded applications had a clear focus on diabetes. We settled on SMBG as the main inclusion criterion in order to filter out applications intended exclusively for medical professionals rather than patients, as well as other general health and lifestyle applications. We excluded applications without English-language user interfaces and those designed exclusively for health care professionals. We also excluded hardware-based solutions geared toward blood glucose tracking or insulin pumps only. Applications with their latest updates or publications prior to 2006 were excluded.

Search Strategy

The search was based on two main source types. The first source was online journal databases, indexers, and reference lists. We searched for prototypes and work in progress using the search terms "diabetes," "mobile," "PDA," "cell," "phone," and "application". We constructed a search string using both the conjunction "AND" and the disjunction "OR" logical operators (diabetes AND [mobile OR PDA OR cell OR phone OR application]). The search was based on the metadata—that is, title, abstract, and keywords. We targeted both original research papers and review articles indexed by Medline, ScienceDirect, ACM (Association for Computing Machinery) Digital Library, IEEE (Institute of Electrical and Electronics Engineers) Xplore Digital Library, Google Scholar, and DBLP (Digital Bibliography & Library Project) Computer Science Bibliography. The databases reflect the multidisciplinary nature of the research involving both medical and computer science fields. We identified three recent relevant reviews by Årsand et al [18], Tatara et al [7], and Liang et al [4], where we cross-checked descriptions. We also searched the gray literature: technical reports, Internet blogs, and portals.

The second source was online stores for mobile applications, using the search terms "diabetes" and "glucose" with the disjunction "OR" logical operator (diabetes OR glucose). We identified online stores for four leading platforms: Apple iPhone, Google Android, BlackBerry, and Nokia Symbian. We searched these two source types—namely, the online journal databases and the online markets—independently of each other. We searched the online market first, and then subsequently searched the related literature.

Evaluation and Assessment of Application Features

We analyzed the following features: (1) self-monitoring: (1.1) blood glucose, (1.2) weight, (1.3) physical activity, (1.4) diet, (1.5) insulin and medication, and (1.6) blood pressure), (2) education, (3) disease-related alerts and reminders, (4) integration of social media functions, (5) disease-related data export and communication, and (6) synchronization with personal health record (PHR) systems or patient portals. These features are the result of iterated brainstorming sessions among the coauthors and discussions in focus group meetings with patients and physicians. The emphasis in these sessions was put on translating guideline recommendations into a requirements specification implementable on a mobile phone platform. We created a list with multiple features and in iteration reduced the list to six main features, which we believed had the most potential for enhancing future mobile applications.

These features are individually quite distinct, but they have the potential to work as an integrated self-management tool. For example, the user could log weight, physical activity, meals, or carbohydrate intake, and have an easy-to-understand visual display to see how they correlate or affect the blood glucose. It should be noted that the "insulin" feature in most applications was part of a customizable "medication" feature for managing other medications as well.

We installed the available applications and recorded the functionality in a spreadsheet (see Multimedia Appendix 1). For those that we were not able to install, we cross-referenced the function descriptions in published articles. We noted whether each of the functions required manual interaction with the user, or whether wired or wireless sensors were used to import data into the application automatically. We then compared the prevalence of features with the recommendations in several clinical guidelines (see Discussion section for references to guidelines). Guideline recommendations can provide a good basis for requirements analysis and specification during the design and development of diabetes applications.

The process of extracting the data presented a major risk of error and uncertainty. For example, the literature is in most instances implicit about the functionality, and it is easy to miss or misunderstand feature descriptions within the text. To avoid potential problems, we enhanced the assessment process with independent verification. While we cannot claim the process we designed is entirely infallible, we avoided likely pitfalls that might otherwise have invalidated our findings. The next paragraph explains the process in some detail.

One author (EÅ) conducted the market and literature search in another related study [18] and another author (TC) conducted another independent search and installation of the selected applications. A third author (LFL) inspected the installation and assessed the feature on randomly sampled applications. The forth author (GH) independently undertook another round of data inspection and verification. Disagreements were settled by discussion and when necessary by redefining categories.

Results

The breakdown of the search process from online journal databases, gray literature, and online markets is shown in Figure 1. As illustrated in the figure, the total matches were 485 for literature and 488 for online markets, bringing the total matches for this study to 973. We went through a sifting process, with 36 applications from the literature and 101 from the online markets remaining, ending at a total of 137 mobile applications. Of the selected 101 market applications, 40 were available for free. The mean and modal price for the rest of the applications was the equivalent of €2.50 and €1.50, respectively. Of the 40 free applications, 12 had some premium functionality available only at an additional cost.

Some applications were counted multiple times—that is, for each platform or source on which they appeared. Of the total 137 eligible applications, we installed 82 on mobile devices for further analyses and classified the rest as either work in progress or unavailable for installation. Two of the 82 installed applications—namely "Tag-It-Yourself" and "Few Touch"—were from the literature. The former was also available on iPhone and the latter was developed in-house. It is important to note that some studies used commercially available applications but did not explicitly refer to the application names or features, and were thus excluded from this study.

We labeled many studies as having irrelevant titles; because the search terms included diabetes and mobile, it was common to find purely medical studies on diabetes in medical journals. Our search was based on the title, abstract, or keywords, but even this streamlined search criterion is bound to yield many irrelevant articles. On the other hand, most of the articles that matched the search criteria in information and communication technology journals turned out to contain relevant data for this study. Abstracts that were judged to have low probability of containing relevant data were labeled as unpromising and excluded from this study.

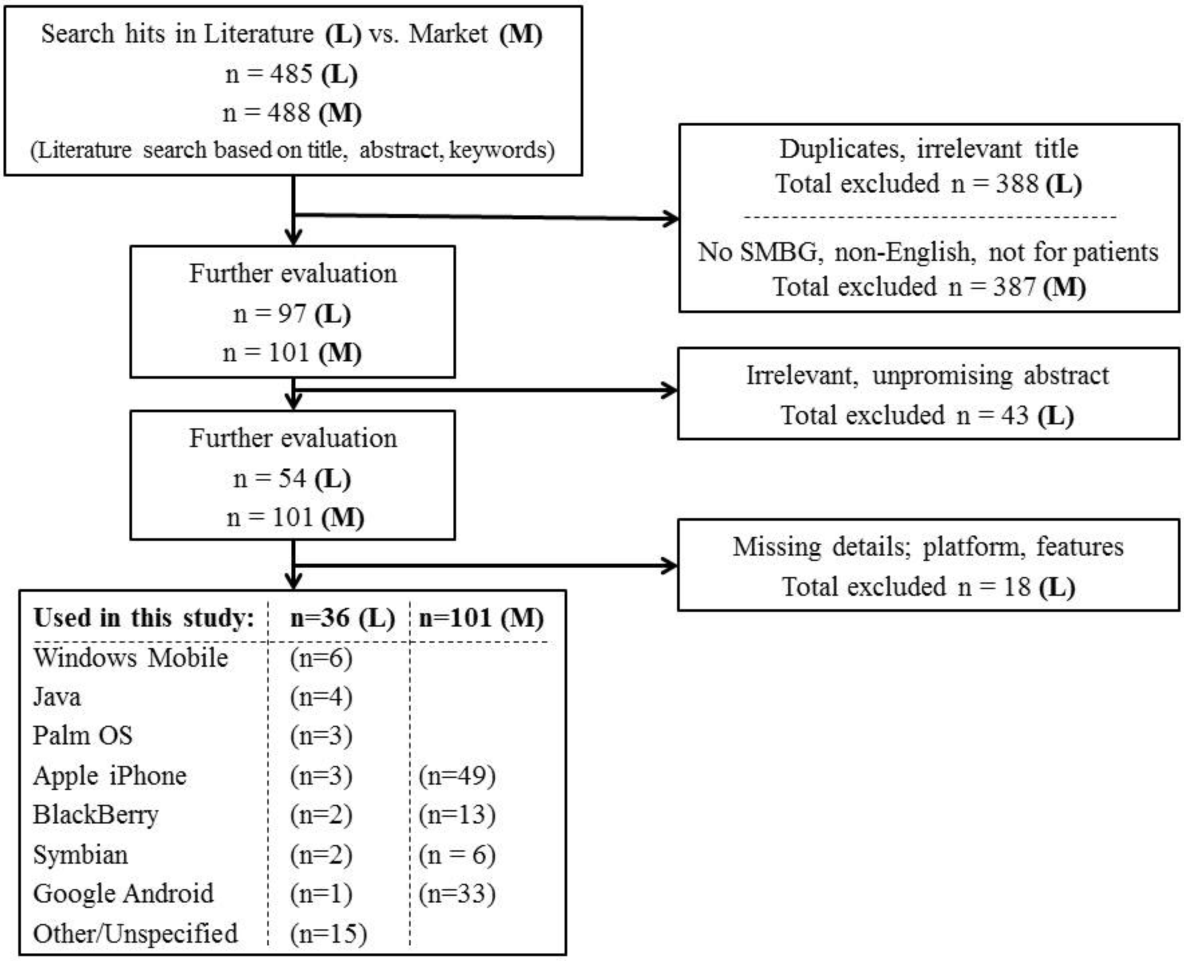
The features of the mobile applications per mobile platform and source are summarized in Table 1. The figures include the total results from the online stores, journal databases, and gray literature. Explanations of the functionalities are given in Multimedia Appendix 2. The blood glucose monitoring feature is not shown in Table 1 because it is a part of all applications as implied by the selection criteria.

Table 1. Numbers and percentages of applications (n = 137) with the respective features of insulin, communication (Comm), diet, physical activity (PA), weight, blood pressure (BP), personal health record (PHR), education (Edu), social media (SM), and alerts

Application	Insulin	Comm	Diet	PA	Weight	BP	PHR	Edu	SM	Alerts
Apple iPhone (n = 49)	35 (71%)	36 (73%)	26 (53%)	17 (35%)	19 (39%)	13 (27%)	7 (14%)	8 (16%)	12 (24%)	7 (14%)
Google Android (n = 33)	19 (58%)	17 (52%)	15 (45%)	10 (30%)	16 (48%)	16 (48%)	7 (21%)	3 (9%)	0 (0%)	0 (0%)
BlackBerry (n = 13)	5 (38%)	6 (46%)	3 (23%)	2 (15%)	5 (38%)	4 (31%)	1 (8%)	2 (15%)	4 (31%)	0 (0%)
Nokia Symbian (n = 6)	3 (50%)	2 (33%)	4 (67%)	4 (67%)	4 (67%)	3 (50%)	2 (33%)	2 (33%)	1 (17%)	1 (17%)
Average for online markets (n = 101)	63 (62%)	61 (60%)	47 (47%)	34 (34%)	43 (43%)	36 (36%)	17 (17%)	16 (16%)	17 (17%)	8 (8%)
Average for literature (n = 26)	17 (65%)	16 (62%)	17 (65%)	15 (58%)	7 (27%)	6 (23%)	18 (69%)	10 (38%)	3 (12%)	7 (27%)
Average for gray literature (n = 10)	9 (90%)	4 (40%)	7 (70%)	5 (50%)	3 (30%)	2 (20%)	5 (50%)	2 (20%)	0 (0%)	1 (10%)
Total weighted average	89 (65%)	81 (59%)	71 (52%)	55 (40%)	53 (39%)	44 (32%)	40 (29%)	27 (20%)	21 (15%)	16 (12%)

Table 1 shows that tools for tracking insulin or other medication were present in 89 (65%) of the applications, although most online market applications did not specify whether the application was meant for type 1 or type 2 diabetes. Just over half of the applications had some form of diet management, either by tracking carbohydrate intake or by providing meal suggestions. Physical activity and weight tracking had 55 (40%) and 53 (39%) applications, respectively. A component for synchronizing with PHRs or Web portals was present in 40 (29%) of the applications. Only seven of the 27 applications with an educational module had personalized education, tips, feedback, or advice. Few applications were sensitive to the age or gender of the users; important specific factors for special user groups such as pregnant women, for example, were largely ignored. Some form of lightweight integration with social media was present in 21 (15%) applications, while 16 (12%) had disease-related reminders. Of the applications randomly sampled for verification checking, 7 (5%) of the 130 features analyzed were in disagreement. None of the disagreements concerned features related to our main findings.

Figure 1. Selection process for online journal databases and online markets (SMBG = self-monitoring of blood glucose).



Discussion

The results in Table 1 are revealing in several ways. Perhaps the most significant outcome apparent in Table 1 is that education is a feature present in only a few diabetes-related mobile applications. Second, we can observe that a small percentage of applications have social media, suggesting that the influence of social media on the development of diabetes mobile applications is so far negligible. Another interesting finding emerging from this study is that most online market applications are based on manual entry of data such as blood glucose levels and weight, while 16 (62%) of the 26 applications found in the literature used wireless (Bluetooth, ZigBee, or Wi-Fi) automatic data acquisition. Wireless sensors are now widely available, but proprietary rights and vendor restrictions hamper their use in some commercial markets (eg, Apple iPhone). Manual data input not only exposes the user to erroneous input, but it can also be a daunting task and may lower compliance [14]. In the remaining subsections, we discuss the details of these results.

Core Functionality versus Requirements

To discover whether the requirements from clinical guidelines were necessarily met, we turned to what was available on the online markets. However, it was impossible to accurately determine how many of the applications available on the commercial market were used in research or were founded on evidence-based principles. The four most prevalent features can be seen from the data in Table 1. We highlight these features in Table 2 with a slightly different perspective. We omit results from the gray literature in the table because of its potential to obfuscate important elements that the table illustrates.

Table 2. The most prevalent features (n = 137 applications) on the online markets versus in the literature

Order	Online stores	Literature	Overall weighted prevalence
1.	Insulin, 63 (62%)	Personal health record, 18 (69%)	Insulin, 89 (65%)
2.	Communicating, 61 (60%)	Insulin, 17 (65%)	Communication, 81 (59%)
3.	Diet, 47 (47%)	Diet, 17 (65%)	Diet, 71 (52%)
4.	Weight, 43 (43%)	Communication, 16 (62%)	Physical activity, 55 (40%)

The results obtained from online journals and markets are compared and weighted in Table 2. From the data, we can see that most applications used in the literature integrated with a PHR, despite the intricacies associated with PHR integration. Outside well-controlled research, however, it is typically more difficult to offer PHR functionalities for facilitating collaborative care. In addition to application development challenges, more threatening are legislative and organizational barriers related to communicating patient data. The existing restrictive environment has hampered adoption and discouraged potential innovators such as Google Health. In Table 2, the PHR feature disappears from the weighted list because of the sample size imbalance between the market (n = 101) and the literature (n = 26).

It is easy to see that the rankings from Table 2 are biased toward patients with more severe illness, where the patients use insulin. This is unexpected because between 90% and 95% of people with diabetes have type 2 diabetes. Most people with type 2 diabetes do not use insulin, but rather oral medication and lifestyle changes such as diet and physical activity. This bias toward insulin-based solutions initially seems counterintuitive in light of the stated statistical data, but the underlying premise is based on known results [19,20]. Although there is still much disagreement among researchers, evidence of the benefit of intensive blood glucose monitoring for patients not using insulin seems rather weak [19,20,21].

Recent advances reflected in these clinical guidelines [22-24] recommend the following features (in random order) as part of important variables for diabetes self-management:

1. Education and personalized feedback;
2. Diet management;
3. Weight management;
4. Physical activity;
5. Communication and patient monitoring by primary care providers;
6. Insulin and medication management;
7. Other therapeutics (foot, eye care);
8. Psychosocial care;
9. Immunization;
10. Complication management.

It is important to note that current applications meet the functional requirements list only partially. Some of the features are shown in the screenshots of two sampled applications in Figure 2 and Figure 3. Figure 2 shows an iPhone application, Glucose Buddy. The application has a touch-sensitive interface and can also be installed on compatible mobile devices such as Apple iPods. Figure 3 shows a Windows Mobile application, Few Touch. This application also has a touch-sensitive interface and can be installed only on devices with the Windows Mobile operating system. The applications in both Figure 2 and Figure 3 are platform specific, as are most mobile applications.

Figure 2. Glucose Buddy iPhone application screenshots showing the main menu (left), blood glucose logging (center), and medication logging (right).

Date 20. juli 2011 22:48
Notes maybe too soon after meal
BG (mg/dL) 6

Date 20. juli 2011 22:42
Notes Small headache after meal
insulin (units) 9

Figure 3. Few Touch Windows Mobile application screenshots showing the main menu (left), food registration (center), and feedback on diet goals (right).

Low carb. snacks
Today: 1
My goal: 5
Meals
Today: 1
My goal: 5
High carb. meals
Prev. 7 days: 2
Goal: max 2 per week

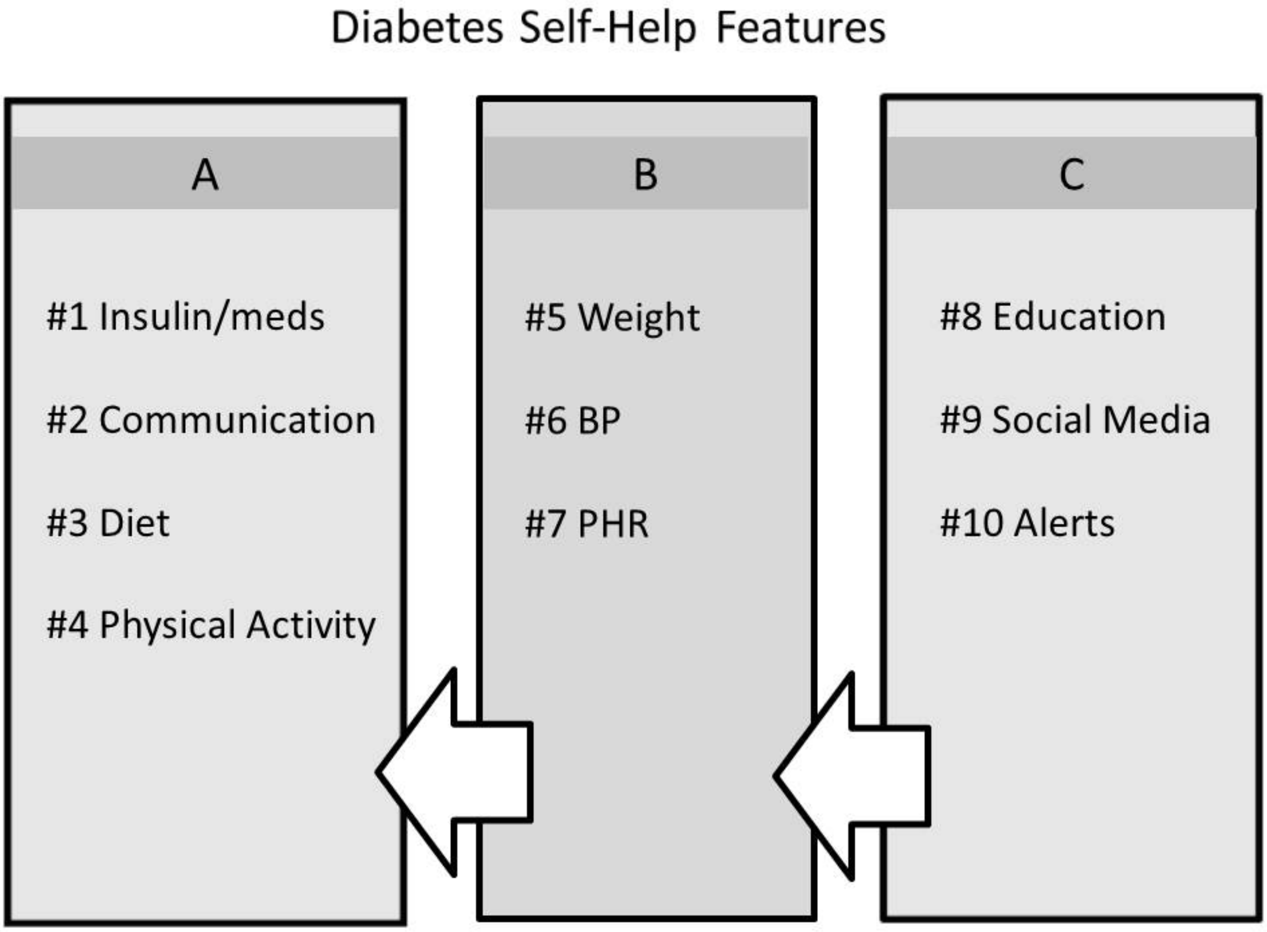
Classification of Functionality

Figure 4 illustrates an arbitrary classification of the surveyed mobile applications on the basis of prevalence. Class A comprises the four major features. Class B functionality comprises weight management, blood pressure monitoring, and PHR integration. These have a significantly higher prevalence than the class C features, which comprise education, social media integration, and alerts. In the future, we expect the ideal application to have all the features available as part of the core application, resulting in an integrated, feature-rich, and adaptive system. The presented classification may be useful for application developers and intervention designers when considering the features to implement. In addition, the classification is intended to draw attention to the least prevalent and less well-studied features.

Although the Action to Control Cardiovascular Risk in Diabetes (ACCORD) study group [25] found no significant benefits of intensive antihypertensive therapy for patients with type 2 diabetes at risk of cardiovascular events, blood pressure monitoring is likely to be part of class A functionality as wireless blood pressure sensors for home and personal use become more ubiquitous. Weight management is important for overweight and obese patients with diabetes. However, a good percentage of the patients do not struggle with weight; therefore, weight management seems rationally placed in class B. A PHR integration feature was implemented in most scientific studies, but because it was not available in online markets where more applications were sampled, it remained in class B.

It is somewhat surprising that education is conspicuously underrepresented, even though clinical guidelines suggest it belongs in class A. Structured and personalized education and actionable feedback are widely suggested as the missing link for people with diabetes who do not use insulin. It is not entirely clear why social media and alerts had very low prevalence, but it could partially be because they are difficult to implement. In the succeeding two subsections we discuss education and social media features in some detail.

Figure 4. Arbitrary classification of functionality based on prevalence in the surveyed mobile applications (BP = blood pressure; PHR = personal health record).



Personalized Structured Education: The Missing Link

Current results reveal something completely unexpected: only 27 (20%) of the applications had an education module, and only 7 (26%) of these met our criteria for personalized education or feedback. A recent study [26] showed that, although Internet health information is growing rapidly, the average person lacks the skills for finding and using the health information strategically for his or her benefit. For people with diabetes who do not use insulin, personalized structured education may be the missing link to deriving benefits from SMBG. Some reviews have supported the view that SMBG does not benefit patients who are not using insulin [19,20,21,27]—having the main end point as glycosylated hemoglobin. In contrast, Polonsky et al [28] argue that previous research has not considered structured SMBG with proper education [28,29], where participants are empowered with actionable knowledge on how to deal with different circumstances related to blood glucose variability. Clar et al [19] also agree with the view that proper actionable knowledge will result in SMBG benefits for people not using insulin or oral agents [19,30,31]. Existing accounts fail to resolve the contradiction between the two views. Larger clinical trials with well-defined interventions are needed to provide definitive evidence.

Nonetheless, in terms of design and development of personalized education modules, the task is challenging and the research field is still undeveloped. There is considerable scope for personalization because the mobile applications have access to some data about the users and their health status. However, using these user data for personalizing education is obviously not trivial. Personalizing health education is a rich and interesting field of inquiry that deserves urgent attention.

Social Media: Emerging Use Cases

In a recent survey, Chen [32] showed the importance of social aspects and experience-sharing among people with diabetes. Chen's findings underscore the importance of individuality and the need for tailored social interactions, which resonate with the concept of "PatientsLikeMe" [33,34], which has recently received enormous attention.

Findings from this study suggest very little influence of social media on current diabetes mobile applications. Most applications that claim to include social media features only provide a link to their groups in well-known social networking sites such as Facebook and Twitter. Some applications also provide the user with an account to a forum. However, there are no functional links or integration between information in the mobile application and the social media application. For instance, it is not easy to share graphs and data in the mobile applications with friends or relatives in social networks.

Integrating mobile applications with social media presents an opportunity for finding similar users and communities in a dynamic fashion. The health data that these applications store can be used to model the health status, which can then be used to find peers. These new techniques can be applied to create new personalized features, such as recommender systems of educational content [35,36]. Relying on peers for practical support and not entirely on primary health care may lessen the strain on health care resources.

Some applications, however, do synchronize their data with an online portal or PHR automatically. Despite previous research having reported the benefits of patient-accessible electronic health records [37], the rate of adoption of PHR is less than initially expected, partially because of usability issues [38]. Parts of the PHRs and portals can be shared with friends, family members, or physicians. Securely integrating health data and social media holds a potential for enhanced peer support. In the project TuAnalyze [39], researchers from Harvard Medical School have created an application for the diabetes online community TuDiabetes that gathers real time information about different health aspects. That social application is based on the PHR platform Indivo, and is a good example of a successful integration agenda.

Limitations

Many of the applications found outside the official online stores were not available for installation. As a result, some of the functionality was recorded from only the description or from published articles. Often there are discrepancies between the text description and the actual features, and some functionality is not apparent until the application is installed and tested. In research articles, some authors are not clear about whether any of the work they describe has been done. For example, Buranarach et al [40] and Chang et al [41] discuss Web-based portals for diabetes self-management, but we could not find the portals or verify some of the reported claims. We therefore had to cross-reference the descriptions using three independent review articles. The scope of this study did not include analysis of patient privacy and security issues. Few studies addressed this issue, which is a field in definite need of research.

Another limitation may be that the main inclusion criterion of selecting only applications with an SMBG feature had the potential to preclude other potentially relevant applications. However, in practice, most of the initial 973 search hits included general health information and health news aggregators. Also included in the initial hits were applications for managing physical activity, diet, or weight, but without special regard to diabetes. Most of the more general lifestyle applications such as diet management have a wider audience and are geared toward keeping healthy and fit. In order to maintain some level of focus, blood glucose tracking appeared to be an attractive inclusion criterion for diabetes mobile health applications, and also seemed to be a generally accepted core feature among developers of diabetes applications.

Since this study only analyzed the availability of applications and their features, it lacks information about the users. One cannot easily make generalizations to populations, since smartphones are not equally used among people of different socioeconomic status and age groups. In addition, it is very hard to know whether the applications are mostly used in developing countries where the available mobile technology may be less sophisticated.

Conclusions

A main finding from this review is that a critical feature strongly recommended by clinical guidelines—namely, personalized education—is not assimilated in current applications. Polonsky et al [28] and Klonoff [42] argue that studies fail to show SMBG benefits for patients who are not taking insulin because the studies have not integrated well-structured education as part of the intervention, and our results seem to support their premise.

The other major finding emerging from this study is that potentially interesting new use cases from social media are largely undeveloped. Although there is some evidence of the use of PHR in augmenting social engagement with peers, we found that the concept is still seldom recognized in the surveyed applications.

The impact of specific application features on clinical outcomes is not easy to determine, but current findings enhance our understanding of how the lack of some designated core features may influence clinical outcomes. The presented findings contribute evidence that shows personalized education and decision-support features not being integrated in most current blood glucose monitoring interventions, despite the evidence-based recommendations and requirement for them.

Acknowledgments

This work was supported in part by the Research Program for Telemedicine, Helse Nord RHF, Norway, and the Centre for Research-based Innovation, Tromsø Telemedicine Laboratory (TTL), Norwegian Research Council Grant No. 174934.

Conflicts of Interest

None declared

Multimedia Appendix 1

Mobile Apps - Data Supplement.

[XLS file (Microsoft Excel File), 110 KB - jmir_v13i3e65_app1.xls]

Multimedia Appendix 2

The different functions and features explained.

[PDF file (Adobe PDF File), 12 KB - jmir_v13i3e65_app2.pdf]

References

1. Lim S, Kang SM, Shin H, Lee HJ, Won Yoon J, Yu SH, et al. Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. Diabetes Care 2011 Feb;34(2):308-313 [FREE Full text] [doi: 10.2337/dc10-1447] [Medline: 21270188]
2. Lim S, Kim SY, Kim JI, Kwon MK, Min SJ, Yoo SY, et al. A survey on ubiquitous healthcare service demand among diabetic patients. Diabetes Metab J 2011 Feb;35(1):50-57 [FREE Full text] [doi: 10.4093/dmj.2011.35.1.50] [Medline: 21537413]
3. Lyles CR, Harris LT, Le T, Flowers J, Tufano J, Britt D, et al. Qualitative evaluation of a mobile phone and web-based collaborative care intervention for patients with type 2 diabetes. Diabetes Technol Ther 2011 May;13(5):563-569. [doi: 10.1089/dia.2010.0200] [Medline: 21406018]
4. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011 Apr;28(4):455-463. [doi: 10.1111/j.1464-5491.2010.03180.x] [Medline: 21392066]
5. 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;4(3):E16 [FREE Full text] [doi: 10.2196/jmir.4.3.e16] [Medline: 12554547]
6. Smith JC, Schatz BR. Feasibility of mobile phone-based management of chronic illness. AMIA Annu Symp Proc 2010;2010:757-761. [Medline: 21347080]
7. Tatara N, Arsand E, Nilsen H, Hartvigsen G. A review of mobile terminal-based applications for self-management of patients with diabetes. In: Proceedings. New York, NY: IEEE Computer Society; 2009 Feb 1 Presented at: International Conference oneHealth, Telemedicine, and Social Medicine (eTELEMED '09); 2009; Cancun, Mexico. [doi: 10.1109/eTELEMED.2009.14]
8. Kouris I, Mougiakakou S, Scarnato L, Iliopoulou D, Diem P, Vazeou A, et al. Mobile phone technologies and advanced data analysis towards the enhancement of diabetes self-management. Int J Electron Healthc 2010;5(4):386-402. [Medline: 21041177]
9. Kollmann A, Riedl M, Kastner P, Schreier G, Ludvik B. Feasibility of a mobile phone-based data service for functional insulin treatment of type 1 diabetes mellitus patients. J Med Internet Res 2007;9(5):e36 [FREE Full text] [doi: 10.2196/jmir.9.5.e36] [Medline: 18166525]
10. Ramadas A, Quek KF, Chan CK, Oldenburg B. Web-based interventions for the management of type 2 diabetes mellitus: a systematic review of recent evidence. Int J Med Inform 2011 Jun;80(6):389-405. [doi: 10.1016/j.ijmedinf.2011.02.002] [Medline: 21481632]
11. Avdal EU, Kizilci S, Demirel N. The effects of web-based diabetes education on diabetes care results: a randomized control study. Comput Inform Nurs 2011 Feb;29(2 Suppl):TC29-TC34. [doi: 10.1097/NCN.0b013e3182155318] [Medline: 21372642]
12. Raiff BR, Dallery J. Internet-based contingency management to improve adherence with blood glucose testing recommendations for teens with type 1 diabetes. J Appl Behav Anal 2010;43(3):487-491. [doi: 10.1901/jaba.2010.43-487] [Medline: 21358907]
13. Farmer AJ, Gibson OJ, Dudley C, Bryden K, Hayton PM, Tarassenko L, et al. A randomized controlled trial of the effect of real-time telemedicine support on glycemic control in young adults with type 1 diabetes (ISRCTN 46889446). Diabetes Care 2005 Nov;28(11):2697-2702 [FREE Full text] [Medline: 16249542]

14. Seto E, Istepanian RS, Cafazzo JA, Logan A, Sungoor A. UK and Canadian perspectives of the effectiveness of mobile diabetes management systems. Conf Proc IEEE Eng Med Biol Soc 2009;2009:6584-6587. [doi: 10.1109/IEMBS.2009.5333998] [Medline: 19964700]
15. Lutes KD, Baggili IM. Diabetic e-management system (DEMS). In: Proceedings. New York, NY: IEEE Computer Society; 2006 Apr 10 Presented at: Third International Conference on Information Technology: New Generations.; 2006; Las Vegas, Nevada, USA. [doi: 10.1109/itng.2006.53]
16. Quinn CC, Gruber-Baldini AL, Shardell M, Weed K, Clough SS, Peeples M, et al. Mobile diabetes intervention study: testing a personalized treatment/behavioral communication intervention for blood glucose control. Contemp Clin Trials 2009 Jul;30(4):334-346. [doi: 10.1016/j.cct.2009.02.004] [Medline: 19250979]
17. Faridi Z, Liberti L, Shuval K, Northrup V, Ali A, Katz DL. Evaluating the impact of mobile telephone technology on type 2 diabetic patients' self-management: the NICHE pilot study. J Eval Clin Pract 2008 Jun;14(3):465-469. [doi: 10.1111/j.1365-2753.2007.00881.x] [Medline: 18373577]
18. Årsand E, Tatara N, Østengen G, Hartvigsen G. Wireless and Mobile Technologies Improving Diabetes Self-Management: Handbook of Research on Mobility and Computing: Evolving Technologies and Ubiquitous Impacts. Hershey, PA: IGI Global; 2011.
19. Clar C, Barnard K, Cummins E, Royle P, Waugh N, Aberdeen Health Technology Assessment Group. Self-monitoring of blood glucose in type 2 diabetes: systematic review. Health Technol Assess 2010 Mar;14(12):1-140 [FREE Full text] [doi: 10.3310/hta14120] [Medline: 20226138]
20. Davidson MB. Evaluation of self monitoring of blood glucose in non-insulin-treated diabetic patients by randomized controlled trials: little bang for the buck. Rev Recent Clin Trials 2010 Sep;5(3):138-142. [Medline: 20482496]
21. Welschen LM, Bloemendal E, Nijpels G, Dekker JM, Heine RJ, Stalman WA, et al. Self-monitoring of blood glucose in patients with type 2 diabetes who are not using insulin: a systematic review. Diabetes Care 2005 Jun;28(6):1510-1517 [FREE Full text] [Medline: 15920083]
22. Paulweber B, Valensi P, Lindström J, Lalic NM, Greaves CJ, McKee M, et al. A European evidence-based guideline for the prevention of type 2 diabetes. Horm Metab Res 2010 Apr;42 Suppl 1:S3-36. [doi: 10.1055/s-0029-1240928] [Medline: 20391306]
23. Sibal L, Home PD. Management of type 2 diabetes: NICE guidelines. Clin Med 2009 Aug;9(4):353-357. [Medline: 19728510]
24. Funnell MM. Standards of care for diabetes: what's new? Nursing 2010 Oct;40(10):54-56. [doi: 10.1097/01.NURSE.0000388314.83393.bd] [Medline: 20844399]
25. ACCORD Study Group, Cushman WC, Evans GW, Byington RP, Goff DC, Grimm RH, et al. Effects of intensive blood-pressure control in type 2 diabetes mellitus. N Engl J Med 2010 Apr 29;362(17):1575-1585 [FREE Full text] [doi: 10.1056/NEJMoa1001286] [Medline: 20228401]
26. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
27. Stephens JW, Carman JE, Brooks CJ, Lyons RA, Ford DV, Price DE, et al. Comment on: Polonsky et al. Structured self-monitoring of blood glucose significantly reduces A1C levels in poorly controlled, noninsulin-treated type 2 diabetes: results from the structured testing program study. Diabetes Care 2011;34:262-267. Diabetes Care 2011 May;34(5):e57; author reply e58 [FREE Full text] [doi: 10.2337/dc11-0258] [Medline: 21525497]
28. Polonsky WH, Fisher L, Schikman CH, Hinnen DA, Parkin CG, Jelsovsky Z, et al. Response to Comment on: Polonsky et al. Structured self-monitoring of blood glucose significantly reduces A1C levels in poorly controlled, noninsulin-treated type 2 diabetes: results from the structured testing program study. Diabetes Care 2011;34:262–267. Diabetes Care 2011;34:e58. [doi: 10.2337/dc11-0365]
29. Polonsky WH, Fisher L, Schikman CH, Hinnen DA, Parkin CG, Jelsovsky Z, et al. Structured self-monitoring of blood glucose significantly reduces A1C levels in poorly controlled, noninsulin-treated type 2 diabetes: results from the Structured Testing Program study. Diabetes Care 2011 Feb;34(2):262-267 [FREE Full text] [doi: 10.2337/dc10-1732] [Medline: 21270183]
30. Bonomo K, De Salve A, Fiora E, Mularoni E, Massucco P, Poy P, et al. Evaluation of a simple policy for pre- and post-prandial blood glucose self-monitoring in people with type 2 diabetes not on insulin. Diabetes Res Clin Pract 2010 Feb;87(2):246-251. [doi: 10.1016/j.diabres.2009.10.021] [Medline: 19954855]
31. García-Mayor R. [Self-monitoring of blood glucose in type 2 diabetic patients non treated with insulin]. Med Clin (Barc) 2010 May 22;134(15):688-691. [doi: 10.1016/j.medcli.2009.11.022] [Medline: 20116074]
32. Chen Y. Take it personally: accounting for individual difference in designing diabetes management systems. In: Proceedings. 2010 Presented at: The 8th ACM Conference on Designing Interactive Systems; August 16 - 20; Aarhus, Denmark. [doi: 10.1145/1858171.1858218]
33. Frost JH, Massagli MP. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. J Med Internet Res 2008;10(3):e15 [FREE Full text] [doi: 10.2196/jmir.1053] [Medline: 18504244]
34. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]

35. Wiesner M, Pfeifer D. Adapting recommender systems to the requirements of personal health record systems. In: Proceedings. 2010 Presented at: The 1st ACM International Health Informatics Symposium; November 11-12, 2010; Arlington, Virginia, USA. [doi: 10.1145/1882992.1883053]
36. Fernandez-Luque L, Karlsen R, Vognild LK. Challenges and opportunities of using recommender systems for personalized health education. Stud Health Technol Inform 2009;150:903-907. [Medline: 19745443]
37. Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, Canadian Committee for Patient Accessible Health Records. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. J Med Internet Res 2008;10(4):e34 [FREE Full text] [doi: 10.2196/jmir.1061] [Medline: 18974036]
38. Liu LS, Shih PC, Hayes GR. Barriers to the adoption and use of personal health record systems. In: Proceedings. 2011 Presented at: The 2011 iConference; February 8-11, 2011; Seattle, Washington. [doi: 10.1145/1940761.1940811]
39. Weitzman ER, Adida B, Kelemen S, Mandl KD. Sharing data for public health research by members of an international online diabetes social network. PLoS One 2011;6(4):e19256 [FREE Full text] [doi: 10.1371/journal.pone.0019256] [Medline: 21556358]
40. Buranarach M, Thatphithakkul N, Supnithi T, Phetsawat R, Phienphanich P, Kawtrakul A, et al. Development of a personalized knowledge portal to support diabetes patient self-management. In: Proceedings. 2010 Presented at: International Conference on Management of Emergent Digital EcoSystems; October 26-29, 2010; Bangkok, Thailand. [doi: 10.1145/1936254.1936290]
41. Chang IC, Shih-Jung H, Hui-Mei H, Tzu-Hsuan C. Building mPHR to assist diabetics in self-healthcare management. In: Proceedings. 2010 Presented at: 7th International Conference on Service Systems and Service Management (ICSSSM); June 28-30, 2010; Tokyo. [doi: 10.1109/ICSSSM.2010.5530092]
42. Klonoff DC. Benefits and limitations of self-monitoring of blood glucose. J Diabetes Sci Technol 2007 Jan;1(1):130-132 [FREE Full text] [Medline: 19888392]

Abbreviations

ACCORD: Action to Control Cardiovascular Risk in Diabetes
mHealth: mobile health
PHR: personal health record
SMBG: self-monitoring of blood glucose

Edited by G Eysenbach; submitted 27.06.11; peer-reviewed by C Lyles; comments to author 11.07.11; revised version received 07.08.11; accepted 26.08.11; published 22.09.11.

Please cite as:
Chomutare T, Fernandez-Luque L, Årsand E, Hartvigsen G
Features of Mobile Diabetes Applications: Review of the Literature and Analysis of Current Applications Compared Against Evidence-Based Guidelines
J Med Internet Res 2011;13(3):e65
URL: http://www.jmir.org/2011/3/e65/
doi:10.2196/jmir.1874
PMID:21979293

Short Paper

Estimation of Physical Activity Levels Using Cell Phone Questionnaires: A Comparison With Accelerometry for Evaluation of Between-Subject and Within-Subject Variations

Christin Bexelius[1], PhD; Sven Sandin[1], MSc; Ylva Trolle Lagerros[2], MD PhD; Jan-Eric Litton[1], PhD; Marie Löf[1,3], PhD

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[2]Unit of Clinical Epidemiology, Department of Medicine, Karolinska Institutet, Stockholm, Sweden
[3]Department of Clinical and Experimental Medicine, University of Linköping, Linköping, Sweden

Corresponding Author:
Marie Löf, PhD
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet
PO 281
Stockholm, 17177
Sweden
Phone: 46 101034486
Fax: 46 8314975
Email: marie.lof@ki.se

Abstract

Background: Physical activity promotes health and longevity. Further elaboration of the role of physical activity for human health in epidemiological studies on large samples requires accurate methods that are easy to use, cheap, and possible to repeat. The use of telecommunication technologies such as cell phones is highly interesting in this respect. In an earlier report, we showed that physical activity level (PAL) assessed using a cell phone procedure agreed well with corresponding estimates obtained using the doubly labeled water method. However, our earlier study indicated high within-subject variation in relation to between-subject variations in PAL using cell phones, but we could not assess if this was a true variation of PAL or an artifact of the cell phone technique.

Objective: Our objective was to compare within- and between-subject variations in PAL by means of cell phones with corresponding estimates using an accelerometer. In addition, we compared the agreement of daily PAL values obtained using the cell phone questionnaire with corresponding data obtained using an accelerometer.

Methods: PAL was measured both with the cell phone questionnaire and with a triaxial accelerometer daily during a 2-week study period in 21 healthy Swedish women (20 to 45 years of age and BMI from 17.7 kg/m^2 to 33.6 kg/m^2). The results were evaluated by fitting linear mixed effect models and descriptive statistics and graphs.

Results: With the accelerometer, 57% (95% confidence interval [CI] 40%-66%) of the variation was within subjects, while with the cell phone, within-subject variation was 76% (95% CI 59%-83%). The day-to-day variations in PAL observed using the cell phone questions agreed well with the corresponding accelerometer results.

Conclusions: Both the cell phone questionnaire and the accelerometer showed high within-subject variations. Furthermore, day-to-day variations in PAL within subjects assessed using the cell phone agreed well with corresponding accelerometer values. Consequently, our cell phone questionnaire is a promising tool for assessing levels of physical activity. The tool may be useful for large-scale prospective studies.

(J Med Internet Res 2011;13(3):e70) doi:10.2196/jmir.1686

KEYWORDS

Cell phone; Internet; physical activity; epidemiology

Introduction

Large epidemiological studies on physical activity and health require accurate methods that are easy to use, cheap, and possible to repeat. Recently, the potential of using cell phones, either through short message service (SMS) or Web-like applications, in behavior change intervention studies have been explored (for example, [1-4]). However, cell phones also open new possibilities for data collection in large-scaled prospective studies [5,6].

When assessing physical activity, one important aspect is the total amount of energy expended due to physical activity. This variable can be obtained as total energy expenditure divided by basal metabolic rate, that is, the so-called physical activity level (PAL) [7]. We have recently developed a Java-based cell phone questionnaire to assess PAL that places little demand on either the study center or the participants [8]. Once every day for 2 weeks subjects are asked two questions via their cell phones about their daily physical activity. In an earlier study, the mean PAL during 2 weeks using this cell phone questionnaire agreed well with corresponding estimates based on the doubly labeled water method and indirect calorimetry (mean difference = 0.014, 1 standard deviation [SD] = 0.15) [8]. However, our earlier study showed a low variation in PAL between subjects (20% of the total variation) [8], indicating that the rest (80%) was caused by variations within subjects. Since the doubly labeled water method does not provide daily PAL values, we could not evaluate whether these variations were true or not. If they are not true, a low between-subject variation in our cell phone questionnaire may indicate that these estimates are not able to distinguish PAL between individuals very well. Thus, the aim of this study was to compare within- and between-subject variations in PAL by means of cell phones with corresponding estimates using an accelerometer. In addition, we compared the agreement of daily PAL values obtained using the cell phone questionnaire with corresponding data obtained using an accelerometer.

Methods

In all, 22 healthy nonsmoking Swedish women were recruited during August 2007 through May 2008 as previously described [8]. The women were 35.1 (range 20-45, SD 8.3) years of age, and their BMI was 23.7 kg/m^2 (range 17.7 kg/m^2 to 33.6 kg/m^2, SD 3.8 kg/m^2). PAL was measured daily during a 14-day study period using the cell phone questions as well as a tri-axial accelerometer, the RT3 (Stayhealthy Inc, Monrovia, CA, USA). The first day of the study period was excluded since the women did not wear the accelerometer until the afternoon that day; thus, results are reported for 13 days. One woman accidently broke her accelerometer; thus, results are reported for only 21 women. The study was approved by the central ethics board in Stockholm, Sweden.

Our JAVA-based questionnaire for assessing PAL using cell phones has previously been described in detail [8]. Briefly, at 9 pm, each woman was asked two questions about her physical activity during the same day (Table 1). For each woman and for each day, the answers to the two short questions were converted to PAL by adding the PAL points obtained for work/day time activities [9] and the PAL points obtained for leisure/evening activities [10] (Table 1).

Table 1. The cell phone questionnaire consisting of two questions

Cell Phone Question	PAL Points
Question 1: How physically active have you been during work/the daytime today?	
Mostly sitting	1.55
Sitting/standing/walking	1.65
Standing/walking most of the time	1.85
Heavy work	2.2
Question 2: How physically active have you been during leisure time/the evening today?	
Mostly sitting	+0
Light/walking 30 minutes	+0.06
Moderate/cycling ≥ 30 minutes	+0.15
Sport/cycling ≥ 60 minutes	+0.29

Each woman was instructed to wear the accelerometer during all her waking hours except when in water. She recorded in a notebook when she took off the device and the activities performed without it (eg, showering or sleeping). Recorded movements were converted to total energy expenditure according to the manufacturer. All women wore the RT3 all 14 days and the recordings obtained during these days covered 97% ± 2% of time awake. For activities reported in the notebook, energy expenditure was estimated based on published energy costs of specific activities [10]. PAL was obtained as total energy expenditure divided by resting metabolic rate (calculated using in-built equations provided by the RT3 manufacturer) for each 24-hour period during days 1 to 13 (starting at 9 pm the first study day).

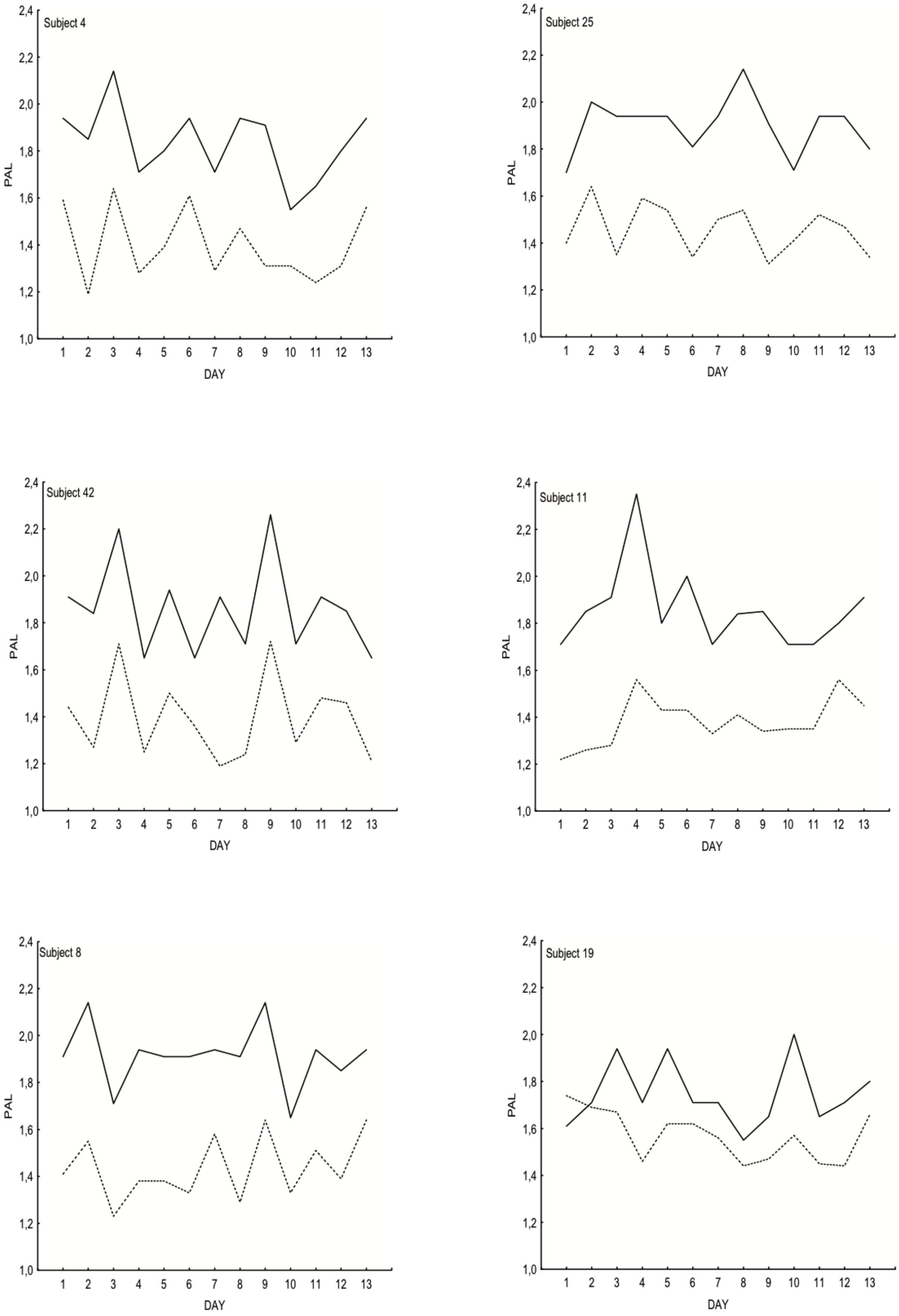
Intraclass correlation coefficients (ICC) were calculated as the between-subject variation divided by the total variation.The coefficients were estimated by means ofa linear mixed effect model with the two components of variance and a fixed effect for days 1 to 13. The ICC was used to compare the between-subject variation in relation to total variation for PAL values at days 1 to 13. The ICC was calculated from models fitted separately for the cell phone data and the accelerometer data. To test the robustness against the influence of single gross outliers, we repeated the ICC calculations for both methods when we had replaced the 10 as well as 20 most deviating daily PAL values out of 273 values with the mean PAL for these women. Furthermore, the agreement of daily PAL values obtained using the cell phone questionnaire and corresponding accelerometer data was evaluated using descriptive statistics. All statistical analyses were carried out using the SAS software proc mixed, version 9.2 (SAS Institute, Inc, Cary, NC, USA).

Results

For PAL obtained using the cell phone, the variance between and within subjects was estimated to be 0.009 and 0.029, respectively. Thus, 24% of the total variation was between subjects, and 76%, within subjects. The corresponding estimates for the accelerometer were 0.017 and 0.023, showing that 43% of the variation was between subjects and 57%, within subjects. The 95% CI for the within-subject variation was 59% to 83 % for the cell phone while it was 40% to 66% for the accelerometer. The variance components for PAL obtained using the cell phone and the accelerometer remained similar when we replaced the 10 and 20 most deviating values with the mean PAL of these women.

Figure 1 compares daily PAL values obtained with cell phones and accelerometers from day 1 to day 13 for 6 randomly selected subjects. Daily PAL values obtained with the accelerometer were generally on a lower level than the corresponding cell phone estimates, but the day-to-day changes in PAL observed using the cell phone questions followed the corresponding changes in the accelerometer PAL. Similar agreement between daily PAL values obtained using the two methods was observed for the other 15 women (data not shown).

Figure 1. Daily measures of PAL over a 13-day period for 6 randomly chosen individuals (the dotted line indicates accelerometer data and the straight line indicates PAL data from cell phones).



Discussion

Both the cell phone questionnaire and the accelerometer showed high within-subject variations, indicating that there is considerable true day-to-day variation in PAL. First, this makes it more likely that the relative low between-subject variations we reported in our earlier paper are not artifacts of the cell phone technique. Second, this is an interesting finding since few reports have described the specific sources of variance in daily physical activity using objective measures in adults [11-13].

PAL obtained using the accelerometer was generally on a lower level than the cell phone value. The RT3 as well as several other accelerometers have shown a tendency to underestimate energy expenditure during free-living conditions compared with the doubly labeled water method [14-15] also in the women in this study (Löf et al, unpublished data). Since the changes in daily PAL values obtained using the RT3 accelerometer agreed well with the corresponding cell phone estimates, we hypothesize that the underestimation in PAL results from an underestimation of energy expenditure from movements. Future studies should explore why the RT3 underestimates energy expenditure and how the monitor may be refined in order to improve its predictions.

With the accelerometer, 43% of the variation in PAL was between subjects, while between-subject variation was 24% with the cell phone. For comparison, Matthews et al [12] reported that differences between subjects accounted for 55% to 60% of the variation in accelerometer counts in 92 healthy adults, but their study covered a wider age range (18 to 79 years of age) than our study. For a paper questionnaire, the between-subject variation in overall physical activity was 20% to 30% in 580 healthy adults [16]. As discussed by Matthews et al, the variance structure between self-reported and objective measures may differ due to different inherent errors [12]. A limitation of accelerometers is that they are not sensitive to activities involving upper body movements such as weight-lifting or carrying. Weight-lifting was not common, but we have no information about carrying. If carrying was evenly distributed among our women, the variance components would be unchanged, but if carrying was unevenly distributed, the true between-subject variation using the accelerometer may be even higher. Thus, the results likely indicate that the cell phones underestimate the between-subject variation. One plausible explanation is the answer options. The highest category for the second question only mentions sports and cycling, while home chores like gardening are not included. Such activities could be added to refine the capability of the cell phone questionnaire to assess between-subject variation in PAL.

The major limitation of this study is the small sample size. First, it makes our analysis sensitive to single gross outliers since these may largely increase the within-subject variations. However, we found no evidence for any important effect of such outliers in our analysis since the variance components were stable when we replaced the 10 or 20 most deviating PAL values (out of totally 273) with the average PAL for those women. Second, the small sample size limits our ability to make firm conclusions about the between-subject variations. We cannot exclude that the low between-subject variation obtained by both methods to some extent is due to the fact that our participants were a relatively small group of women. Other limitations are that we estimated energy expenditure for activities when the women were not wearing the RT3, but this amount of time was small, and that this study was conducted in healthy, moderately active women. The results should be repeated in other populations including men as well as elderly and obese subjects.

In conclusion, (1) Both the cell phone questionnaire and the accelerometer showed high within-subject variations in PAL and (2) changes in daily cell phone PAL values agreed well with corresponding accelerometer values. This study adds further evidence to our earlier findings [8] that the cell phone questionnaire is a promising tool for assessing PAL in epidemiological studies.

Acknowledgments

The authors would like to thank all the women who participated in the study and Novatelligence for Medipal, with special thanks to Johan Cederlund. This study was supported by the following grants awarded to Marie Löf: the Swedish Research Council, FORMAS, the Magnus Bergvall Foundation, the Thuring Foundation, and the Swedish Society of Medicine.

Conflicts of Interest

None declared

Authors' Contributions

ML was responsible for the recruitment of subjects and data collection for the accelerometer. CB collected data using cell phones and conducted the data analyses (in collaboration with SS who was responsible for the statistical analyses). CB together with ML prepared the manuscript, which was subsequently reviewed by JEL, YTL, and SS. All authors contributed to the interpretation of the results, and all authors have approved the final manuscript.

References

1. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
2. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]
3. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
4. Woolford SJ, Clark SJ, Strecher VJ, Resnicow K. Tailored mobile phone text messages as an adjunct to obesity treatment for adolescents. J Telemed Telecare 2010;16(8):458-461. [doi: 10.1258/jtt.2010.100207] [Medline: 20959393]

5. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
6. Aanensen DM, Huntley DM, Feil EJ, al-Own F, Spratt BG. EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection. PLoS One 2009;4(9):e6968 [FREE Full text] [doi: 10.1371/journal.pone.0006968] [Medline: 19756138]
7. FAO/WHO/UNU Expert Consultation. Human Energy Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation. Rome, Italy: Food & Agriculture Organization of the United Nations; 2004.
8. Bexelius C, Löf M, Sandin S, Trolle Lagerros Y, Forsum E, Litton JE. Measures of physical activity using cell phones: validation using criterion methods. J Med Internet Res 2010;12(1):e2 [FREE Full text] [doi: 10.2196/jmir.1298] [Medline: 20118036]
9. Black AE, Coward WA, Cole TJ, Prentice AM. Human energy expenditure in affluent societies: an analysis of 574 doubly-labelled water measurements. Eur J Clin Nutr 1996 Feb;50(2):72-92. [Medline: 8641250]
10. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S498-S504. [Medline: 10993420]
11. Gretebeck RJ, Montoye HJ. Variability of some objective measures of physical activity. Med Sci Sports Exerc 1992 Oct;24(10):1167-1172. [Medline: 1435166]
12. Matthews CE, Ainsworth BE, Thompson RW, Bassett DR. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc 2002 Aug;34(8):1376-1381. [Medline: 12165695]
13. Levin S, Jacobs DR, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. Ann Epidemiol 1999 Nov;9(8):481-488. [Medline: 10549881]
14. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity (Silver Spring) 2007 Oct;15(10):2371-2379 [FREE Full text] [doi: 10.1038/oby.2007.281] [Medline: 17925461]
15. Maddison R, Jiang Y, Hoorn SV, Mhurchu CN, Lawes CM, Rodgers A, et al. Estimating energy expenditure with the RT3 triaxial accelerometer. Res Q Exerc Sport 2009 Jun;80(2):249-256. [Medline: 19650390]
16. Matthews CE, Hebert JR, Freedson PS, Stanek EJ, Merriam PA, Ebbeling CB, et al. Sources of variance in daily physical activity levels in the seasonal variation of blood cholesterol study. Am J Epidemiol 2001 May 15;153(10):987-995 [FREE Full text] [Medline: 11384955]

Abbreviations

CI: confidence interval
ICC: Intraclass correlation coefficient
PAL: physical activity level
SMS: Short Message Service

Edited by G Eysenbach; submitted 17.11.10; peer-reviewed by E Hekler; comments to author 19.02.11; revised version received 19.05.11; accepted 19.05.11; published 25.09.11.

Please cite as:
Bexelius C, Sandin S, Trolle Lagerros Y, Litton JE, Löf M
Estimation of Physical Activity Levels Using Cell Phone Questionnaires: A Comparison With Accelerometry for Evaluation of Between-Subject and Within-Subject Variations
J Med Internet Res 2011;13(3):e70
URL: http://www.jmir.org/2011/3/e70/
doi:10.2196/jmir.1686
PMID:21946080

Original Paper

Can Online Consumers Contribute to Drug Knowledge? A Mixed-Methods Comparison of Consumer-Generated and Professionally Controlled Psychotropic Medication Information on the Internet

Shannon Hughes[1], PhD; David Cohen[2], PhD

[1]Utah State University, Department of Sociology, Social Work, and Anthropology, Logan, UT, United States

[2]Florida International University, Robert Stempel College of Public Health and Social Work, Miami, FL, United States

Corresponding Author:
Shannon Hughes, PhD
Utah State University
Department of Sociology, Social Work, and Anthropology
0730 Old Main
Room 240A
Logan, UT, 84322
United States
Phone: 1 435 797 8636
Fax: 1 435 797 1240
Email: shughes5@msn.com

Abstract

Background: Ongoing initiatives to filter online health searches exclude consumer-generated content from search returns, though its inferiority compared with professionally controlled content is not demonstrated. The antidepressant escitalopram and the antipsychotic quetiapine have ranked over the last 5 years as top-selling agents in their respective drug classes. Both drugs have various off-label mental health and non–mental health uses, ranging from the relief of insomnia and migraines to the treatment of severe developmental disorders.

Objective: Our objective was to describe the most frequently reported effects of escitalopram and quetiapine in online consumer reviews, to compare them with effects described in professionally controlled commercial health websites, and to gauge the usability of online consumer medication reviews.

Methods: A stratified simple random sample of 960 consumer reviews was selected from all 6998 consumer reviews of the two drugs in 2 consumer-generated (www.askapatient.com and www.crazymeds.us) and 2 professionally controlled (www.webmd.com and www.revolutionhealth.com) health websites. Professional medication descriptions included all standard information on the medications from the latter 2 websites. All textual data were inductively coded for medication effects, and intercoder agreement was assessed. Chi-square was used to test for associations between consumer-reported effects and website origination.

Results: Consumers taking either escitalopram (n = 480) or quetiapine (n = 480) most frequently reported symptom improvement (30.4% or 146/480, 24.8% or 119/480) or symptom worsening (15.8% or 76/480, 10.2% or 49/480), changes in sleep (36% or 173/480, 60.6% or 291/480) and changes in weight and appetite (22.5% or 108/480, 30.8% or 148/480). More consumers posting reviews on consumer-generated rather than professionally controlled websites reported symptom worsening on quetiapine (17.3% or 38/220 versus 5% or 11/220, $P < .001$), while more consumers posting on professionally controlled websites reported symptom improvement (32.7% or 72/220 versus 21.4% or 47/220, $P = .008$). Professional descriptions more frequently listed physical adverse effects and warnings about suicidal ideation while consumer reviews emphasized effects disrupting daily routines and provided richer descriptions of effects in context. The most recent 20 consumer reviews on each drug from each website (n = 80) were comparable to the full sample of reviews in the frequency of commonly reported effects.

Conclusion: Consumer reviews and professional medication descriptions generally reported similar effects of two psychotropic medications but differed in their descriptions and in frequency of reporting. Professional medication descriptions offer the advantage of a concise yet comprehensive listing of drug effects, while consumer reviews offer greater context and situational examples of how effects may manifest in various combinations and to varying degrees. The dispersion of consumer reviews across websites limits their integration, but a brief browsing strategy on the two target medications nonetheless retrieved representative consumer content. Current strategies for filtering online health searches to return only *trusted* or *approved* websites may inappropriately address the challenge to identify quality health sources on the Internet because such strategies unduly limit access to an entire complementary source for health information.

(J Med Internet Res 2011;13(3):e53) doi:10.2196/jmir.1716

KEYWORDS

Psychotropic drugs; mental health; consumer health information; Internet; pharmacoepidemiology; drug monitoring; product surveillance, postmarketing

Introduction

Consumers and clinicians increasingly consult consumer-generated health content on the Internet [1-3], but there are no direct comparisons of such content with that found on professionally controlled commercial health websites. Ongoing initiatives in Internet searching aim to filter health-related searches to return only sources meeting medical grading system requirements such as depth, timeliness, transparency, and readability [4,5]. These so-called trusted sources typically include broadly networked and well-resourced commercial, institutional, and government websites representing a professional knowledge base but exclude consumer-generated content [5-7]. Despite widespread discussion and speculation about the varying quality of health information on the Internet, such initiatives may be premature in the absence of reliable evidence suggesting that nonprofessionally delivered content is necessarily inferior to that provided by professionally controlled health sites [8,9].

In parallel, there is increasing momentum to gather patient-reported health and treatment outcomes [10,11], with the Internet identified as a major mechanism to accomplish this efficiently [12-14]. While much research has focused on developing and implementing new Internet-based technologies to collect patient-reported outcomes, studies on the practical uses of existing consumer-generated online health content remain limited in number and scope. In the mental health arena, for example, researchers have described discussion themes in online support groups [15], the efficacy of such groups to reduce depressive symptoms [16], online help-seeking behaviors [17], and completeness of drug information on pharmaceutical company websites [18-20]. An analysis of 1 year of comments from an online discussion forum identified 238 drug-related problems with antiparkinsonian agents, noting incongruences with clinical trial data [21]. Online consumer comments were also employed to analyze the subjective effects of older and newer antipsychotic medications [22]. Finally, the online patient community PatientsLikeMe collects longitudinal data on treatment effects for various conditions from their members and presents these data to the scientific community [23,24].

Although consumer-generated content about psychiatric medications may take many forms, much of it appears online as brief (usually 1 to 3 paragraphs) first-person accounts or reviews of experiences around the ingestion of a prescribed medication. This study describes, for two widely prescribed psychotropic drugs, the most frequently reported effects in online consumer reviews found both in consumer-generated and professionally controlled commercial health websites. It also compares consumer-reported effects of the two drugs to the authoritative account of these drugs’ effects found in professionally controlled commercial health websites. Moreover, it does so by privileging neither source as an a priori standard for quality and accuracy or by using standardized drug effect terminology. The findings provide the first empirical basis to evaluate possible advantages and disadvantages of using each online source (consumer-generated and professionally controlled health sites) for making medication-related treatment decisions.

Methods

Website Sampling

Since consumer-generated health content is unlikely to be returned among top search engine results, search engines and an index of online mental health resources compiled over the last 16 years by www.psychcentral.com were both employed to identify 2 consumer-generated websites. Combinations of the following search terms were used in Google and Yahoo search engines: patient, consumer, review, rating, support, and Lexapro (or Seroquel). The top 50 returns in each search engine as well as the online index of mental health resources previously cited were screened according to the following inclusion criteria: (1) all consumer commentary was viewable without requiring registration or membership conditioned on moderator approval, and (2) the website contained at least 200 consumer comments for each drug. This search resulted in the consumer-generated websites www.askapatient.com and www.crazymeds.us. The former website contains pre-defined fields for users to input a 1 to 5 numerical rating of their satisfaction of the drug as well as their diagnosis, drug side effects, open-ended comments, sex, age, time taken, and dosage. The drug reviews then accumulate in a tabular format with little additional content on the website. The latter website hosts a discussion forum in which conversation threads are structured according to drug class and brand name. Neither website is monitored or edited by medical or health professionals, nor are postings edited for any reasons other than inappropriate content (ie, vulgar language or threats of self harm).

Professionally controlled commercial health websites (hereinafter referred to as “professionally controlled websites”) take the form of information portals monitored by health professionals and intended for a general audience of consumers and clinicians seeking a broad variety of online health information. The key criterion used in this study to identify professionally controlled websites was the oversight of content by a team of medical professionals, usually medical journalists who gather and write the content and medical doctors who provide oversight and consultation. Other inclusion criteria were: (1) the website was a commercial health portal (not operated by a governmental group or organization), (2) the given disclosures provided no evidence that the selected websites were owned by the same company or that they shared professional contributors, and (3) the website had received accolades for excellence in providing online health content. Governmental websites were excluded based on the reasoning that commercial health sites might be more likely to make concerted efforts toward assessing and adjusting website structure, functionality, and content in order to appeal to a broad and general audience and thereby increase site traffic. Google and Yahoo search engines were used to identify 2 professionally controlled health websites that were listed among the top 20 returns in searches of drugs’ brand names and met all inclusion criteria. This sampling resulted in the websites www.webmd.com and www.revolutionhealth.com. Content in both websites is pulled from a network of partners, including clinics, other health news sources, and health publishers, with oversight provided by health professionals and medical writers. Both websites are highly trafficked and lauded as reputable resources for up-to-date, authoritative health and treatment information.

Both professionally controlled websites also provide space for consumers to post ratings and reviews of drug treatment experiences. On WebMD, consumers are prompted to share a numerical 1 to 5 rating on the effectiveness, ease of use, and satisfaction of the drug, their diagnosis, age range, sex, how long they have taken the drug, and an open-ended comment. On RevolutionHealth, consumers are prompted to share a 1 to 10 rating on the effectiveness, ease of use, tolerability, and recommendation for the drug, their diagnosis, and an open-ended comment.

Case Sampling

Escitalopram and quetiapine were selected as the points of entry for this study because both were top-selling drugs in the antidepressant and antipsychotic classes, respectively, at the time of this research [25-28]. Escitalopram was first approved by the US Food and Drug Administration (FDA) in 2002 for the treatment of depression and since 2006 has consistently earned an average US $2.5 billion in annual US retail sales. Quetiapine was first approved by the FDA in 1997 for the treatment of schizophrenia and was ranked as the ninth best-selling drug in 2006 with US $3 billion in US retail sales. It has maintained and exceeded that level of revenue in more recent years. Both drugs are also commonly used for numerous off-label purposes, including developmental disorders, anxiety, depression, and insomnia for quetiapine, and panic, social anxiety, premenstrual dysphoric disorder, and migraines for escitalopram [29].

All consumer reviews and commentary about the 2 drugs from the 4 websites through the end of February 2009 were imported into QDA Miner 3.2 data analysis software (see Multimedia Appendix 1) [30]. Each individual consumer was considered a single case. The comparison group of professional medication descriptions was retrieved from the 2 professionally controlled websites by importing all main text (excluding advertisements) returned from a search of the medications into QDA Miner 3.2 software. On WebMD, this text included the professionally controlled information on drug warnings, uses, side effects, precautions, interactions, and overdose. On RevolutionHealth, it included drug uses, side effects, dosage, interactions, and a section titled “Important information.”

Data collection resulted in a sampling frame of 6998 consumer cases (see Table 1) and the professional medication descriptions (all text for 2 medications on 2 websites). A stratified simple random sample of 120 consumer cases per drug per website (13.7% of the sampling frame) resulted in a coding sample of 960 cases (escitalopram, n = 480; quetiapine, n = 480). Since the sampling frame was not evenly distributed across websites, as illustrated in Table 1, this sampling strategy had the effect of oversampling consumer reviews on consumer-generated websites. Equal representation of consumer reviews from each website was thus ensured, and the coding sample became more manageable in size. All 4 professional medication descriptions were included in the analysis.

Online consumer reviews were regarded in this study to be part of the public domain [31], and no personally identifiable information was collected. The Florida International University Office of Research Integrity approved this study.

Table 1. Website description and sampling frame for consumer reviews

Description	Selected Websites	N Consumer Reviews	
		Escitalopram	Quetiapine
Professionally controlled commercial health websites			
Created and monitored by health professionals Reflect recognized standards of scientific/medical excellence Intended for lay and professional audiences May include pages where site users review and rate treatments	www.webmd.com[a] www.revolutionhealth.com[a]	1402 1873	722 624
Consumer-generated health websites			
Not monitored or edited by medical or health professionals Contain only or mostly consumer-generated contributions (but may display some ads) Include ≥ 200 consumer reviews for each of escitalopram and quetiapine	www.askapatient.com www.crazymeds.us	1093 266	791 227
	Total	4634	2364

[a] The comparison group of professional medication descriptions was retrieved by copying all textual drug information returned from searches of each medication on the professionally controlled commercial health websites.

Coding

Author SH developed a codebook by inductively coding 85 randomly selected consumer cases from the sampling frame using initial and focused coding procedures [32]. This strategy was selected because a primary research aim was to explore consumer medication reviews on their own terms rather than fit them into a standardized vocabulary. Initial coding aimed to capture and condense literal meanings of reported medication effects with as little interpretation as possible. In keeping with the grounded theory approach, consumer text language was preserved. For example, descriptions such as “extreme sleepiness” were used as code names instead of the standard professional codes *drowsiness* or *somnolence*. Next, focused coding involved refining the initial codes to develop more definitive effect categories. Constant comparisons of data to data were used to ensure consistency in grouping drug effects. The final codebook identified 70 drug effects in consumer and professional text (eg, low libido, increased libido, trouble achieving orgasm) grouped into 11 effect categories (eg, sexual effects, see Figure 1). The present analysis describes the 5 most frequently reported drug effects.

Coding Agreement Analysis

A coding agreement analysis was conducted by author SH and another independent coder on 191 (20%) randomly selected cases. Intercoder agreement was calculated in QDA Miner 3.2 for each effect category on the level of code occurrence within a case using Scott’s pi (≥ .70 prespecified to indicate acceptable intercoder agreement) [33]. Both coders coded the first 100 cases and a Scott’s pi was calculated. The coders together reviewed each disagreement and came to a mutual decision about its resolution. After discussing individual coding decisions, the coders agreed upon collapsing or splitting some codes. The process was repeated with the next 91 cases. Author SH then coded the remaining 769 cases in the sample.

Data Analysis

Frequency tables summarized consumer-reported drug effects. To compare consumer-reported effects and professional medication descriptions, we estimated the relative attention each group gave to specific effects by calculating the proportion of mentions of an effect out of all mentions of effects. Chi-square was calculated to test the null hypothesis of no association between website origination and consumer-reported drug effects. Significance tests were two-tailed and corrections were made for multiple comparisons by dividing the alpha level of .05 by k number of comparisons. Excerpts from text were extracted to illustrate differences in descriptions between consumer-generated and professionally controlled text.

An online health seeker is likely to visit only a few pages from each of 2 to 5 websites when researching health information online [3]. Therefore, the systematic evaluation of hundreds of consumer reviews performed in the present analysis has limited relevance to the everyday use of online consumer reviews for making treatment decisions. To simulate how a typical Internet user might consult consumer reviews while searching for medication-related information, then, the most recent 20 consumer comments on each drug from each website (n = 80) were compared for representativeness to all remaining consumer comments on that drug (n = 400). Chi-square was calculated to test the null hypothesis of no difference between recent and all remaining comments.

Results

Consumer Characteristics

Most consumers on AskaPatient and WebMD reported their gender, age, and length of time on the drug, while most consumers on the remaining 2 websites did not report gender or age, and at least half did not report length of time on the drug. Table 2 provides demographic characteristics for the consumers in this sample according to website on which the comment was posted. Table 3 shows the same information according to medication.

Table 2. Consumer characteristics according to website

	AskaPatient	CrazyMeds	WebMD	Revolution-Health	Total
Total n in each sample	240	240	240	240	960
Characteristics	n (%)	n (%)	n (%)	n (%)	n (%)
Gender					
Female	158 (66)	64 (27)	179 (75)	22 (9)	423 (44)
Male	81 (34)	27 (11)	50 (21)	10 (4)	168 (17.5)
Not given	1 (< 1)	149 (62)	11 (5)	208 (87)	369 (38.4)
Age in years					
≤ 18	15 (6)	3 (1)	5 (2)	2 (1)	25 (2.6)
19–34	116 (48)	3 (1)	82 (34)	3 (1)	204 (21.3)
35–54	93 (39)	10 (4)	103 (43)	6 (3)	212 (22.1)
≥ 55	13 (5)	2 (1)	36 (15)	1 (<1)	52 (5.4)
Not given	3 (1)	224 (93)	14 (6)	228 (95)	469 (48.9)
Length of time on drug					
< 1 month	67 (28)	23 (10)	51 (21)	43 (18)	184 (19.2)
1–6 months	72 (30)	38 (16)	59 (25)	42 (18)	211 (22)
6 months–2 years	58 (24)	37 (15)	55 (23)	17 (7)	167 (17.4)
≥ 2 years	41 (17)	24 (10)	56 (23)	10 (4)	134 (15)
Not given	1 (<1)	120 (50)	17 (7)	128 (53)	266 (27.7)

Table 3. Consumer characteristics according to drug

	Escitalopram	Quetiapine
Total n in each sample	480	480
Characteristics	n (%)	n (%)
Gender		
Female	216 (45)	207 (43.1)
Male	77 (16)	91 (19)
Not given	187 (39)	182 (37.9)
Age in years		
≤ 18	7 (1.5)	18 (3.8)
19–34	102 (21.2)	101 (21.2)
35–54	113 (24)	99 (20.7)
≥ 55	26 (4.4)	26 (5)
Not given	232 (48.3)	235 (49)
Length of time on drug		
< 1 month	106 (22.1)	78 (16.2)
1–6 months	122 (25.4)	89 (18.5)
6 months–2 years	85 (17.6)	82 (16.9)
≥ 2 years	55 (11.5)	79 (14.4)
Not given	112 (23.3)	152 (31.7)

Intercoder Agreement

Table 4 shows that acceptable overall intercoder agreement was obtained (average Scott's pi for all categories = .90 in phase 1, .82 in phase 2). Only for the category of *other effects* was agreement clearly unsatisfactory (< .41) because the miscellaneous effects included in it were grouped only after a substantial amount of coding had been completed.

Table 4. Intercoder agreement results

Drug Effect Categories	Scott's Pi, Part 1 n = 100	Scott's Pi, Part 2 n = 91
Appetite and weight	1	.84
Gastrointestinal and urinary	.88	1
Head and face	.82	.67
Lab tests and chronic conditions	1	.88
Mental and mood	.94	.74
Musculoskeletal and neurological	.82	.80
Nose, throat, and chest	1	.80
Sexual	.95	1
Skin	1	.65
Sleep	.93	.87
Other	.41	.33
Average overall	.90	.82

Consumer Reported Effects

The most frequently mentioned effects by the 480 sampled consumers taking each drug were related to symptom improvement or worsening, and changes in sleep, weight, and appetite (see Multimedia Appendix 1). About one-fifth of escitalopram consumers also reported sexual effects. Approximately 30% (146/480) of consumers taking escitalopram and 25% (119/480) taking quetiapine reported an improvement in anxiety, depression, mania, or other symptoms. Another 15.8% (76/480) of consumers taking escitalopram and 10.2% (49/480) taking quetiapine reported new or worsening symptoms as an effect of the medication, including new or worsened panic attacks, depression, mania, or hallucinations. Significantly more consumers posting medication reviews on the consumer-generated (AskAPatient and CrazyMeds) than on the professionally controlled websites (WebMD and RevolutionHealth) reported quetiapine worsened their symptoms (17.3% or 38/220 versus 5% or 11/220, $P < .001$, significance level set at .002 for k = 25 comparisons), while more consumers posting reviews on professionally controlled sites reported it improved their symptoms (32.7% or 72/220 versus 21.4% or 47/220 $P = .008$). This trend held for escitalopram without reaching statistical significance.

Over 60% (291) of the 480 consumers taking quetiapine reported effects on sleep, with 35.6% (171) commenting the medication helped their sleep and 33.1% (159) that it caused excessive sleep and tiredness. For the 480 consumers taking escitalopram, sleep changes indicating excessive sleep were mentioned by 23.8% (114), while 13.3% (64) of consumers mentioned insomnia. For quetiapine, over 30% (148) of consumers reported a range of appetite and weight effects, most notably weight gain (22.5% or 108/480). For escitalopram, 13.1% (63) of consumers reported weight gain and 4.8% (23) weight loss. Finally, 20.2% (97) of consumers taking escitalopram reported sexual effects, primarily in the form of low libido (10.6% or 51/480) and trouble achieving orgasm (8.5% or 41/480).

Consumer Reported Effects Compared With Professional Medication Descriptions

Figure 1 lists the relative frequency of mentions of drug effects in consumer reviews and professional medication descriptions across the 11 effect categories, and Figure 2 compares consumer reviews and professional descriptions on the most frequently mentioned effects for each drug as a proportion of all mentions of effects in each of the respective texts. For both medications, professional descriptions on WebMD and RevolutionHealth frequently mentioned worsening mental or mood effects, such as agitation and suicidal thinking, as well as physical effects, including dizziness, weakness, and vision problems. Other miscellaneous effects (such as toothache or bronchitis) were also more frequently mentioned in professional descriptions from both websites compared with consumer reviews.

Figure 1. Relative frequency of mentions of effects in consumer reviews and in professional medication descriptions according to effect category

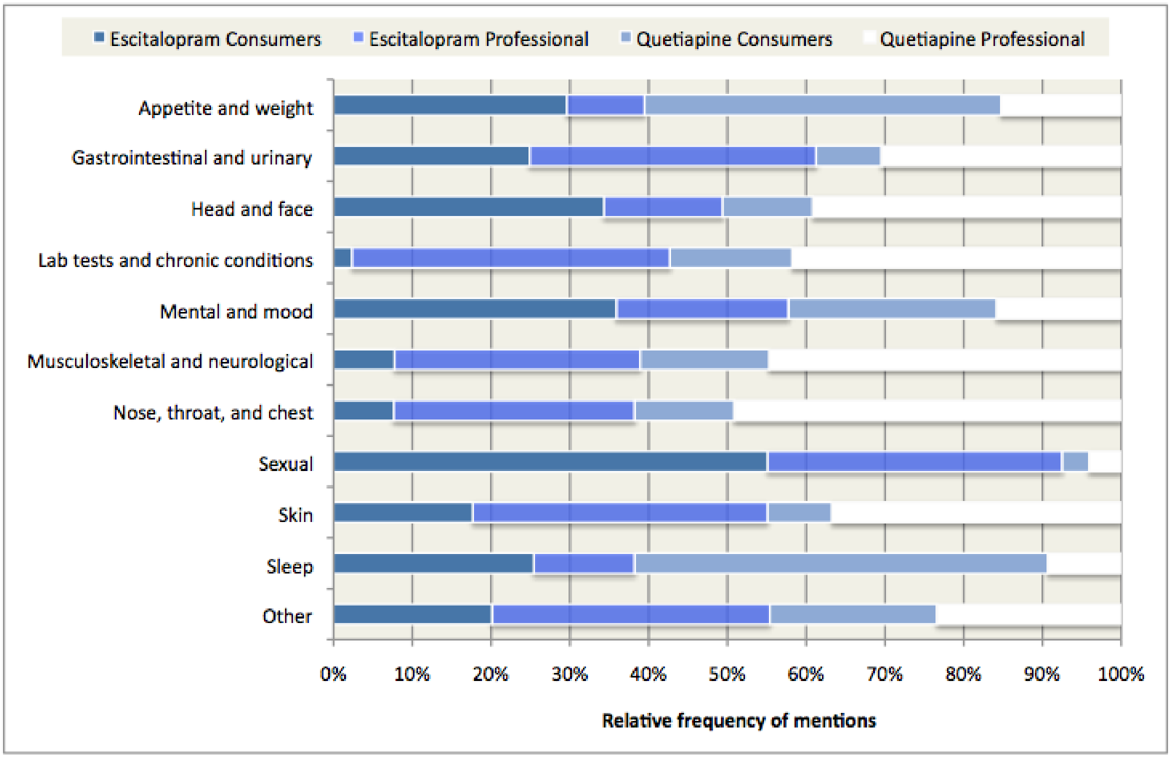


Figure 2. Most frequently mentioned effects as a proportion of all mentions of escitalopram effects in consumer-generated and in professionally controlled text

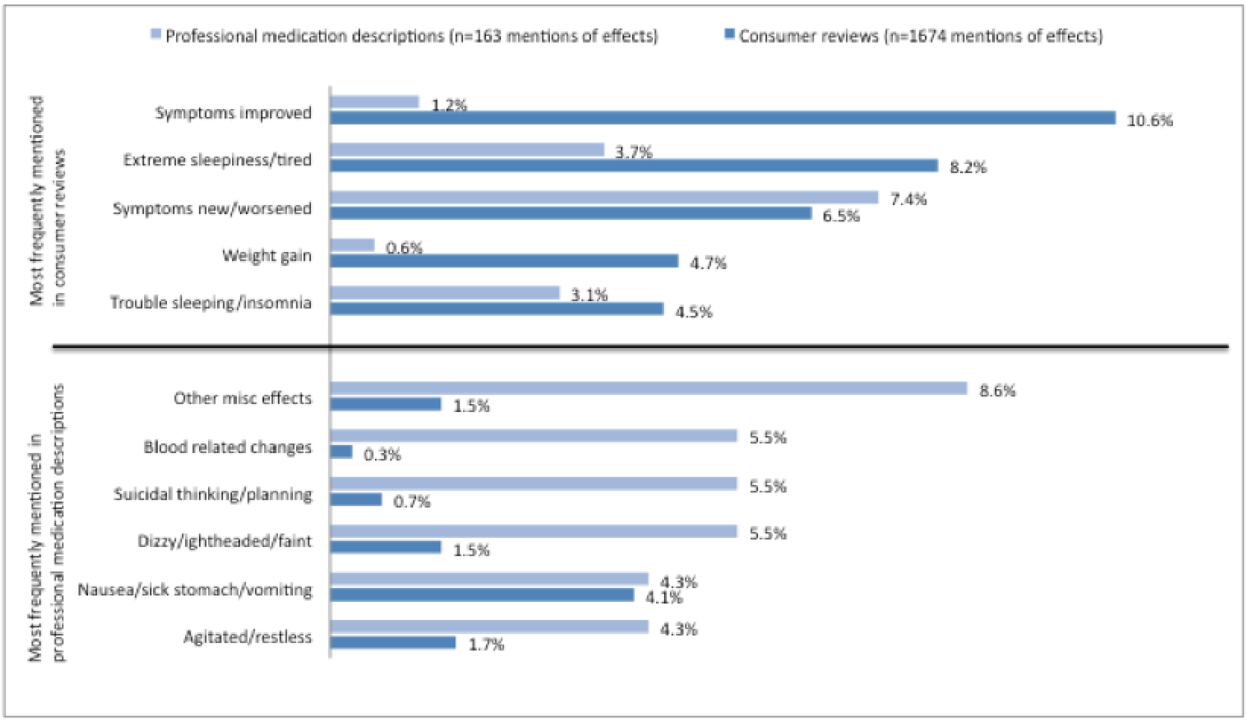
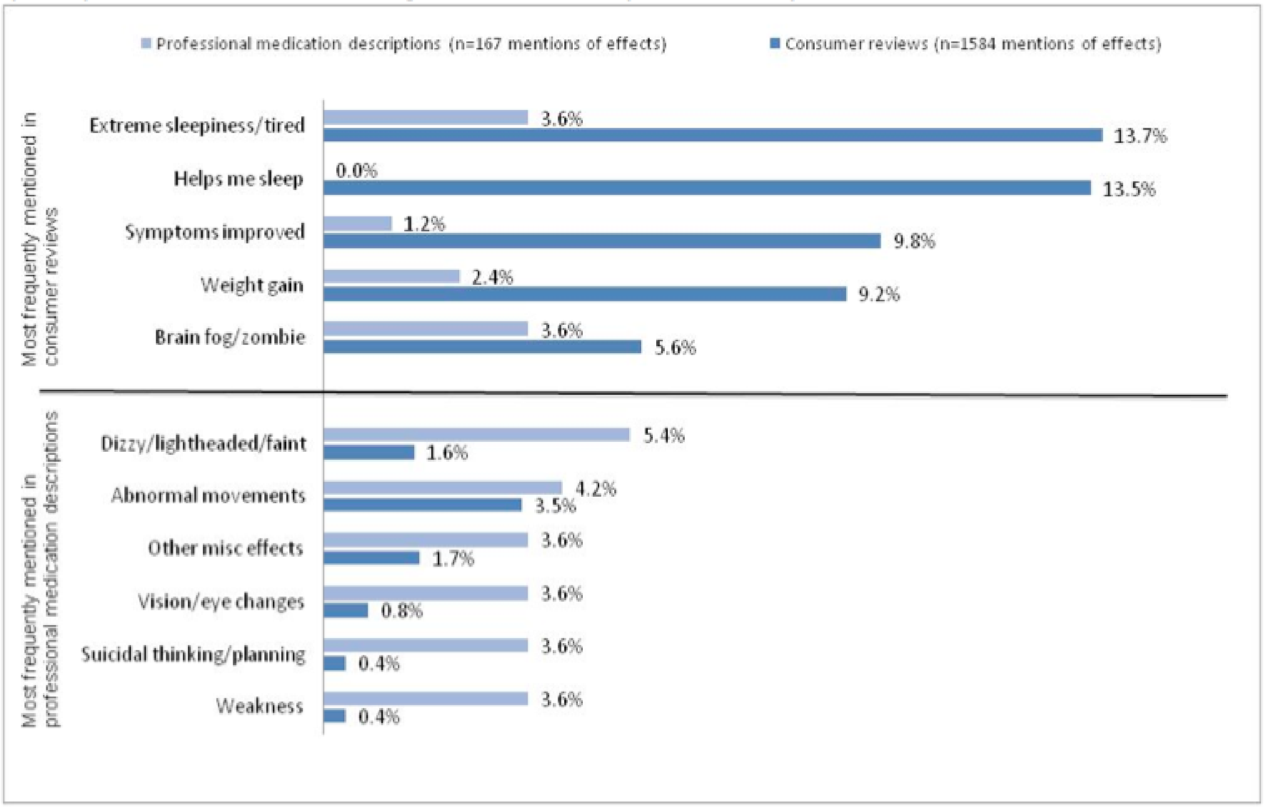


Figure 3. Most frequently mentioned effects as a proportion of all mentions of quetiapine effects in consumer-generated and in professionally controlled text



The following three tables illustrate qualitative differences between consumer-generated and professionally controlled text. Table 5 compares website text with respect to worsening symptoms while taking escitalopram. Listed are the standard warnings in professional medication descriptions for worsened symptoms and suicidality while taking antidepressants (seemingly derived from the FDA-approved drug label) in addition to consumer reviews that identify the same effects but further illustrate their various manifestations.

Table 5. Worsening symptoms on escitalopram according to consumer-generated and professionally controlled text

Website	Standard Warning[a] (Professionally Controlled Text)	Consumer Review[b] (Consumer-Generated Text)
Revolution-Health	*Call your doctor at once if you have any new or worsening symptoms such as mood or behavior changes, anxiety, panic attacks, trouble sleeping, or if you feel impulsive, irritable, agitated, hostile, aggressive, restless, hyperactive (mentally or physically), more depressed, or have thoughts about suicide or hurting yourself*	*A couple of days later I had my first manic experience which lasted about 30 minutes of complete reckless driving, I probably should have gotten arrested. And a few minutes later I came down into deep depression*. [Consumer review #258]
WebMD	*Tell the doctor immediately if you notice worsening depression/other psychiatric conditions, unusual behavior changes (including possible suicidal thoughts/attempts), or other mental/mood changes (including new/worsening anxiety, panic attacks, trouble sleeping, irritability, hostile/angry feelings, impulsive actions, severe restlessness, very rapid speech)*	*I have been very hostile and irritable on this med and my panic attacks have been coming more often and they have been much worse! I have no patience with my kids or my fiancé, or basically anyone around me.* [Consumer review #364]
AskaPatient		…[a]*nd then the worst crippling panic attacks I have ever had to date*… [Consumer review #8] *I seemed to become more aggressive and assertive. I would just speak my mind whenever I got angry, and had no fear. I seemed to become more "mean" and "mad" and I just didn't like myself.* [Consumer review #41]
CrazyMeds		*Had some hypomania then extreme agitation, then suicidality. The agitation was awful, felt like I was going to jump out of my skin—and my mind was racing.* [Consumer review #172] …*2 hours of alternating panic attacks/crying jags*… [Consumer review #130]

[a] Complete text is provided.

[b] Selected illustrative comments are provided

Table 6 shows that most of the mentions of sexual effects of escitalopram in professional medication descriptions were related to *other* sexual effects, such as the nondescript sexual problems and priapism. Of all mentions of sexual effects, consumers most frequently discussed lost sex drive (42.2%) and trouble achieving orgasm (37%), though the former was described in the professionally controlled text on WebMD as infrequent and was absent from RevolutionHealth. Professional descriptions used the terms *less serious*, *less severe*, or *severe* to describe sexual effects, while consumers consistently described these as "the absolute worst," or "extremely frustrating," and made comments such as, "I want to quit…so I can have a frigging orgasm" or "can't perform sexually so you get depressed and anxious."

Table 6. Sexual effects of escitalopram according to consumer-generated and professionally controlled text

	Mentions of Sexual Effects		Professional Medication Descriptions of Sexual Effects		Consumer Reviews
	Professional Descriptions	Consumer Reviews	WebMD[b]	Revolution-Health[b]	All Websites[c]
Total n	9	135			
Sexual effect	n (%)	n (%)			
Lost sex drive	2 (22.2)	57 (42.2)	Unlikely but serious	--	"very bothersome"
Trouble achieving orgasm	3 (33.3)	50 (37)	Common, less severe	Less serious	"the absolute worst" "extremely frustrating"
Other sexual effects[a]	4 (44.4)	16 (11.9)	Infrequent, less severe	--	"significant sexual effects"

[a] In professionally controlled text, this code included only the terms *priapism* and *sexual dysfunction*. In consumer reviews, this code included the terms *sexual side effects*, *sexual dysfunction*, and *sexual problems*.

[b] Complete text is provided.

[c] Illustrative comments are provided.

Table 7 illustrates notable qualitative differences also observed in the described sleep effects of quetiapine. While approximately one-third of consumers reported the drug helped them sleep, this benefit was absent from professional medication descriptions, which only mentioned drowsiness or tiredness as a "less severe" side effect of quetiapine. The typical excerpts from consumer reviews listed in Table 7 describe the sleep effect as sometimes helpful and sometimes burdensome, depending on the individual's circumstances and needs at the time.

Table 7. Sleep effects of quetiapine according to consumer-generated and professionally controlled text

Website	Sleep Effect[a] (Professionally Controlled Text)	Sleep Effect[b] (Consumer-Generated Text)
Revolution-Health	*The following warnings are available for this medication…may cause drowsiness*	*[It] puts me to sleep. It’s that simple. I take it and within an hour I’m out—unwakable—for the next 12 or more hours.* [Consumer review #739] *…helped very, very effectively with sleep: 30 minutes max after taking 125-150 mg at night, I am out for good.* [Consumer review #808] *…the worst side effect is the sleepiness—I sleep 10-12 hours a day and still have periods when I have to nap (or could fall asleep standing up).* [Consumer review #773]
WebMD	*Common side effects: drowsiness…less severe, tiredness*	*It helped me sleep very well, but I was very groggy in the morning* [Consumer review #871]
AskaPatient		*So while it does provide me sleep…it’s the kind of sleep that wouldn’t allow me to be woken, even if my house is on fire. I am not able to be woken from this coma-like sleep for hours. That scares me.* [Consumer review #515] *…extreme sleeping…* [Consumer review #570]
CrazyMeds		*I like what this drug does to me (sleepy bye bye land).* [Consumer review #172] *You’ll sleep until next Tuesday. Of course, that could be a good thing, depending on how your life is at this moment.* [Consumer review #1084]

[a] Complete text is provided.

[b] Selected illustrative comments are provided.

Representativeness of Recent Consumer Reviews

Table 8 compares comments from the most recent 20 consumer reviews from each of the 4 websites to comments in all remaining 400 consumer reviews for each drug on effects mentioned by more than 10% of consumers. For all but 2 effects (extreme sleepiness/tired for escitalopram and brain fog/zombie for quetiapine), frequencies of reported effects in recent comments were quite comparable to frequencies in all remaining reviews.

Table 8. Twenty most recent consumer reviews compared with all remaining consumer reviews for each drug on each website for effects mentioned by more than 10% of consumers

Drug effects	Escitalopram			Quetiapine		
	80 Recent Reviews	Remaining 400 Reviews	*P* Value	80 Recent Reviews	Remaining 400 Reviews	*P* Value
	n %	n %		n %	n %	
Symptoms reduced/improved	26 (32.5)	145 (30.4)	.67	23 (28.8)	96 (24.8)	.37
Symptoms new/worsened	15 (18.8)	75 (15.8)	.43	9 (11.3)	40 (10.2)	.74
Extreme sleepiness/tired	26 (32.5)	113 (23.8)	.045	26 (32.5)	133 (33.1)	.90
Weight gain	11 (13.8)	62 (13.1)	.84	16 (20)	92 (22.5)	.56
Brain fog/zombie	10 (12.5)	51 (10.8)	.58	18 (22.5)	55 (15.2)	.047

Discussion

Principal Results

Online consumer-generated and professionally controlled text bearing on the same psychotropic drugs reported many of the same drug effects but differed substantially in their descriptions and in the relative frequency of mentions of certain effects. Consumers more frequently discussed effects with an obvious manifestation and immediate impact on their daily lives, such as excessive sleeping and weight gain. Other than repetitions of regulatory warnings about serious adverse mental or mood effects (increased suicidal ideation, for example), professional medication descriptions most often mentioned physical side effects, such as dizziness and vision problems. Additionally, descriptive labels applied in professional text, such as *less serious* or *severe*, rarely matched with the perceived importance or severity of common effects according to consumers. For example, *less severe drowsiness* caused by quetiapine, as described in professional text, can translate to "coma-like sleep" or having to miss work because of the inability to stay awake, as described in consumer reviews.

Consumer reporting of medication effects also varied across health websites, with consumers posting reviews on professionally controlled health websites more often reporting greater symptom improvement, less symptom worsening, and fewer side effects. These differences may be partly explained by visual cues and normative themes present on websites that may attract drug consumers who share a particular perspective or attitude. For example, WebMD receives substantial revenue from pharmaceutical company sponsored advertisements, which may in turn attract users who hold a favorable disposition towards medication taking.

Finally, a cursory examination of only recent consumer comments on a particular medication, as might be viewed by the "typical" Internet user seeking online information from consumer-generated text, reflected commonly reported drug effects in proportion to a full representative sample of consumer reviews.

Overall, consumer-generated and professionally controlled medication descriptions each offer distinct advantages and disadvantages in helping to make treatment decisions or gauge the predictability of one's personal medication experience. Professional medication descriptions on commercial health portals provide succinct and comprehensive summaries of possible effects, but the meaningfulness of this information is limited by the lack of context. Consumer reviews, on the other hand, provide abundant concrete descriptions and situational examples of how specific effects may manifest in various combinations and to varying degrees. While the lack of organization of consumer reviews—which are individually dispersed across many websites and sometimes quite numerous on a single website—limits their integration into coherent wholes for consumers and clinicians consulting them to aid treatment decisions, this research provides initial empirical evidence for the representativeness and usability of a typical brief browsing strategy involving consumer reviews.

Nevertheless, unless the online health searcher who uses consumer reviews actively seeks a variety of sources to retrieve consumer reviews, differences in reporting across websites—such as those observed in this study where professionally controlled websites contained more positive consumer comments and consumer-generated websites contained more negative comments—could unknowingly hinder informed decision-making. At the same time, if professional medication descriptions could more richly describe the range or impact of drug effects in ordinary situations and contexts, then online consumer reviews might not constitute such a necessary innovation for the many active and potential drug consumers who consult them. In the current environment, clinicians and consumers seeking medication information on the Internet may want to be open to consulting consumer-generated content but vigilant when reviewing it and are encouraged to maximize their exposure to a variety of drug accounts by utilizing a diversity of online consumer-generated and professionally controlled sources.

Limitations

A major limitation to all research relying on Internet data is the inherent anonymity of online users. While the accuracy of self-report data is naturally a concern in all research designs, the anonymity of Internet users adds the possibility of data contributions from persons with vested interests. Pharmaceutical industry literature, for example, has expressed a clear interest in utilizing online patient communities to build brand trust [34]. No method exists to distinguish genuine from possibly unauthentic consumer accounts, and few studies have attempted to address this problem [35]. Despite unknown authenticity and credibility, however, consumer-generated health content is quickly gaining popularity and carries utility for its users, making its description an important initial step for continued research. The present study further found that consumer-generated data does correspond with professional medication descriptions, which may add validity to these anonymous consumer Internet postings.

Also, the present study did not explore differences in drug effects according to diagnosis, reason for use, or indication, partly due to inconsistent reporting of this information by online consumers. When this information was provided, it was further difficult to parcel out diagnosis (ie, bipolar disorder) from individuals' stated reason for using a drug (ie, to help with sleep). With large proportions of consumers reporting, for example, sleep changes on quetiapine, it appears that some effects are experienced globally regardless of diagnosis or indication [36]. Further, most consumers (64%) did not report the dose of the drug they were taking, and many who did described trying multiple doses, which made it difficult to isolate any dose-effect relationships for the purposes of this analysis. Finally, while data collection strategies aimed to capture information on the immediate-release, brand-name versions of the two selected drugs, consumers may not have made the distinction in their reviews between brand name versus generic or immediate versus extended release. It is, therefore, possible that some consumer reviews described experiences of different versions of the selected medications.

Despite these limitations, this research used a mixed qualitative and quantitative analysis of a large representative sample of Internet data from a purposively varied selection of websites. All textual data thus obtained were submitted to coding. Three strategies to minimize interpretive biases in qualitative coding methods were used: the research grounded codes in the data by preserving consumers' language in developing code names and categories, maintained utmost transparency by using tracking features in QDA Miner 3.2 software, and tested for the reliability of assigned codes by measuring agreement with a second independent coder.

Comparison With Prior Work

Notable similarities and differences exist between consumer-reported effects in this sample and other estimates of drug effects. An online service that collects drug safety information from its patient community, iGuard.org, surveyed a random sample of 700 members taking 1 of 5 antidepressants, including escitalopram [37]. Congruent with the present findings, the most frequently reported side effects were sexual dysfunction (24.5%), sleepiness (23.1%), and weight gain (21.4%). The FDA-approved label for escitalopram lists lower rates of these effects, reporting that 1% to 7% of clinical trial participants with major depressive disorder and anxiety experienced decreased libido or impotence, and 6% to 13% experienced somnolence, while no clinically important changes in body weight were observed. Postmarketing studies of antidepressants have estimated higher but varying rates of sexual effects, affecting 20% to 80% of users [38-40]. Research on escitalopram-induced weight gain has shown the effect to be minor [41], and data on sleep show, as the present findings, both sedative and stimulant effects [42,43].

Similarly, for quetiapine, the FDA drug label cites 4% to 22% of participants in clinical trials experiencing weight gain, an effect mentioned by 22.5% of consumers in this study. Since the release of quetiapine on the US market in 2002, weight gain and metabolic disorders have been recognized as significant problems for all atypical antipsychotics, though quetiapine is typically regarded as causing less weight gain than other medications in its class [44-46]. Reports from consumers in this study also seem to reflect real-world use of quetiapine as a sleep aid, among other frequent off-label uses [47-49].

Conclusions

If online consumer medication reviews can offer meaningful information to those contemplating or making treatment decisions, as this research suggests, then such reviews may further be useful for postmarketing safety surveillance. Current safety surveillance systems, such as the FDA's MedWatch, are known to capture only a fragment of medically defined serious adverse events. The dispersion of consumer reviews within and across websites, their lack of a standardized vocabulary for reporting drug effects, and sparse detailing of the main elements of a conventional adverse event report currently limit their practical value for surveillance. Technology to integrate and organize in a searchable format the mass of dispersed consumer medication reviews may partially address these limitations and hold the potential to be an innovative addition to a currently deficient system [50]. In the meantime, informed discussion, creative suggestions, as well as guidance from the FDA regarding the responsibility of website owners and pharmaceutical companies over monitoring and reporting adverse events found in online consumer reviews and patient communities are needed [51].

The findings of this study suggest avenues for continued research. First, the present analysis could be replicated using a standardized medical coding vocabulary (ie, MedDRA) in order to facilitate comparison with other pharmacoepidemiological databases. The present analysis could also be replicated (1) to determine if online consumers report effects in similar proportion for additional medications and websites and (2) to search for temporal trends and patterns in types of effects reported and their associations with large-scale events such as warnings from regulatory agencies or direct-to-consumer ad campaigns for medications. Secondly, it is unclear if discrepancies in drug effects between consumer reviews and information derived from conventional drug research represent an overestimation of effects by online consumers or an underestimation of effects in drug research. To address this, controlled clinical trials could incorporate simple targeted measures for weight, sleep, and sexual effects, rather than continue to rely on spontaneous or unsolicited participant self-report for such data (a method that tends to underestimate the true frequency of events) [52,53]. Lastly, this research suggests that current strategies for filtering online health searches to return only *trusted* or *approved* websites [5,6] may inappropriately address the challenge to identify quality health sources on the Internet because such strategies unduly limit access to an entire complementary source for health information.

Acknowledgments

The authors thank the JMIR reviewers who provided valuable suggestions. S Hughes received a dissertation year fellowship from Florida International University and a doctoral dissertation grant from the Fahs-Beck Fund for Research and Experimentation. D Cohen received funding from the National Institute of Mental Health. These funders had no role in the study design, writing of the report, or in the decision to submit the article for publication. S Hughes also received the JMIR Medicine 2.0 award and IMIA Medicine 2.0 award for a presentation at the 3rd Medicine 2.0 World Congress on Social Media in Health and Medicine in Maastricht, 2010, involving a portion of this manuscript. The former award provided a waiver for the article publication fee by JMIR.

Conflicts of Interest

None declared

Authors' Contributions

Both authors contributed equally to the planning and reporting of the work described in the article. S Hughes conducted the analysis of the material reviewed. Both authors had full access to all of the data and take responsibility for the integrity of the data. S Hughes is the guarantor and accepts full responsibility for the finished article and controlled the decision to publish.

Multimedia Appendix 1

Powerpoint presentation: Can online consumers contribute to drug knowledge and drug safety? An examination of consumer reporting of drug effects across health websites

[PPT file (Microsoft Powerpoint File), 1024 KB - jmir_v13i3e53_app1.ppt]

References

1. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2009 Jun. URL: http://pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2010-12-20] [WebCite Cache ID 5v7h0Skmz]
2. Kummervold P, Chronaki C, Lausen B, Prokosch H, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
3. Von Knoop C, Lovich D, Silverstein M, Tutty M. Vital signs: E-health in the United States. Boston, MA: The Boston Consulting Group; 2003. URL: http://www.bcg.com/documents/file14063.pdf [accessed 2010-12-20] [WebCite Cache ID 5v7i9vt4m]
4. Brown P. WHO calls for a health domain name to help consumers. BMJ 2002 Mar 9;324(7337):566 [FREE Full text] [Medline: 11884314]
5. Fox S. Health sites: some are more equal than others.: Pew Internet & American Life Project; 2010 Jan 21. URL: http://www.pewinternet.org/Commentary/2010/January/Health-Sites-Some-Are-More-Equal-Than-Others.aspx [accessed 2010-12-20] [WebCite Cache ID 5v7iFMYn4]
6. The Google algorithm.: The New York Times; 2010 Jul 14. URL: http://www.nytimes.com/2010/07/15/opinion/15thu3.html?_r=1 [accessed 2010-12-20] [WebCite Cache ID 5v7iJIBsz]
7. Introna L, Nissenbaum H. Defining the Web: the politics of search engines. Computer 2000;33(1):54-62.
8. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 12020305]
9. Lewis T. Seeking health information on the Internet: lifestyle choice or bad attack of cyberchondria. Media, Culture & Society 2006;28(4):521-539.
10. Speight J, Barendse SM. FDA guidance on patient reported outcomes. BMJ 2010;340:c2921. [Medline: 20566597]
11. Wise J. NHS publishes new data on patient reported health outcomes. BMJ 2010;341:c5143. [doi: 10.1136/bmj.c5143]
12. Bhinder S, Chowdhury N, Granton J, Krahn M, Tullis D, Waddell T, et al. Feasibility of internet-based health-related quality of life data collection in a large patient cohort. J Med Internet Res 2010;12(3):e35 [FREE Full text] [doi: 10.2196/jmir.1214] [Medline: 20719740]
13. Jones JB, Snyder CF, Wu AW. Issues in the design of Internet-based systems for collecting patient-reported outcomes. Qual Life Res 2007 Oct;16(8):1407-1417. [doi: 10.1007/s11136-007-9235-z] [Medline: 17668293]
14. Madan A, Borckardt JJ, Connell A, Book SB, Campbell S, Gwynette MF, et al. Routine assessment of patient-reported outcomes in behavioral health: room for improvement. Qual Manag Health Care 2010;19(1):70-81. [doi: 10.1097/QMH.0b013e3181ccbc53] [Medline: 20042935]
15. Barker KK. Electronic support groups, patient-consumers, and medicalization: the case of contested illness. J Health Soc Behav 2008 Mar;49(1):20-36. [Medline: 18418983]
16. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
17. Powell J, McCarthy N, Eysenbach G. Cross-sectional survey of users of Internet depression communities. BMC Psychiatry 2003 Dec 10;3:19 [FREE Full text] [doi: 10.1186/1471-244X-3-19] [Medline: 14664725]
18. Davis JJ, Cross E, Crowley J. Pharmaceutical websites and the communication of risk information. J Health Commun 2007;12(1):29-39. [doi: 10.1080/10810730601091326] [Medline: 17365347]
19. Macias W, Lewis L. A content analysis of direct-to-consumer prescription drug websites. J Advert 2003;32(4):43-56.
20. Waack KE, Ernst ME, Graber MA. Informational content of official pharmaceutical industry web sites about treatments for erectile dysfunction. Ann Pharmacother 2004 Dec;38(12):2029-2034. [doi: 10.1345/aph.1E251] [Medline: 15507499]
21. Schröder S, Zöllner YF, Schaefer M. Drug related problems with Antiparkinsonian agents: consumer Internet reports versus published data. Pharmacoepidemiol Drug Saf 2007 Oct;16(10):1161-1166. [doi: 10.1002/pds.1415] [Medline: 17486665]
22. Moncrieff J, Cohen D, Mason JP. The subjective experience of taking antipsychotic medication: a content analysis of Internet data. Acta Psychiatr Scand 2009 Aug;120(2):102-111. [doi: 10.1111/j.1600-0447.2009.01356.x] [Medline: 19222405]
23. Wicks P, MacPhee GJ. Pathological gambling amongst Parkinson's disease and ALS patients in an online community (PatientsLikeMe.com). Mov Disord 2009 May 15;24(7):1085-1088. [doi: 10.1002/mds.22528] [Medline: 19353722]
24. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]
25. The Association of the British Pharmaceutical Industry. 2007. Facts and statistics from the pharmaceutical industry, top pharmaceutical products in the UK URL: http://www.abpi.org.uk/statistics/section.asp?sect=2 [accessed 2010-12-20] [WebCite Cache ID 5v7lKHP85]
26. IMS Health. 2008. Top 15 global products URL: http://www.imshealth.com/deployedfiles/imshealth/Global/Content/StaticFile/Top_Line_Data/Global_Top_15_Products.pdf [accessed 2010-12-20] [WebCite Cache ID 5v7lPD3Qm]

27. IMS Health. 2008. Top 15 products by U.S. sales URL: http://www.imshealth.com/deployedfiles/imshealth/Global/Content/StaticFile/Top_Line_Data/2008_Top_15_Products_by_U.S._Sales.pdf [accessed 2010-12-20] [WebCite Cache ID 5v7lSw8sr]
28. Drugs.com. 2010. Top 200 drugs for 2009 by sales URL: http://www.drugs.com/top200.html [accessed 2011-04-05] [WebCite Cache ID 5xj9CVw2h]
29. Stahl S. Stahl's Essential Psychopharmacology: The Prescriber's Guide. Cambridge, England: Cambridge University Press; 2009.
30. Peladeau, N. QDA Miner Version 3.2 [computer program]. Montreal: Provalis Research; 2009.
31. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: 11701577]
32. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 3rd edition. Thousand Oaks, CA: Sage Publications, Inc; 2008.
33. Krippendorf K. Content Analysis: An Introduction to Its Methodology. 2nd edition. Thousand Oaks, CA: Sage Publications; 2004.
34. O'Neill A. DTC Perspectives. 2007. The patient trust deficit in pharmaceutical marketing URL: http://www.dtcperspectives.com/content/editor/files/The%20Patient%20Trust%20Deficit%20in%20Pharmaceutical%20Marketing.pdf [accessed 2011-04-05] [WebCite Cache ID 5xj9OXPAz]
35. David S, Pinch TJ. Six degrees of reputation: the use and abuse of online review and recommendation systems.: Social Science Research Network; 2005. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=857505 [accessed 2010-12-20] [WebCite Cache ID 5v7laTWh1]
36. Moncrieff J, Cohen D. How do psychiatric drugs work? BMJ 2009;338:b1963. [Medline: 19482870]
37. Cascade E, Kalali AH, Kennedy SH. Real-World Data on SSRI Antidepressant Side Effects. Psychiatry (Edgmont) 2009 Feb;6(2):16-18. [Medline: 19724743]
38. Landén M, Högberg P, Thase ME. Incidence of sexual side effects in refractory depression during treatment with citalopram or paroxetine. J Clin Psychiatry 2005 Jan;66(1):100-106. [Medline: 15669895]
39. Lee KU, Lee YM, Nam JM, Lee HK, Kweon YS, Lee CT, et al. Antidepressant-Induced Sexual Dysfunction among Newer Antidepressants in a Naturalistic Setting. Psychiatry Investig 2010 Mar;7(1):55-59. [doi: 10.4306/pi.2010.7.1.55] [Medline: 20396434]
40. Serretti A, Chiesa A. Treatment-emergent sexual dysfunction related to antidepressants: a meta-analysis. J Clin Psychopharmacol 2009 Jun;29(3):259-266. [doi: 10.1097/JCP.0b013e3181a5233f] [Medline: 19440080]
41. Uher R, Mors O, Hauser J, Rietschel M, Maier W, Kozel D, et al. Changes in body weight during pharmacological treatment of depression. Int J Neuropsychopharmacol 2011 Apr;14(3):367-375. [doi: 10.1017/S1461145710000933] [Medline: 20716398]
42. Mayers A, Baldwin D. Antidepressants and their effect on sleep. Hum Psychopharmacol 2005 Dec;20(8):533-559. [doi: 10.1002/hup.726] [Medline: 16229049]
43. Wilson S, Argyropoulos S. Antidepressants and sleep: a qualitative review of the literature. Drugs 2005;65(7):927-947. [Medline: 15892588]
44. Komossa K, Rummel-Kluge C, Schmid F, Hunger H, Schwarz S, Srisurapanont M, et al. Quetiapine versus other atypical antipsychotics for schizophrenia. Cochrane Database Syst Rev 2010(1):CD006625. [doi: 10.1002/14651858.CD006625.pub2] [Medline: 20091600]
45. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 2005 Sep 22;353(12):1209-1223 [FREE Full text] [doi: 10.1056/NEJMoa051688] [Medline: 16172203]
46. Rummel-Kluge C, Komossa K, Schwarz S, Hunger H, Schmid F, Lobos CA, et al. Head-to-head comparisons of metabolic side effects of second generation antipsychotics in the treatment of schizophrenia: a systematic review and meta-analysis. Schizophr Res 2010 Nov;123(2-3):225-233. [doi: 10.1016/j.schres.2010.07.012] [Medline: 20692814]
47. Leslie DL, Mohamed S, Rosenheck RA. Off-label use of antipsychotic medications in the department of Veterans Affairs health care system. Psychiatr Serv 2009 Sep;60(9):1175-1181 [FREE Full text] [doi: 10.1176/appi.ps.60.9.1175] [Medline: 19723731]
48. Philip NS, Mello K, Carpenter LL, Tyrka AR, Price LH. Patterns of quetiapine use in psychiatric inpatients: an examination of off-label use. Ann Clin Psychiatry 2008;20(1):15-20. [doi: 10.1080/10401230701866870] [Medline: 18297582]
49. Stowell KR, Ghinassi FA, Fabian TJ, Nash KC, Haskett RF. Best practices: an intervention to promote evidence-based prescribing at a large psychiatric hospital. Psychiatr Serv 2009 Mar;60(3):294-296 [FREE Full text] [doi: 10.1176/appi.ps.60.3.294] [Medline: 19252040]
50. Strom BL. How the US drug safety system should be changed. JAMA 2006 May 3;295(17):2072-2075. [doi: 10.1001/jama.295.17.2072] [Medline: 16670415]
51. Davies M. Listening to consumers in a highly regulated environment: How pharmaceutical manufacturers can leverage consumer-generated media.: The Nielsen Company; 2008. URL: http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/11/Nielsen-Online-Healthcare-Practice_Social-Media-Adverse-Event-Reporting_nov09.pdf [accessed 2011-04-05] [WebCite Cache ID 5xj9dEyD3]

52. Hughes S, Cohen D. Understanding the assessment of psychotropic drug harms in clinical trials to improve social workers' role in medication monitoring. Soc Work 2010 Apr;55(2):105-115. [Medline: 20408352]
53. Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. J Nerv Ment Dis 2002 Sep;190(9):583-592. [doi: 10.1097/01.NMD.0000030522.74800.0D] [Medline: 12357091]

Abbreviations

FDA: Food and Drug Administration

Edited by G Eysenbach; submitted 20.12.10; peer-reviewed by F Pestello, L Fernandez-Luque, P Wicks; comments to author 10.03.11; revised version received 06.04.11; accepted 04.05.11; published 29.07.11.

Please cite as:
Hughes S, Cohen D
Can Online Consumers Contribute to Drug Knowledge? A Mixed-Methods Comparison of Consumer-Generated and Professionally Controlled Psychotropic Medication Information on the Internet
J Med Internet Res 2011;13(3):e53
URL: http://www.jmir.org/2011/3/e53/
doi:10.2196/jmir.1716
PMID:21807607

Original Paper

Prevalence and Global Health Implications of Social Media in Direct-to-Consumer Drug Advertising

Bryan A Liang[1,2], MD, PhD, JD; Timothy K Mackey[2,3], MAS

[1]Department of Anesthesiology, San Diego Center for Patient Safety, San Diego School of Medicine, University of California, San Diego, CA, United States

[2]Institute of Health Law Studies, California Western School of Law, San Diego, CA, United States

[3]Joint Doctoral Program in Global Health, San Diego State University-University of California, San Diego, CA, United States

Corresponding Author:
Timothy K Mackey, MAS
Institute of Health Law Studies
California Western School of Law
350 Cedar Street
San Diego, CA, 92101
United States
Phone: 1 619 515 1568
Fax: 1 619 515 1599
Email: tmackey@ucsd.edu

Abstract

Background: Direct-to-consumer advertising (DTCA), linked to inappropriate medication use and higher health care expenditures, is the fastest growing form of pharmaceutical marketing. DTCA is legal only in the United States and New Zealand. However, the advent of online interactive social media "Web 2.0" technologies—that is, eDTCA 2.0—may circumvent DTCA legal proscriptions.

Objective: The purpose of this study was to assess the prevalence of DTCA of leading pharmaceutical company presence and drug product marketing in online interactive social media technologies (eDTCA 2.0).

Methods: We conducted a descriptive study of the prevalence of eDTCA 2.0 marketing in the top 10 global pharmaceutical corporations and 10 highest grossing drugs of 2009.

Results: All pharmaceutical companies reviewed (10/10, 100%) have a presence in eDTCA 2.0 on Facebook, Twitter/Friendster, sponsored blogs, and really simple syndication (RSS) feeds. In addition, 80% (8/10) have dedicated YouTube channels, and 80% (8/10) developed health care communication-related mobile applications. For reviewed drugs, 90% (9/10) have dedicated websites, 70% (7/10) have dedicated Facebook pages, 90% (9/10) have health communications-related Twitter and Friendster traffic, and 80% (8/10) have DTCA television advertisements on YouTube. We also found 90% (9/10) of these drugs had a *non*-corporate eDTCA 2.0 marketing presence by illegal online drug sellers.

Conclusion: Pharmaceutical companies use eDTCA 2.0 to market themselves and their top-selling drugs. eDTCA 2.0 is also used by illicit online drug sellers. Regulators worldwide must take into account the current eDTCA 2.0 presence when attempting to reach policy and safety goals.

(J Med Internet Res 2011;13(3):e64) doi:10.2196/jmir.1775

KEYWORDS

Illegal pharmacies; social media; pharmaceutical marketing; direct-to-consumer-advertising; internet pharmacies; global health; law; health policy

Introduction

Pharmaceutical direct-to-consumer advertising (DTCA) is legal only in the United States and New Zealand among industrialized countries [1]. It is linked with inappropriate medication use, overutilization, and increased spending on expensive branded drugs, and it may endanger public health due to promotion of potentially dangerous products [2]. International DTCA legal proscriptions indicate that sovereign states have deemed DTCA risks to outweigh its benefits [3]. Yet it is the fastest growing form of marketing, rising 330% from 1996 to 2005 [2], with about US $4.3 billion 2009 expenditures [4], outpacing physician marketing and research and development spending over the past decade [5].

With the Internet's rapid development, users have migrated from passive information sources, using read-only "Web 1.0" technology, to interactive, dynamic, and custom-built relationships, using "Web 2.0" technologies [6]. Along with this digital revolution, new potential DTCA marketing opportunities haven recently emerged that include Web 2.0 social networking sites and other interactive systems ("eDTCA 2.0"), which cross geopolitical borders [7]. The US Food and Drug Administration (FDA) has not issued guidelines on eDTCA 2.0 marketing, nor have regulators recognized eDTCA 2.0 and its potential global spillover.

With DTCA public health concerns, lack of specific US regulation, global DTCA legal prohibition, and the Internet's extensive reach via social networking and other Web 2.0 technologies, we assessed global pharmaceutical company eDTCA 2.0 presence for potential marketing. Here, the focus was the corporate presence in eDTCA 2.0 social media for potential marketing, rather than the accuracy of DTCA claims, which has been found problematic elsewhere [5,8]. In addition, we wished to assess whether top-selling drugs have an established eDTCA 2.0 presence, potentially avoiding global DTCA prohibitions.

Methods

We identified the 10 largest pharmaceutical companies and the top 10 grossing medicines worldwide in 2009 using IMS Health sales data, which tracks revenue by company and product [9].

Google searches were run to identify their eDTCA 2.0 presence. Important to note is that eDTCA 2.0 presence contains a mix of DTCA categories, including reminder advertisements that only contain information about a disease or condition and recommend health care provider consultation; help-seeking advertisements that include product name and therapeutic claims; and product-claim advertisements that only provide the product name [10]. The prevalence of social media marketing tools used by pharmaceutical corporations was determined by searching for dedicated corporate social media sites, Facebook pages, Twitter or Friendster accounts, blogs or really simple syndication (RSS) feeds, and YouTube channels, and whether corporations had developed and sold health-related mobile applications in the Apple iTunes store.

Dedicated corporate social media sites were defined as manufacturer-hosted websites launching multiple company social media marketing tools, including Facebook and Twitter or Friendster. Facebook is the leading social network interactive service connecting 500 million users [11]. Twitter, a popular US-dominant social media site with 106 million users [12], and Friendster, with a strong international presence, particularly in Asia, with 100 million users [13], are social networking and microblog sites using short Internet posts. Corporate blogs and RSS feeds are company-sponsored Web feed formats providing communication to users. Corporate YouTube channels were defined as company-sponsored sites with dedicated channels for its marketing videos. Corporate mobile application presence was defined as mobile applications identified by pharmaceutical company name and copyright for smartphones and other mobile technologies in the iTunes store. Foreign corporate subsidiaries and affiliates were also included in these searches. A diverse mix of eDTCA 2.0 categories occurs in this space, but the predominant category of DTCA is help-seeking advertisements.

In assessing eDTCA 2.0 prevalence for drugs, searches were conducted to determine whether each specific product had a dedicated website, product-specific Facebook pages, user-generated Twitter/Friendster postings, and YouTube-available DTCA. Dedicated product pages were identified as manufacturer websites that solely marketed the product. A corporate Facebook page where the profile was identified by product name and description was defined as a product-specific Facebook page. In determining whether products were the subject of Twitter/Friendster postings, all postings uploaded by all users that discussed health-related communications were reviewed for drug-specific discussions. YouTube-available DTCA was identified as any generated viewable video content upload of video-broadcasted DTCA relating to the specific drug. eDTCA 2.0 for drug products also includes a mix of DTCA types; however, the predominant category in this space is product-claim advertisements, which identify the product and the therapeutic claim along with safety and efficacy information.

Searches were conducted from September 2010 to December 2010 using the Google search engine. The search strategy for identifying dedicated social media marketing tools used by pharmaceutical companies included the following key word searches. For dedicated manufacturer-hosted social media sites the identified corporation name and the term "social media" were used for the search (eg, "Pfizer+social media"). For specific social media tools used by pharmaceutical companies, we used social media tool links from the identified manufacturer's hosted social media website results and key word searches consisting of the identified corporation name and the social media tool type (eg, "Pfizer+Facebook"). In assessing eDTCA 2.0 presence for drug products, searches for dedicated product websites used the identified drug name and the key word "official website" (eg, "Lipitor+official website"). For product-specific social media tools we used the identified drug name and the social media tool type (eg, "Lipitor+Facebook").

Results

Table 1 summarizes the eDTCA 2.0 presence of the top 10 pharmaceutical corporations. Of these, 40% (4/10) had dedicated social media corporate websites linking to all the company's other social media marketing tools. The world's largest pharmaceutical corporation, Pfizer, had the most robust social media website, including links to YouTube, Facebook, and Twitter, as well as SlideShare, LinkedIn, Flickr, and blog resources. All (10/10, 100%) corporations had corporate Facebook pages, Twitter feeds, and some type of sponsored blogs or RSS feeds, while 80% (8/10) had dedicated YouTube channels and health care-related mobile applications.

Table 1. Top 10 pharmaceutical companies by sales and eDTCA 2.0[a] presence

Company	Dedicated social media site	Facebook page	Twitter/Friendster	Sponsored blogs/RSS[b]	YouTube channel	Mobile applications
Pfizer	Yes	Yes	Yes	Yes	Yes	Yes
Merck & Co.	No	Yes	Yes	Yes	No	Yes
Novartis	No	Yes	Yes	Yes	Yes	Yes
Sanofi-Aventis	Yes	Yes	Yes	Yes	Yes	Yes
Glaxo-SmithKline	No	Yes	Yes	Yes	Yes	Yes
AstraZeneca	Yes	Yes	Yes	Yes	Yes	Yes
Roche	Yes	Yes	Yes	Yes	Yes	Yes
Johnson & Johnson	No	Yes	Yes	Yes	Yes	Yes
Eli Lilly	No	Yes	Yes	Yes	No	No
Abbott	No	Yes	Yes	Yes	Yes	No

[a] Direct-to-consumer advertising developed for interactive social media “Web 2.0” technologies.

[b] Really simple syndication.

Table 2 summarizes eDTCA 2.0 for 10 blockbuster branded drugs. Of these, 90% (9/10) have dedicated product pages; 70% (7/10) have product-specific dedicated Facebook pages; 90% (9/10) have health communications-related Twitter/Friendster traffic; and 80% (8/10) have DTCA advertisements on YouTube.

Table 2. Top 10 grossing drugs and eDTCA 2.0[a] presence

Drug	Dedicated product page	Facebook page	Twitter/Friendster	YouTube	eDTCA 2.0 online pharmacy link
Lipitor	Yes	Yes	Yes	Yes	Yes
Plavix	Yes	No	Yes	Yes	Yes
Nexium	Yes	Yes	Yes	Yes	Yes
Seretide	Yes	No	Yes	Yes	Yes
Seroquel	Yes	Yes	Yes	Yes	Yes
Enbrel	Yes	Yes	Yes	Yes	No
Remicade	Yes	Yes	Yes	No	No
Crestor	Yes	Yes	Yes	Yes	Yes
Zyprexa	No (taken down)	No	Yes	No	Yes
Humira	Yes	Yes	No	Yes	No

[a] Direct-to-consumer advertising developed for interactive social media “Web 2.0” technologies.

During searches for drug-specific, corporate eDTCA 2.0 presence, we also observed an unexpected finding: use of eDTCA 2.0 pharmaceutical marketing by illicit online pharmacies. Illicit online pharmacies are websites or links to websites identified as marketing the sale of drug products without a prescription [14]. In addition, identified websites and links to websites did not meet the verification criteria under the National Association of Boards of Pharmacy Verified Internet Pharmacy Practice Sites (VIPPS) program [15]. These illicit online pharmacies have historically used online marketing tools such as search engine optimization and search engine marketing tools as an illicit form of DTCA [7]. When performing these searches, we discovered that 90% (9/10) had a noncorporate eDTCA 2.0 presence, either advertising the purchasing of drugs online without a prescription or linking directly to illegal online pharmacies promoting sales without a prescription. The search was expanded to the top 20 globally marketed DTCA products. The expanded search used The Nielsen Company data to identify the top 20 drugs brands by DTCA spending [16]. The trend remained, with 80% (16/20) of the top 20 products either advertising or linking to illegal online drug sellers using eDTCA 2.0 tools (Table 3). Most included graphical advertisements and direct links to illicit sellers. The prevalence of illegal online drug sellers on Facebook alone was 60% (6/10) for the top 10 drugs and 50% (10/20) for the top 20.

Table 3. Top 20 drugs by global sales and eDTCA 2.0[a] link to illegal online drug sellers

Drug	Facebook page link	Twitter/Friendster link
Lipitor	Yes	Yes
Ablify	Yes	Yes
Advair Diskus	No	Yes
Cymbalta	Yes	Yes
Cialis	Yes	Yes
Lyrica	No	No
Plavix	Yes	Yes
Symbicort	No	Yes
Ambien CR	Yes	Yes
Crestor	No	Yes
Viagra	Yes	Yes
Pristiq	Yes	Yes
Flomax	No	Yes
Chantix	No	Yes
Yaz	Yes	Yes
Enbrel	No	No
Celebrex	Yes	Yes
Boniva	No	No
Spiriva	No	No
Caduet	Yes	Yes

[a] Direct-to-consumer advertising developed for interactive social media “Web 2.0” technologies.

Discussion

Pharmaceutical Digital Drive

The Internet’s global expanse has led to significant patient use. The National Center for Health Statistics reported 51% of adults searched for health information on the Internet from January to June 2009 [17], and a Harris poll estimated 175 million adults use the Internet for health care information [18]. With hundreds of millions of users of social media and other interactive tools [19], this market cannot be ignored.

Pharmaceutical companies appear to agree, adapting and engaging eDTCA 2.0 technologies to promote themselves and their highest grossing drugs. With near universal adoption of the most popular social media marketing represented by Facebook, Twitter/Friendster, and RSS feeds and blogs, companies are firmly committed to eDTCA 2.0. As well, 80% of firms with YouTube-dedicated channels and mobile health applications indicate they are also investing in other eDTCA 2.0 tools, such as multimedia formats with videos and health communications-related software.

However, these results indicate that eDTCA 2.0 marketing may not be limited to where DTCA is permitted by law. eDTCA 2.0 sites such as GlaxoSmithKline’s blog site and AstraZeneca’s community Facebook page indicate they are “intended for US residents/customers only,” but appear to offer no access restrictions to non-US users [7]. Further, mobile applications such as those advertised on Apple’s iTunes media store are marketed by large pharmaceutical firms such as Pfizer, Novartis, and Roche, and target Canadian, French, and Korean audiences. Hence, eDTCA 2.0 marketing may occur outside the United States and New Zealand [7].

Estimates of 2009 total online DTCA spending (including eDTCA 2.0) are between US $117 million and $1 billion. Although these figures appear lower than traditional DTCA expenditures, these figures may both underestimate [16,20] and not reflect Internet advertising’s significantly lower costs and greater benefits in reaching advertising markets [21,22]. These benefits include flexibility in marketing blockbuster and niche therapies, ability to reach larger audiences and target specific patients, and better financial analytics of social media and marketing return on investment [23]. Further, the scope of eDTCA 2.0 efforts is still not well defined. For example, some estimates include banners, advertisement displays, and advertisement time on streaming videos [24] but fail to capture eDTCA 2.0 presence in sponsored links, advertisement works, website development and hosting, and other interactive social media tools that have proliferated for promoting drugs [16].

Regulatory Challenges

For public health policymakers and patients, the Internet’s escalating use is of concern. First, pharmaceutical manufacturers’ eDTCA 2.0 development and presence may set back even further FDA efforts to effectively regulate DTCA

[8]. Internationally, global DTCA prohibitions are being emasculated by these promotion efforts. eDTCA 2.0 technology has moved DTCA outside geopolitical regulatory boundaries into a system of potentially unfettered promotion, openly accessible to manufacturers and patients worldwide. Indeed, eDTCA 2.0 has spread DTCA to anyone, anywhere with an Internet connection.

This may serve as a stark warning for public health policy, given earlier work that concluded that prohibited television-based US DTCA transmitted to Canada increased prescriptions for tegaserod by 42% after it began—a drug that was later withdrawn from US and Canadian markets due to safety concerns [25]. These developments are also of public health concern because reform efforts in the United States and abroad have not recognized issues specifically relating to eDTCA 2.0 [5,8,26]. This places these regions at risk for higher pharmaceutical prices [27] and increased costs associated with excessive prescriptions [28].

In addition, assessments of online drug advertisements show suspect quality and overemphasis of potential benefits [8]. Second, the most-advertised drugs have been those with large patient markets and/or that treat chronic or long-term conditions, placing many, particularly vulnerable, patients at risk [24]. But further, this risk may be disproportionately increased because these newer drugs are most heavily marketed early in product life cycles—when they may not be adequately assessed for safety [29-31]. Indeed, in the United States, of the top 20 DTCA-advertised drugs, 90% (18/20) were subject to a black box warning, recall, or other safety notification (Table 4). The United States's adverse experience relating to drug promotion may become a global health concern under eDTCA 2.0, particularly in developing countries with growing chronic disease burdens [32].

Table 4. Prevalence of safety issues for top 20 products advertised direct to consumers, 2009

Drug	Safety warning[a]
Lipitor	Subject to recall; other safety notification
Ablify	Black box warning; subject to recall
Advair Diskus	Black box warning
Cymbalta	Black box warning
Cialis	Other safety notification
Lyrica	None
Plavix	Black box warning; subject to recall; other safety notification
Symbicort	Black box warning; other safety notification
Ambien CR	Other safety notification
Crestor	Other safety notification
Viagra	Other safety notification
Pristiq	Black box warning; other safety notification
Flomax	Other safety notification
Chantix	Black box warning
Yaz	Black box warning; other safety notification
Enbrel	Black box warning; subject to recall; other safety notification
Celebrex	Black box warning
Boniva	Other safety notification
Spiriva	Subject to recall
Caduet	None

[a] "Other safety warnings" include voluntary recalls, counterfeit warnings, warnings on contraindications, warnings about combining with other drugs, and other US Food and Drug Administration-issued alerts.

Rogue eDTCA 2.0

Adding even more public health risks and worries is the discovery that rogue online pharmacies are already entrenched in eDTCA 2.0. The presence of illegal online pharmacies using interactive social media to sell illicit products is of great worry [7]. First, from a legal point of view, this eDTCA 2.0 presence represents another way it can undermine legal proscriptions (eg, prohibition of medication sales without a prescription, license, or oversight). But more important, avoidance of such legal proscriptions goes deeper as a safety risk, as patients may purchase drugs not approved or that are the subject of safety concerns (eg, Zyprexa).

Indeed, illegal online pharmacies have succeeded in selling tainted or fake drugs globally [14]. They have used search engine-sponsored links to sell illegally and circumvent search engine mandates for legitimacy verification [14]. But these illicit sellers are highly nimble; after investigations by us and others [14,32] showing lack of search engine oversight and continued illicit sales, Google, followed by Yahoo and Bing/Microsoft, adopted recommendations to use only VIPPS accreditation [33,34]. However, after this change, although eliminating sponsored link presence, illicit online pharmacies deftly moved to be listed in search engine shopping pages. Only after notification of this development by one of us (Andrew Kline, Senior Advisor, Intellectual Property Enforcement, Executive Office of the President, oral communication, October 8, 2010) to a government representative was this loophole closed. However, it now appears that illicit online drug sellers have infiltrated eDTCA 2.0, including Facebook with its hundreds of millions of users worldwide.

Such illicit presence and online purchasing of drugs are rife with patient risks. Counterfeit, diverted, and unregulated drugs are sold by these illegal vendors causing patient harm [14]. With drug supply globalization, risks associated with unregulated, non-VIPPS sellers are legion, with a host of public health and regulatory agencies warning about online drug sellers [14,35,36,37,38].

Global Health Problem

The combination of eDTCA 2.0 presence by drug companies and by illicit drug sellers creates even greater patient risks than each alone, and a more urgent imperative for public action [7]. Global eDTCA 2.0 from whatever source, licit or illicit, may inappropriately increase demand, and then compound harm by permitting self-prescribing and purchasing from online drug sellers directly linked from sites such as Facebook. Hence, patient exposure to potentially risky drugs or drug forms is fueled by both unregulated eDTCA 2.0 corporate marketing and illicit drug sellers populating social media.

Regulatory priority for eDTCA 2.0 is needed for other reasons. The relatively overstated benefits associated with traditional DTCA [39,40] may drive a corporate shift to more efficient eDTCA 2.0 marketing. This is especially relevant given studies showing that traditional forms of DTCA such as television, magazines, and the radio may have only a minor impact on sales of DTCA drugs [40]. Governments hence, may wish to anticipate this potential upswing by responding with aggressive regulation. Theoretically, this would also capture illicit online drug sellers, although the ease by which they can recreate their presence thwarts any single solution to addressing this concern [14]. Further, because search engines drive consumers to content [6], eDTCA 2.0 interactive media forms receive higher traffic due to repeat use or links from other websites. They will therefore often appear higher in search results and may have a disproportionate patient impact [41]. eDTCA 2.0 may consequently result in pharmaceutical companies becoming the dominant source of health-related information, replacing health care professionals [39,42].

Reform Considerations

Given limited recognition and inadequate regulation of eDTCA 2.0, and difficulty instituting a global ban due to US legal considerations [43], alternative approaches are warranted. We believe solutions must at least include eDTCA export prevention, eDTCA funding-source transparency, and patient safety integration into eDTCA communications.

In preventing eDTCA 2.0 illegal export, similar to disease outbreak surveillance and control, the United States should take an active role in ensuring that eDTCA 2.0 originating from US-based companies or information technology infrastructure is not transmitted across its borders [7]. Given that eDTCA is a public health concern, The United States, as the largest producer and eDTCA exporter, needs effective policy to proactively regulate DTCA as a global export.

To do this, the United States should require marketers, pharmaceutical manufacturers, Web content providers (such as Internet service providers and registrars), and social media sites to engage in active surveillance and block foreign internet protocol (IP) address holders or users from viewing eDTCA 2.0 content. Such efforts would be analogous to prior activities between a joint task force between the FDA and Federal Trade Commission to combat illegal Web activity involving the sale of fraudulent H1N1 supplements [44]. Enforcement would specifically include social media sites such as Facebook and Twitter. Further, illicit online pharmacy presence in the most popular social media forms poses a serious public health risk to both individuals searching for health-related information and users increasingly dependent on social networks for information. Social media service providers must recognize the serious implications of not actively policing this illicit content. Public–private partnerships in developing filtering software can be part of this effort, which can be disseminated to domestic and global social media entities and regulators [7].

Another key component is clear eDTCA 2.0 sponsorship identification. Patients must be appraised about online sources, including financial underwriting, of eDTCA 2.0 content [7]. Often reliability of eDTCA 2.0 cannot be determined because consumers cannot discern or lack information about which parties own, administer, or fund it [8]. This is especially problematic when pharmaceutical companies use third parties such as bloggers, “consumer opinion leaders,” or other employed individuals or companies [45] to promote products without disclosing financial ties [8]; moderated forums or sites that appear interactive but offer only one-sided communication; and “unbranded websites” without sponsorship disclosure [46]. The latter may garner considerable consumer attention, complement other social media, and be effective in promoting downstream revenue [46]. Such potential hidden drug marketing should be exposed, taking into account the vagaries of social network tools including short entry limitations [47] and allowing for correction of user-generated content that makes unverified, negative comments [18,48] or gives inappropriate clinical advice [49].

Lack of marketing transparency is not new, with inappropriate pharmaceutical financial funding leading to US reform [50]. Similar conflicts of interest arising in digital forums require similar disclosure and transparency. Building on the Affordable Care Act [51], policymakers can expand disclosure provisions to include funding and compensation of third parties who engage in eDTCA 2.0 on behalf of pharmaceutical manufacturers.

Disclosure would include recipient information, products promoted, payments, and identification of supported Web addresses, social media sites, and other eDTCA 2.0 presence. Arguably, industry online DTCA funding provided to health care professionals already falls under the Act and is subject to disclosure requirements. In any event, financial information should be prominently disclosed in eDTCA 2.0 media. Such information should be registered and placed in a public database so that consumers and policymakers have access to information about eDTCA 2.0 funding, which is an important first step in identifying its scope and impact. Concomitantly, this registration would also promote detection and shutdown of illegal online pharmacies in interactive social media.

There are potential benefits of eDTCA 2.0, similar to those identified for traditional DTCA, including motivating discussion between patients and providers, increasing patient education, and encouraging patients to seek treatment [52]. In addition, others include communication of patient safety and public health information, such as requiring manufacturer integration of patient safety information into eDTCA 2.0 by coordinating existing regulatory tools such as FDA risk evaluation and mitigation strategies, currently required to market certain products [53]. Drugs subject to risk evaluation and mitigation strategies can have this noted in eDTCA 2.0. As well, eDTCA 2.0 can integrate safety information from reliable sources, including the FDA’s MedWatch system, into online searches for health information and pharmaceutical products [8]. Access to such information could be via a link in eDTCA 2.0 tools. Furthermore, eDTCA 2.0 could be jointly used by drug regulatory authorities and pharmaceutical companies to warn about dangers associated with buying drugs online without a prescription. Official drug regulatory banners linking to the agency site with a list of approved, legitimate online pharmacies could be incorporated into such efforts, as well as public–private joint warnings about newly discovered safety issues associated with drugs [14].

Finally, regulatory efforts should provide specific manufacturer guidance as to eDTCA 2.0 limits. Although the FDA has not issued guidelines, it has held hearings discussing risk-versus-benefit reporting, manufacturer liability, and social media roles in DTCA, and announced plans to issue some guidance on the topic [8]. The FDA has only issued warning letters, including a dozen sent to pharmaceutical companies regarding online advertising deemed misbranded, but lacks a comprehensive policy [8]. More recently, Novartis was informed that its use of a Facebook media-sharing widget violated US drug marketing regulations [47]. eDTCA 2.0 regulations for companies should be clarified, but should also take into account patient privacy [54] and avoid the unintended advertising booms that followed earlier DTCA liberalization [19].

Study Limitation

This study has several limitations. Results are descriptive and only provide a snapshot of the current state of eDTCA 2.0. The online environment is in constant flux, and findings may not reflect changes in marketing trends.

Advertising spending is likely only a rough indicator of social media presence and DTCA use. DTCA expenditures were not stratified specifically for eDTCA 2.0, since these figures are not readily available. Further, risks associated with DTCA may not be directly analogous to eDTCA 2.0 regarding product and patient safety, and can vary based on media and content.

Further, we examined only global sales revenue of the top 10 firms and products, not regional or country variations in spending or firms outside of these high sales markets. These sales data are also difficult to validate. Examination of smaller grossing firms and drugs could show different prevalence due to marketing spending limits and differing promotion strategies. We also did not validate whether sites advertising illegal online drug sales fill orders without prescriptions, since purchasing drugs over the Internet for nonmedically appropriate reasons for a fictitious patient creates legal and ethical concerns.

As well, we examined high-volume social media sites based only on popularity, but did not validate the actual number of impressions or volume of traffic on these sites. Smaller, less-visited sites may be used more by illicit drug sellers and consumers that may lead to underreporting of prevalence. Indeed, we found many social media sites of different content and origin promoting illegal online pharmacies.

Overall, DTCA globalization through eDTCA 2.0 by both drug companies and illegal online pharmacies is a global public health challenge. Licit and illicit entities have embraced the evolution in eDTCA 2.0, leaving regulatory efforts languishing. The new consumer is one that is global and connected online, a profile that precisely fits the patient/consumer of eDTCA 2.0 [55]. Public health policy must take into account this new consumer and the rapidly developing digital environment, and act to ensure that legal proscriptions against DTCA and illegal online pharmacies are followed and patient safety is protected on a global scale.

Conflicts of Interest

None declared

References

1. World Health Organization. Direct-to-consumer advertising under fire. Bull World Health Organ 2009 Aug;87(8):576-577 [FREE Full text] [Medline: 19705005]
2. Donohue JM, Cevasco M, Rosenthal MB. A decade of direct-to-consumer advertising of prescription drugs. N Engl J Med 2007 Aug 16;357(7):673-681 [FREE Full text] [doi: 10.1056/NEJMsa070502] [Medline: 17699817]

3. Mintzes B. European Commission. 2008. Response to European Commission Public Consultation, Legal Proposal on Information to Patients URL: http://ec.europa.eu/health/files/patients/respons_publ_consult_200805/healthcare_professionals/mintzes_barbara_canada_en.pdf [accessed 2011-07-13] [WebCite Cache ID 60BCWTbaw]
4. IMS Health. IMS Integrated Promotional Services. 2009. Total US Promotional Spend by Type, 2009 URL: http://www.imshealth.com/deployedfiles/imshealth/Global/Content/StaticFile/Top_Line_Data/PromoUpdate2009.pdf [accessed 2011-07-14] [WebCite Cache ID 60BE0sgRx]
5. United States Government Accountability Office. GAO. 2006. Prescription Drugs: Improvements Needed in FDA's Oversight of Direct-to-Consumer Advertising URL: http://www.gao.gov/new.items/d0754.pdf [accessed 2011-07-14] [WebCite Cache ID 60BE3HvA6]
6. Toner R. Web 2.0: a new vocabulary for health policy. Health Policy Newsl 2008;21(2) [FREE Full text]
7. Liang BA, Mackey T. Direct-to-consumer advertising with interactive internet media: global regulation and public health issues. JAMA 2011 Feb 23;305(8):824-825. [doi: 10.1001/jama.2011.203] [Medline: 21343583]
8. Kuehn BM. FDA weighs limits for online ads. JAMA 2010 Jan 27;303(4):311-313. [doi: 10.1001/jama.2009.1961] [Medline: 20103748]
9. IMS Health. IMS Health Incorporated. 2011. Top-line Market Data: 2009 U.S. Sales and Prescription Information URL: http://www.imshealth.com/portal/site/imshealth/menuitem.a46c6d4df3db4b3d88f611019418c22a/?vgnextoid=4e260fc5b45b7210VgnVCM100000ed152ca2RCRD&cpsextcurrchannel=1 [WebCite Cache ID 5yXZJLgYW]
10. Gellad ZF, Lyles KW. Direct-to-consumer advertising of pharmaceuticals. Am J Med 2007 Jun;120(6):475-480. [doi: 10.1016/j.amjmed.2006.09.030] [Medline: 17524744]
11. Pepitone J. CNNMoney.com. 2010 Jul 21. Facebook Hits 500 Million Users URL: http://money.cnn.com/2010/07/21/technology/facebook_500_million/index.htm [accessed 2011-05-08] [WebCite Cache ID 5yXZMgeEq]
12. Website-Monitoring.com. Nick Burcher. 2010. Twitter Facts and Figures URL: http://www.nickburcher.com/2010/05/twitter-facts-and-figures-latest.html [accessed 2011-05-08] [WebCite Cache ID 5yXZPZKKL]
13. Lee M. Reuters. 2009 Dec 3. Friendster to be Sold to Asian Buyer by Month-End: Source URL: http://www.reuters.com/article/2009/12/03/friendster-idCNSHA19695420091203?rpc=44 [accessed 2011-05-08] [WebCite Cache ID 5yXZX5qEJ]
14. Liang BA, Mackey T. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(1):125-184. [Medline: 19534258]
15. National Association of Boards of Pharmacy. NABP. 2011. VIPPS URL: http://www.nabp.net/programs/accreditation/vipps/ [accessed 2011-05-08] [WebCite Cache ID 5yXe63Ecf]
16. Arnold M. Medical Marketing & Media. 2010 Mar 15. DTC Report: Flat is the New Up URL: http://www.mmm-online.com/dtc-report-flat-is-the-new-up/article/166958/ [accessed 2011-07-14] [WebCite Cache ID 60BENae5C]
17. Cohen RA, Stussman B. National Center for Health Statistics. 2010. Health Information Technology Use Among Men and Women Aged 18-64: Early Release of Estimates From the National Health Interview Survey, January-June 2009 URL: http://www.cdc.gov/nchs/data/hestat/healthinfo2009/healthinfo2009.pdf [accessed 2011-07-14] [WebCite Cache ID 60BESabLm]
18. Taylor H. Harris Polls. 2010 Aug 4. “Cyberchondriacs” on the rise? URL: http://www.harrisinteractive.com/NewsRoom/HarrisPolls/tabid/447/mid/1508/articleId/448/ctl/ReadCustom%20Default/Default.aspx [accessed 2011-07-14] [WebCite Cache ID 60BEVrXTQ]
19. Greene JA, Kesselheim AS. Pharmaceutical marketing and the new social media. N Engl J Med 2010 Nov 25;363(22):2087-2089 [FREE Full text] [doi: 10.1056/NEJMp1004986] [Medline: 21105789]
20. Hobson K. The Wall Street Journal: WSJ Blogs. 2010 Aug 26. Pharma Digital Spending Growing-But Slowly URL: http://blogs.wsj.com/health/2010/08/26/report-pharma-digital-spending-growing-but-slowly/ [accessed 2011-05-08] [WebCite Cache ID 5yXZZKyd8]
21. Angelides MC. Implementing the Internet for business: A global marketing opportunity. International Journal of Information Management 1997;17(6):405-419. [doi: 10.1016/S0268-4012(97)00024-8]
22. Copeland H, Mitzel M. New Politics Institute. 2007 Oct 22. New Tools: Advertise Online URL: http://www.newpolitics.net/node/167?full_report=1 [accessed 2011-05-08] [WebCite Cache ID 5yXZdlsef]
23. Manhattan Research. Manhattan Research, LLC. 2008. Cybercitizen Health v8. 0: Catch the New Pharmaceutical Marketing Wave: Trends and Strategies for Reaching Today's Healthcare Consumer URL: http://manhattanresearch.com/files/White_Papers/Catch_the_New_Pharmaceutical_Marketing_Wave.pdf [accessed 2011-07-14] [WebCite Cache ID 60BEnwQJe]
24. Campbell S. Congressional Budget Office. 2009. Economic and Issue Brief: Promotional Spending for Prescription Drugs URL: http://www.cbo.gov/ftpdocs/105xx/doc10522/DrugPromo_Brief.shtml [accessed 2011-07-14] [WebCite Cache ID 60BEps5Kx]
25. Law MR, Majumdar SR, Soumerai SB. Effect of illicit direct to consumer advertising on use of etanercept, mometasone, and tegaserod in Canada: controlled longitudinal study. BMJ 2008;337:a1055 [FREE Full text] [Medline: 18765444]
26. EUR-Lex. Europa Publications Office. 2008 Dec 10. Proposal for a Directive of the European Parliament and of the Council Amending, As Regards Information to the General Public on Medical Products Subject to Medical Prescription Directive

2001/83/EC on the Community Code Relating to Medicinal Products for Human Use URL: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008PC0663:EN:not [accessed 2011-05-08] [WebCite Cache ID 5yXZiUWAQ]

27. Brekke KR, Kuhn M. Direct to consumer advertising in pharmaceutical markets. J Health Econ 2006 Jan;25(1):102-130. [doi: 10.1016/j.jhealeco.2005.06.003] [Medline: 16280182]
28. Findlay S. National Institute for Health Care Management Research and Educational Foundation. 2001. Prescription Drugs and Mass Media Advertising, 2000 URL: http://www.nihcm.org/pdf/DTCbrief2001.pdf [accessed 2011-07-14] [WebCite Cache ID 60BEviliU]
29. Frosch DL, Grande D, Tarn DM, Kravitz RL. A decade of controversy: balancing policy with evidence in the regulation of prescription drug advertising. Am J Public Health 2010 Jan;100(1):24-32. [doi: 10.2105/AJPH.2008.153767] [Medline: 19910354]
30. Kolata G. The New York Times. 2010 Oct 16. When Drugs Cause Problems They Are Supposed to Prevent URL: http://www.nytimes.com/2010/10/17/health/policy/17drug.html [accessed 2011-07-14] [WebCite Cache ID 60BExTG7q]
31. Vastag B. US aims to tighten rules on direct-to-consumer drug ads. Nat Biotechnol 2007 Mar;25(3):267. [doi: 10.1038/nbt0307-267] [Medline: 17344868]
32. World Health Organization. WHO. 2010. Facts Related to Chronic Disease Fact Sheet: The Global Strategy on Diet, Physical Activity and Health URL: http://www.who.int/hpr/gs.fs.general.shtml [accessed 2011-05-08] [WebCite Cache ID 5yXavxd0V]
33. Cooperman T, Lewitt G. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(4):682-3; author reply 683. [Medline: 20196287]
34. Hogans A. CBSNews.com. 2010 Jul 28. Unverified Pharmacies Easy to Search Online URL: http://www.cbsnews.com/8301-31727_162-20011928-10391695.html [accessed 2011-05-08] [WebCite Cache ID 5yXb1fOtI]
35. Liang BA. Parallel trade in pharmaceuticals: injecting the counterfeit element into the public health. NC J Int Law Commer Regul 2006;31(4):847-899.
36. World Health Organization. WHO. 2010 Jan. Medicines: Counterfeit Medicines. Fact Sheet No URL: http://www.who.int/mediacentre/factsheets/fs275/en/ [accessed 2011-05-08] [WebCite Cache ID 5yXb4AoXv]
37. National Association of Boards of Pharmacy. NABP. 2011. Buying Medicine Online URL: http://www.nabp.net/programs/consumer-protection/buying-medicine-online/ [accessed 2011-05-08] [WebCite Cache ID 5yXb8qYPJ]
38. U.S. Food and Drug Administration. FDA. 2010. Buying Prescription Medicine Online: A Consumer Safety Guide URL: http://www.fda.gov/drugs/resourcesforyou/ucm080588.htm [accessed 2011-05-08] [WebCite Cache ID 5yXbADcm2]
39. Fogel J, Novick D. Direct-to-consumer advertisements of prescription medications over the Internet. Health Mark Q 2009;26(4):347-371. [doi: 10.1080/07359680903304310] [Medline: 19916099]
40. Cameron D. Harvard Medical School Office of Public Affairs. 2008. News Alert: Value of Direct-to-Consumer Advertising Oversold, Study Finds URL: http://hms.harvard.edu/public/news/ss0808/090108_soumerai.html [accessed 2011-05-08] [WebCite Cache ID 5yXbIqSPM]
41. Mondry B. Driving Traffic: DTC online marketing: the online content strategy imperative. DTC Perspect 2010;9(3):21-24.
42. Lorence D, Churchill R. A study of the web as DTC drug marketing agent. J Med Syst 2007 Dec;31(6):551-556. [Medline: 18041291]
43. Shuchman M. Drug risks and free speech--can Congress ban consumer drug ads? N Engl J Med 2007 May 31;356(22):2236-2239 [FREE Full text] [doi: 10.1056/NEJMp078080] [Medline: 17476002]
44. U.S. Food and Drug Administration. FDA. 2009 Oct 19. FDA, FTC Issue Joint Warning Letter to Web Site Offering Fraudulent H1N1 Flu Supplements: Agencies Continue Efforts to Protect Public Health From Illegal Web Activity URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm187142.htm [accessed 2011-05-08] [WebCite Cache ID 5yXbLzS6D]
45. Barrette J, Becker B. Yahoo Insights Library. 2007 Mar. Social Media: Pharmaceutical Marketing in the Age of Engagement URL: http://advertisingcentral.yahoo.com/insights/?category=pharmaceuticals [accessed 2011-07-14] [WebCite Cache ID 60BFF9Fw1]
46. Malloy M. What is the role of unbranded websites in pharmaceutical digital marketing? DTC Perspect 2010;9(3):27-28.
47. Jackson M. Drugmakers shy from social media sites. Los Angeles Daily Journal 2010;10.
48. Zieger A. Pharma social media marketing blunders. In: FierceBiotech editors. Social Media for Pharma Brand Managers. Washington, DC: FierceMarkets; 2010:6.
49. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292. [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
50. Liang BA, Mackey T. Confronting conflict: addressing institutional conflicts of interest in academic medical centers. Am J Law Med 2010;36(1):136-187. [Medline: 20481404]
51. US Congress. The Library of Congress Thomas. 2009. Patient Protection and Affordable Care Act of 2009, HR 3590, §6002 URL: http://thomas.loc.gov/cgi-bin/query/F?c111:1:./temp/~c111GgNLTt:e1913776: [accessed 2011-07-14] [WebCite Cache ID 60BFK7lGh]

52. Oetgen S. US Armed Forces Institute of Pathology. 2003. DTC Advertising of Prescription Drugs: How Does it Impact Patient Behavior? URL: http://www1.askafip.org/pls/askafip/docs/F1836440391/2003%20LMJ%20PDF.pdf [accessed 2011-07-14] [WebCite Cache ID 60BFRUIxa]
53. Liang BA, Mackey T. Reforming direct-to-consumer advertising. Nat Biotechnol 2011 May;29(5):397-400. [doi: 10.1038/nbt.1865] [Medline: 21552236]
54. Chester J, Mierzwinski E, Simpson JM, Dixon P. Center for Digital Democracy, U.S. PIRG, Consumer Watchdog, World Privacy Forum. 2010. Before the Federal Trade Commission In the Matter of: Online Health and Pharmaceutical Marketing that Threatens Consumer Privacy and Engages in Unfair and Deceptive Practices URL: http://www.democraticmedia.org/sites/default/files/2010-11-19-FTC-Pharma-Filing.pdf [accessed 2011-07-14] [WebCite Cache ID 60AmDAikY]
55. The Nielsen Company. Just ask Nielsen [print advertisement]. DTC Perspect 2010;9(3):12.

Abbreviations

DTCA: direct-to-consumer advertising
eDTCA 2.0: direct-to-consumer advertising developed for interactive social media “Web 2.0” technologies
FDA: US Food and Drug Administration
IP: Internet protocol
RSS: really simple syndication
VIPPS: Verified Internet Pharmacy Practice Sites

Edited by G Eysenbach; submitted 17.02.11; peer-reviewed by M Veronin; comments to author 06.05.11; revised version received 09.05.11; accepted 09.05.11; published 31.08.11.

Please cite as:
Liang BA, Mackey TK
Prevalence and Global Health Implications of Social Media in Direct-to-Consumer Drug Advertising
J Med Internet Res 2011;13(3):e64
URL: http://www.jmir.org/2011/3/e64/
doi:10.2196/jmir.1775
PMID:21880574

Original Paper

Predictors of Retention in an Online Follow-up Study of Men Who Have Sex With Men

Christine M Khosropour[1], MPH; Patrick S Sullivan[1], DVM PhD

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

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

Abstract

Background: In the past 10 years, the Internet has emerged as a venue for men who have sex with men (MSM) to meet sex partners. Because online sex seeking has increased among MSM, Internet-based human immunodeficiency virus (HIV) prevention interventions are of interest. However, few online studies to date have demonstrated an ability to retain study participants, specifically MSM of color, in longitudinal online studies.

Objective: The current analysis examines data from a 3-month online prospective study of MSM to determine the association of race and incentive level with two retention outcomes: (1) agreeing to participate in a follow-up survey and providing an email address and (2) linking into the follow-up survey at the follow-up time point.

Methods: Internet-using MSM were recruited through banner advertisements on MySpace.com. White, black, and Hispanic participants from 18 to 35 years of age were randomized to an offer of enrollment in an online follow-up survey at four levels of incentive (US $0, US $5, US $10, and US $20). Multivariable logistic regression models were used to estimate the odds of the two outcome measures of interest controlling for additional independent factors of interest.

Results: Of eligible participants, 92% (2405/2607) agreed to participate in the follow-up survey and provided an email address. Hispanic men had decreased odds (adjusted odds ratio [OR] = 0.66, 95% confidence interval [CI] 0.47-0.92) of agreeing to participate in the follow-up survey compared with white men. Men reporting unprotected anal intercourse with a male sex partner in the past 12 months had increased odds of agreeing to participate in the follow-up survey (adjusted OR = 1.42, 95% CI 1.05-1.93). Of the participants who provided an email address, 22% (539/2405) linked into the follow-up survey at the 3-month follow-up time point. The odds of linking into the follow-up survey for black men were approximately half the odds for white men (adjusted OR = 0.47, 95% CI 0.35-0.63). Participants who were offered an incentive had increased odds of linking into the follow-up survey (adjusted OR = 1.29, 95% CI 1.02-1.62). Email addresses provided by participants that were used for online financial management and email accounts that were checked daily were associated with increased odds of linking into the follow-up survey (adjusted OR = 1.97, 95% CI 1.54-2.52; adjusted OR = 1.51, 95% CI 1.22-1.87, respectively).

Conclusions: This analysis identified factors that predicted retention in an online, prospective study of MSM. Hispanic and black study participants were less likely to be retained in the study compared with white study participants. Because these men bear the greatest burden of HIV incidence among MSM in the United States, it is critical that new research methods be developed to increase retention of these groups in online research studies.

(J Med Internet Res 2011;13(3):e47) doi:10.2196/jmir.1717

KEYWORDS

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

Introduction

Since the earliest reports of AIDS in the United States, men who have sex with men (MSM) have been and continue to be the most heavily impacted risk group in the US human immunodeficiency virus (HIV) epidemic. Since 2000, MSM have been the only risk group in the United States with increasing HIV incidence [1], and from 2005 through 2007, over 250,000 MSM died with AIDS [2]. In 2008, 54% of incident HIV infections were among MSM [2], who only account for an estimated 4% of the population [3]. There are also pronounced disparities in the US MSM HIV epidemic by race and ethnicity. In 2008, approximately 22% of all new HIV diagnoses were among black MSM, and 10% were among Hispanic MSM [2]; however, during the same period, black and Hispanic men accounted for less than 6% and 8% of the US population, respectively [4]. From 2001 through 2006, new HIV diagnoses in black and Hispanic MSM each increased by 1.9% per year compared with a 0.7% annual increase for white MSM [5].

Increasing rates of HIV infection among all MSM in the past 10 years represents a resurgence of the MSM HIV epidemic, which had experienced a plateau in the late 1990s [6]. The reasons for the resurgence are multifaceted [6-14] and likely due to a combination of factors. However, one explanation for the resurgence is the use of the Internet by MSM to meet sex partners. MSM have been using the Internet to meet sex partners at increasing rates [15-20]. A number of studies have indicated that MSM who meet sex partners through the Internet not only have more sex partners than MSM who meet partners offline [21,22], but may also have higher-risk sex with their sex partners [16,18,22,23].

Since the practice of online sex seeking has increased among MSM, Internet-based HIV interventions are of great interest in the prevention community. Although few online behavioral interventions of MSM have been developed, these interventions do show promise as a method to deliver HIV prevention messages. In an uncontrolled trial in which participants viewed an online 9-minute dramatic video at baseline and returned at a 3-month time point, men were significantly more likely to both disclose their HIV status and to ask the HIV status of their partner at follow-up compared with baseline. Viewing the video was also associated with decreased odds of both engaging in unprotected anal intercourse (UAI) during the last sexual encounter and having a casual sex partner during the last sexual encounter [24]. A randomized, controlled trial of an online interactive e-animation found that 6 months after the intervention, the e-animation was associated with a 23% reduction in UAI and a significant reduction in willingness to engage in risk-taking behaviors by the participants [25].

Although online interventions have the potential to reach a large number of MSM, the ability to test online interventions provides unique challenges. The Centers for Disease Control and Prevention (CDC) Prevention Research Branch systematically reviews and summarizes HIV behavioral prevention interventions to identify best-evidence interventions (ie, interventions that demonstrate significant effects in eliminating or reducing sex- or drug-related risk behaviors, reducing the rate of new HIV infections, or increasing HIV-protective behaviors). The CDC evaluates these interventions using published criteria, which include prospective design, adequate retention of participants (> 70%), and preferably adequate representation of racial and ethnic minorities [26]. Although these criteria were established for offline HIV behavioral interventions, they can also be used as guidelines for the development and testing of online HIV interventions. Based on these guidelines, online intervention studies should retain participants, specifically black and Hispanic men, in the research study for a follow-up period sufficient to assess an intervention outcome. To date, however, prospective online studies of MSM that have included a follow-up component have demonstrated limited retention of participants [27] and an inability to retain participants of color [24]. Given that retention in online research of MSM has been modest and is below the 70% required by CDC's criteria for best-evidence interventions, it is critical that factors associated with retention in online HIV studies be evaluated in order to create research design and retention protocols that allow for testing of online interventions.

The current analysis examines the demographic and behavioral characteristics of a subset of MSM enrolled in an online study [28] designed to examine HIV behavioral risk factors of MSM to determine the factors that predicted retention of participants in the longitudinal study.

Methods

Study Design

MSM were recruited online through selective placement of banner advertisements on MySpace.com. The banner advertisements used for the study were provided and described in Sullivan et al [28]. Exposures of advertisements were made at varying times of the day to MySpace members in the United States who reported on their profile that they were male, at least 18 years of age, and were gay, bisexual, or unsure. As we have previously described [28], men who clicked through the banner advertisements were taken to an online survey where they were screened for eligibility. Men aged 18 years or older who reported at least one male sex partner in the 12 months prior to the baseline survey were eligible to participate and were taken to an online informed consent module. Men had to click agreement before enrolling in the study and beginning the baseline survey. The eligibility survey, informed consent module, and baseline survey were hosted on the secure servers of SurveyGizmo, an online survey provider. The 30-minute baseline survey queried participants on demographic information, their use of the Internet to meet sex partners, recent sexual risk behaviors, use of technology, HIV testing history, and interest in specific, new HIV prevention interventions.

Participation in a 3-month follow-up survey was offered to men who, in addition to the baseline eligibility criteria, were 35 years of age or older and reported their race/ethnicity as white non-Hispanic, black non-Hispanic, or Hispanic. The follow-up survey was limited to this group of men in order to compare retention patterns among MSM who are at the highest risk for HIV infection (ie, young black and Hispanic MSM) as these men would benefit the most from future online HIV prevention interventions. Participants who were eligible for the follow-up survey were randomized to be offered enrollment into the follow-up survey at four incentive levels: no incentive, US $5, US $10, and US $20.

Men who agreed to participate in the follow-up survey were asked to provide an email address so that the link (URL) to the follow-up survey could be emailed to them at the 3-month follow-up time point. The only way in which participants could access the follow-up survey was to click on the link provided in the email. If participants did not click on the link to the follow-up survey provided in the email, they were sent three additional reminder emails at 7, 14, and 21 days, respectively, after the first reminder email. These additional reminder emails were only sent to participants who had not linked into the survey from any previous reminder email. Each follow-up email provided participants with the option to withdraw from the study if they did not wish to take the follow-up survey or did not wish to receive additional emails. Participants who completed the follow-up survey received an incentive at the level to which they were randomized. These incentives were paid via PayPal, Amazon.com, or Target.com electronic gift cards that were sent to participants' email addresses. The study was reviewed and approved by the Institutional Review Board (IRB) of Emory University.

Statistical Analysis

We have previously provided a description of all study participants and an analysis of retention within the 30-minute baseline survey [28]. The current analysis is limited to the men who were eligible to participate in the follow-up survey and describes retention of these men in the longitudinal study. Two outcomes of interest were used to assess retention in the longitudinal study: agreeing (during the baseline survey) to participate in the follow-up survey and providing an email address (outcome 1) and linking into the follow-up survey at the 3-month follow-up time point (outcome 2).

Outcome 1: Agreement to Participate in the Follow-up Survey and Provide an Email Address

Eligible participants were asked as part of the baseline survey if they would like to participate in the follow-up survey in 3 months for either no incentive (if they were randomized to this group) or for an incentive (if they were randomized to receive US $5, US $10, or US $20). The survey explained to participants that a link to the follow-up survey would be sent to their email address in 3 months and that they must complete the follow-up survey at that time in order to be paid the incentive (if any). Participants who agreed to participate in the follow-up survey were prompted to enter their email address. Participants who initially agreed to participate in the follow-up survey but did not provide an email address were *not* considered to have agreed to participate in the follow-up survey. Those who agreed and provided an email address of valid format (ie, xxx@xxx.xxx) were considered to have agreed to participate.

Outcome 2: Linking Into the Follow-up Survey at the 3-month Follow-up Time Point

Emails with links to the follow-up survey generated at the 3-month follow-up time point were sent to participants who had provided an email address. Participants who clicked on the link to the follow-up survey were taken to the online survey and were considered to have linked into the follow-up survey regardless of what proportion of the follow-up survey they had completed. Participants who never clicked on the link provided in the email (or subsequent reminder emails) were not considered to have linked into the follow-up survey.

The two factors of primary interest were the association of race and incentive level with the two outcomes. Given that eligibility for the follow-up survey was limited to white non-Hispanic, black non-Hispanic, and Hispanic men, the primary factor of race in this analysis was restricted to those three racial/ethnic groups. To measure the effect of incentive level on retention in the follow-up survey, each level of incentive (ie, US $5, US $10, or US $20) was used as a separate factor to predict retention. However, because initial crude associations of incentive level with retention did not reveal differences in retention by level of incentive, men who were offered any incentive were grouped together for this analysis and compared with those men who were not offered an incentive.

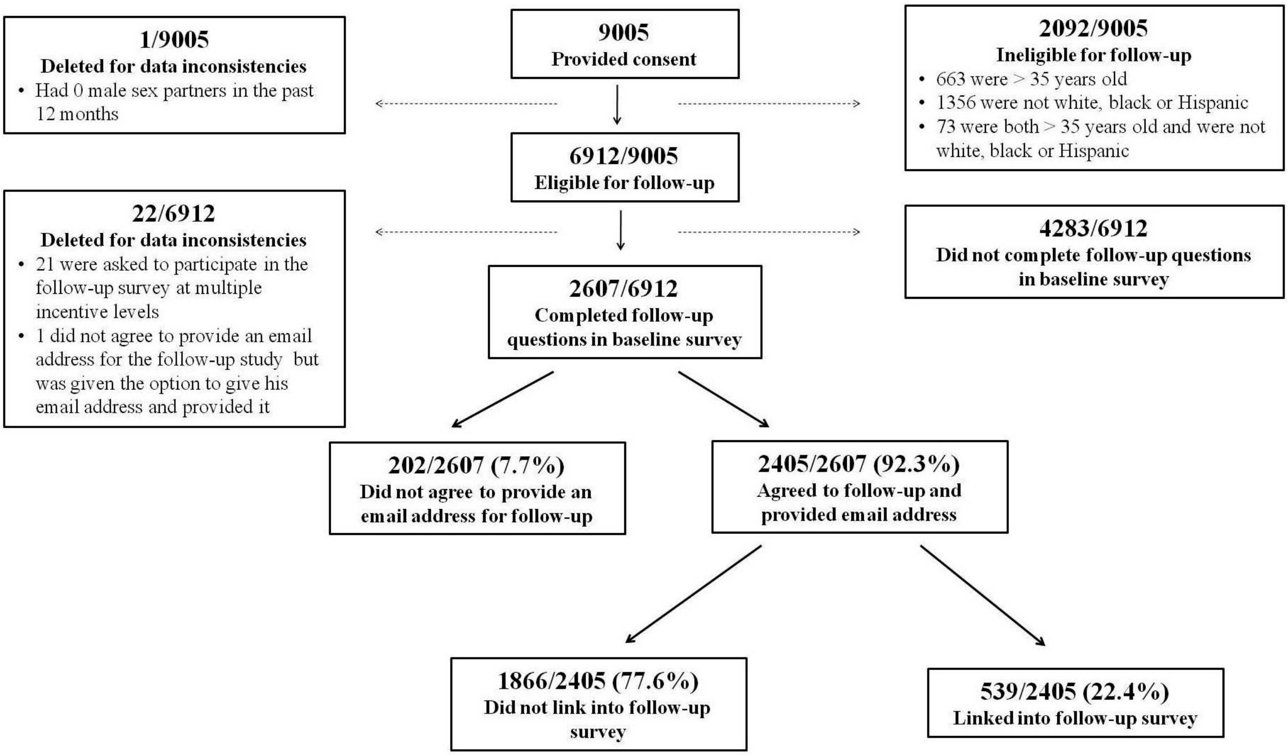
To examine the association of race and incentive with the two outcomes, factors identified as possible confounders in previous longitudinal online research studies of MSM were identified for use in this analysis [24,27]. These independent variables included demographic characteristics (age, education, and sexual identity) and sexual behaviors. Included in sexual behaviors were the number of male sex partners in the past 12 months, having UAI with a male sex partner in the past 12 months, having met a male sex partner on the Internet in the past 12 months, having a female sex partner in the past 12 months, having an exchange partner (defined in the survey as "someone who you have sex with in exchange for money, drugs, food, or something else of value" in the past 12 months), and HIV testing behaviors. For outcome 2, characteristics of the email addresses provided by participants as their preferred method of contact for the follow-up survey (type of email address, how email address was used, how long email address had been used, and frequency of checking email account) were included as independent factors of interest, in addition to the aforementioned variables.

Estimated logit plots for continuous independent factors of interest were analyzed to determine which variables would be more appropriately used as categorical or dichotomous variables. Univariate analyses examined frequencies and crude associations between each independent variable and the outcome variables of interest. These are reported as percentages and crude odds ratios (ORs) with 95% confidence intervals (CI), respectively. Analyzed were two multivariable logistic regression models to determine which factors were significantly associated ($P < .05$) with outcome 1 and outcome 2. All variables (regardless of the results from the univariate analysis) were considered in the multivariable models. Backward elimination was used to determine which variables were significantly associated with the outcome variables of interest. Since race and incentive were the primary independent factors of interest in the analyses, these variables were forced into each model during backward elimination. Independent variables that remained in the model at the conclusion of backward elimination were considered for two-way interactions with the outcome variables. All two-way interactions were considered together with a *P* value adjusted for simultaneous assessment to result in an aggregate alpha of .05 for evaluation of interaction. This interaction assessment revealed that there were no significant two-way interaction terms for either model. Collinearity diagnostics of the final models were analyzed using a SAS macro (Collinearity SAS macro obtained from David Kleinbaum, Emory University, Atlanta, GA). Using a condition index (CNI) of greater than 30 as an indication of a collinearity, there was no collinearity detected in either model. All multivariable model findings are reported as adjusted odds ratios (ORs) with 95% CI. Data analysis was completed using SAS 9.2 (SAS Institute, Cary, North Carolina, USA).

Results

Figure 1 details the enrollment and retention of study participants. Of the 9005 men who consented to participate in the baseline survey, 6912 (77%) were eligible to participate in the follow-up survey. Analysis of eligible participants revealed that a technological survey error resulted in 21 participants being offered multiple incentives for participation in the follow-up survey. Additionally, 1 participant did not agree to provide his email address for the follow-up survey but was mistakenly prompted to enter his email address on a subsequent survey screen, and 1 participant reported no male sex partners but was permitted to proceed in the survey. The results from these 23 participants are not included in this analysis. Of participants eligible for the follow-up survey, 62% (4283/6912) did not complete the final questions regarding participation in the follow-up survey; therefore, this analysis is restricted to the 2607 participants (38%) who were eligible for the follow-up survey and completed the follow-up questions in the baseline survey.

Figure 1. Flow chart of participant enrollment and retention in an online HIV behavioral risk study, United States, 2009



Among those 2607 participants, 1110 (43%) were white non-Hispanic, 550 (21%) were black non-Hispanic, and 947 (36%) were Hispanic. In all, 75% (1957/2607) of participants self-reported their sexual identity as homosexual, and 92% (2405/2607) had at least a high school level of education. Also, 64% (1612/2523) of participants reported having met a male sex partner online in the past 12 months, and 68% (1738/2547) reported having UAI with a male sex partner in the past 12 months. Engaging in sexual activity with a female sex partner or an exchange partner (male or female) were not frequently reported behaviors, at 10% (249/2607) and 7% (175/2582) of participants reporting, respectively.

Outcome 1: Agreement to Participate in the Follow-up Survey and Provide an Email Address

Of the 2607 participants included in this analysis, 2405 (92%) agreed to participate in the 3-month follow-up survey and provided an email address. Participant characteristics associated with agreement to participate in the follow-up survey are provided in Table 1. Hispanic men had significantly decreased odds of agreement to participate in the follow-up survey compared with white men. Approximately equal proportions of those participants who were offered an incentive and were not offered an incentive agreed to participate in the follow-up survey. In the crude analysis, men who agreed to participate in the follow-up survey were less likely to report being bisexual compared with being homosexual and were less likely to report having had a female sex partner in the past 12 months; however, these associations were not significant in the multivariable model. The only independent factor that was significantly associated with the outcome in the multivariable model was having had UAI with a male sex partner in the past 12 months, as participants reporting this activity had increased odds of agreeing to participate in the follow-up survey compared with men who did not report this sexual activity.

Table 1. Number, percentage, and odds of agreement to participate in an online follow-up survey by men who have sex with men enrolled in an online HIV behavioral risk study, by select demographic and behavioral characteristics (n = 2607), United States, 2009

	Agreed to Follow-up (n = 2405)	Did Not Agree to Follow-up (n = 202)		
Characteristic	n (%)	n (%)	Crude Odds Ratio (95% CI[a])	Adjusted Odds Ratio (95% CI)
Age (in years), mean (SD[b])	22.7 (4.4)	22.3 (4.3)	1.20 (0.86 - 1.69)[c]	--[d]
Number of male sex partners in past 12 months, mean (SD)	7.5 (17.4)	7.7 (15.5)	1.00 (0.99 – 1.01)[e]	--
Race [f]				
White, non-Hispanic	1038 (94)	72 (6)	Referent	Referent
Black, non-Hispanic	507 (92)	43 (8)	0.82 (0.55 - 1.21)	0.87 (0.57 - 1.32)
Hispanic	860 (91)	87 (9)	*0.69 (0.50 - 0.95) g*	*0.66 (0.47 - 0.92)*
Offered an incentive				
Yes	620 (93)	44 (7)	1.25 (0.88 - 1.76)	1.25 (0.87 - 1.79)
No	1785 (92)	158 (8)	Referent	Referent
Sexual identity				
Homosexual	1819 (93)	138 (7)	Referent	--
Bisexual	534 (90)	58 (10)	*0.71 (0.51 - 0.97)*	
Heterosexual	8 (80)	2 (20)	0.31 (0.06 - 1.46)	
Other[h]	25 (96)	1 (4)	1.92 (0.26 - 14.25)	
Education				
College/postgraduate	314 (92)	27 (8)	1.12 (0.71 - 1.77)	--
Some college/associate degree	1071 (93)	78 (7)	1.33 (0.96 - 1.83)	
High school/GED[i]	836 (91)	79 (9)	Referent	
Less than high school	170 (92)	15 (8)	1.09 (0.62 - 1.94)	
UAI [j] **with male sex partner in past 12 months**				
Yes	1621 (93)	117 (7)	*1.42 (1.05 - 1.92)*	*1.42 (1.05 - 1.93)*
No	734 (91)	75 (9)	1.00	Referent
Met MSP online in past 12 months				
Yes	1502 (93)	110 (7)	1.33 (0.99 - 1.80)	--
No	830 (91)	81 (9)	Referent	
Exchange partner [k] **in past 12 months**				
Yes	158 (90)	17 (10)	0.76 (0.45 - 1.28)	--
No	2225 (92)	182 (8)	Referent	
Ever been tested for HIV				
Yes	1725 (93)	133 (7)	1.29 (0.95 - 1.76)	--
No	654 (91)	95 (9)	Referent	
One or more female sex partners in past 12 months				
Yes	221 (89)	28 (11)	*0.63 (0.41 - 0.96)*	--
No	2184 (93)	174 (7)	Referent	

[a] CI: confidence Interval

[b] SD: standard deviation

[c] Odds ratio interpreted as odds of agreeing to follow-up every 10 year increase in age

[d] Denotes variable was not significant in the multivariable model and was therefore not included in the final model

[e] Odds ratio interpreted as odds of agreeing to follow-up per one partner increase

[f] Enrollment in the follow-up survey was limited to white non-Hispanic, black non-Hispanic, and Hispanic participants

[g] Italicized odds ratio and confidence interval denote significance at $P < .05$

[h] Participants could write in a text response for sexual identity; most frequent responses were "queer," "curious," and "questioning"

[i] GED: general equivalency diploma

[j] UAI: unprotected anal intercourse

[k] Exchange partner defined in the survey as "someone who you have sex with in exchange for money, drugs, food, or something else of value"

Outcome 2: Linking Into the Follow-up Survey at the 3-Month Follow-up Time Point

Only 539 of the 2405 participants (22%) who provided an email address for the follow-up survey actually linked into the follow-up survey at the 3-month follow-up time point. Table 2 details both the participant characteristics and the email characteristics associated with linking into the follow-up survey. Multivariable analysis revealed that the odds of linking into the survey by black men were about half the odds for linking into the survey by white men. There was no significant difference in the odds of linking into the survey for Hispanic men compared with white men.

Table 2. Number, percentage, and odds of linking into a follow-up survey by men who have sex with men enrolled in an online HIV behavioral risk study who agreed to participate in a follow-up survey, by select demographic, behavioral, and email characteristics (n = 2405), United States, 2009

	Linked Into Follow-up (n=539)	Did Not Link Into Follow-up (n=1866)		
Characteristic	n (%)	n (%)	Crude Odds Ratio (95% CI[a])	Adjusted Odds Ratio (95% CI)
Age (in years), mean (SD[b])	23.5 (4.6)	22.4 (4.3)	*1.74 (1.41 - 2.14)* [cd]	*1.35 (1.07 - 1.70)*
Number of male sex partners in past 12 months, mean (SD)	7.5 (17.4)	7.7 (15.5)	1.00 (0.99 - 1.01)[e]	--[f]
Race [g]				
White, non-Hispanic	264 (25)	774 (75)	Referent	Referent
Black, non-Hispanic	76 (15)	431 (85)	*0.52 (0.39 - 0.69)*	*0.47 (0.35 - 0.63)*
Hispanic	199 (23)	661 (76)	0.88 (0.72 - 1.09)	0.87 (0.70 - 1.08)
Offered an incentive				
Yes	150 (24)	470 (76)	1.15 (0.92 - 1.42)	*1.29 (1.02 - 1.62)*
No	389 (22)	1398 (78)	Referent	Referent
Sexual identity				
Homosexual	428 (24)	1391 (76)	Referent	--
Bisexual	99 (19)	435 (81)	*0.74 (0.58 - 0.94)*	
Heterosexual	3 (38)	5 (62)	1.95 (0.46 - 8.19)	
Other[h]	4 (16)	21 (84)	0.62 (0.21 - 1.81)	
Education				
College/postgraduate	106 (34)	208 (66)	*2.48 (1.85 - 3.33)*	*2.22 (1.61 - 3.05)*
Some college/associate degree	266 (25)	805 (75)	*1.61 (1.28 - 2.01)*	*1.52 (1.20 - 1.93)*
High school/GED[i]	141 (17)	695 (83)	Referent	Referent
Less than high school	22 (13)	148 (87)	0.72 (0.45 - 1.17)	0.73 (0.44 - 1.20)
UAI [j] **with male sex partners in past 12 months**				
Yes	380 (23)	1241 (77)	1.16 (0.94 - 1.44)	--
No	153 (21)	581 (79)	Referent	
Met male sex partners online in past 12 months				
Yes	366 (24)	1136 (76)	*1.29 (1.05 - 1.59)*	--
No	166 (20)	664 (80)	Referent	
Exchange partner [k] **in past 12 months**				
Yes	33 (21)	125 (79)	0.91 (0.61 - 1.35)	--
No	501 (23)	1724 (77)	Referent	
Ever been tested for HIV				
Yes	398 (23)	1327 (77)	1.15 (0.93 - 1.44)	--
No	135 (21)	519 (79)	Referent	
One or more female sex partners in past 12 months				
Yes	42 (19)	179 (81)	0.80 (0.56 - 1.13)	--
No	497 (23)	1687 (77)	Referent	
Type of email address				

Characteristic	Linked Into Follow-up (n=539) n (%)	Did Not Link Into Follow-up (n=1866) n (%)	Crude Odds Ratio (95% CI[a])	Adjusted Odds Ratio (95% CI)
Work address				
Yes	56 (26)	153 (73)	1.30 (0.94 - 1.79)	--
No	483 (22)	1713 (78)	Referent	
Personal address				
Yes	501 (23)	1722 (77)	1.10 (0.76 - 1.60)	--
No	38 (21)	144 (79)	Referent	
School address				
Yes	75 (30)	174 (70)	*1.57 (1.18 - 2.10)*	--
No	464 (22)	1692 (78)	Referent	
How email address used				
Communicate with friends				
Yes	213 (27)	581 (73)	*1.45 (1.19 - 1.76)*	--
No	326 (20)	1285 (80)	Referent	
Communicate with family				
Yes	173 (30)	413 (70)	*1.66 (1.35 - 2.06)*	--
No	366 (20)	1453 (80)		
Manage financial accounts				
Yes	132 (36)	238 (64)	*2.22 (1.75 - 2.82)*	*1.97 (1.54 - 2.52)*
No	407 (20)	1628 (80)		
Communicate with "hook-ups"				
Yes	82 (25)	249 (75)	1.17 (0.89 - 1.53)	--
No	457 (22)	1617 (78)		
Had email address > 1 year				
Yes	384 (23)	1274 (77)	1.04 (0.82 - 1.31)	--
No	117 (23)	403 (77)	Referent	
Check email account > once per day				
Yes	370 (26)	1055 (74)	*1.67 (1.36 - 2.05)*	*1.51 (1.22 - 1.87)*
No	162 (17)	772 (83)	Referent	Referent

[a] CI: confidence Interval

[b] SD: standard deviation

[c] Odds ratio interpreted as odds of linking into follow-up survey per every 10 year increase in age

[d] Italicized odds ratio and confidence interval denotes significance at $P < .05$

[e] Odds ratio interpreted as odds of linking into follow-up survey agreeing to follow-up per one partner increase

[f] Denotes variable was not significant in the multivariable model and was therefore not included in the final model

[g] Enrollment in the follow-up survey was limited to white non-Hispanic, black non-Hispanic, and Hispanic participants

[h] Participants could write in a text response for sexual identity; most frequent responses were "queer," "curious," and "questioning"

[i] GED: general equivalency diploma

[j] UAI: unprotected anal intercourse

[k] Exchange partner defined in the survey as "someone who you have sex with in exchange for money, drugs, food, or something else of value"

Participants who were offered an incentive for follow-up had significantly increased odds of linking into the follow-up survey compared with those who were not offered an incentive. An adjusted analysis of incentive by dollar amount offered was performed to determine which level of incentive (if any) was more likely to predict linking into the follow-up survey. Participants who were offered US $20 for completing the follow-up survey were significantly more likely to link into the follow-up survey (adjusted OR = 1.55, 95% CI 1.08-2.22) compared with men who were not offered any incentive. Participants who were offered $5 and $10 were no more likely to link into the follow-up survey compared with those who were offered no incentive (US $5: adjusted OR = 1.02, 95% CI 0.71-1.48; US $10: adjusted OR = 1.35, 95% CI 0.96-1.90; test for trend, *P* = .049).

A higher proportion of men reporting having completed college or some college linked into the follow-up survey compared with men with a high school level of education and men with less than a high school education. From the multivariable model, having a college education or some college education were associated with increased odds of linking into the follow-up survey compared with a high school level of education. Moreover, older participants were significantly more likely to link into the follow-up survey compared with younger participants.

Table 2 also describes the characteristics of the email addresses provided by participants for the follow-up survey, controlling for the other factors (participant characteristics) in the model of outcome 2. Of the 2405 participants who provided an email address for follow-up, 2223 (92%) reported that the email address they provided was a personal email address, 249 (10%) reported it was a school email address, and 206 (9%) reported it was a work email address (participants were permitted to indicate that the email address was used for multiple purposes so percentages do not equal 100%). Participants who provided a school email address or an email they used to communicate with friends or with family had greater unadjusted odds of linking into the follow-up survey compared with those who did not provide these types of email addresses, but these associations were not significant in the multivariable model. Participants who used their email address for financial management or online ordering had approximately twice the odds of linking into the follow-up survey compared with men who did not provide an email address of this type. Frequency of checking the email account was also associated with increased odds of linking into the survey: participants who indicated they checked their email account at least once a day had greater odds of linking into the follow-up survey compared with participants who reported checking their email account less frequently than once per day.

Discussion

This analysis identified factors that predict retention in an online, prospective follow-up study of MSM. Overall, only 21% (539/2607) of participants who were eligible to participate in the follow-up survey agreed to participate and returned at the 3-month follow-up time point. Hispanic men were less likely to agree to participate in the follow-up survey than white men, but participants reporting high-risk sexual behavior (ie, UAI in the past 12 months) were more likely to agree to follow-up. The latter result perhaps indicates that participants who may benefit the most from an online intervention (ie, are at high-risk for HIV) are willing to participate in longitudinal research. Although offering an incentive did not increase the odds of agreeing to participate in the follow-survey, participants who were offered an incentive were significantly more likely to link into the follow-up survey at the 3-month follow-up time point. Additionally, older men and participants with a higher level of education had increased odds of linking into the follow-up survey, but black men were significantly less likely to link into the follow-up survey compared with white men. Participants overwhelmingly agreed to complete the follow-up survey, with over 90% of respondents providing an email address to be contacted at the 3-month follow-up time point. However, at follow-up only a small proportion (22%) of respondents providing their email address actually returned to complete the survey, indicating that intention to participate in the follow-up survey may not necessarily predict actual participation.

We asked participants who provided their email address to answer a series of questions regarding the type, use, and frequency of use of the email address they were providing. This allowed us, for the first time to our knowledge, to analyze the characteristics of the email addresses provided by participants as possible factors predicting retention in an online study. We found that email addresses provided by participants that were used for online financial management or were addresses that participants checked at least once a day were associated with linking into the survey at the follow-up time point. These data suggest that participants who provide these types of email addresses (ie, an email address that may be important to the participant) are more likely to be retained in the study.

While the overall rate of retention in this study was quite poor, this result is not especially uncommon in online follow-up studies. Retention rates in online longitudinal research have typically been lower than in conventional offline follow-up studies [27], and the difficulty in retaining study participants in online research spans studies from smoking cessation to diabetes to depression [27,29,30]. In fact, the high number of participants lost to follow-up in Internet research has prompted researchers to title this common phenomenon the *law of attrition*, described as the fundamental methodological challenge in longitudinal online research [31]. Although the difficulty in retaining online study participants is widely recognized, our study furthers knowledge by providing an analysis to compare participants who were retained in the study and those who were lost to follow-up.

This analysis suggests that online studies of MSM similar to this one may experience differential attrition of black and Hispanic men, leading to a discrepant underrepresentation of these groups in the final outcome analysis. There are a number of possible reasons for the decreased retention of racial/ethnic minorities in this and similar studies. First, black and Hispanic Americans have less access to private high speed Internet than white Americans. Data from the US Census 2009 Current Population Survey indicated that 46% of black households and 47% of Hispanic households do not have Internet access in their homes, compared with 27% of white households [32]. Among households that have access to Internet in the home, 68% of white households have high speed Internet access compared with only 50% and 48% of black and Hispanic households, respectively [32]. This lack of access to home high speed Internet has resulted in black and Hispanic Americans using the Internet in public locations, such as libraries [33]. If black and Hispanic participants in our study did not have home access to the Internet, it is possible that they may not have seen the email reminders sent with a link to the follow-up study if they were unable to access the Internet during the time in which the reminder emails were sent or checked their email infrequently during that time. However, black and Hispanic participants did not indicate that they checked their email address less frequently than white participants; therefore, it is also possible that black and Hispanic participants were accessing their email at their place of work or in a public venue where they were not comfortable linking into the follow-up survey, given their knowledge of the sensitive nature of the baseline survey questions. This is difficult to discern, however, since participants were not asked for the location at which they most frequently accessed their email account.

While it has been noted that the historic distrust by racial/ethnic minorities of research studies may extend to the realm of online research [34], it is unclear if it is a contributing factor in this study. We speculate that because Hispanic participants were less likely to provide an email address for the follow-up survey, they may not have felt comfortable disclosing their email address or did not trust that their confidentiality would be maintained. Given that after having just completed a personal and sensitive online survey black participants were equally as likely to provide their email address compared with white men, it is it is unlikely that black men did not link into the follow-up survey because of a lack of trust.

This study has a number of limitations. Participants in the study are not representative of MSM who use MySpace.com or MSM in the United States, and we have previously characterized the potential selection bias among the same group of study participants [28]. Because this study relied on self-reported characteristics of participants, we cannot be sure that the reported information from participants such as race or age was correct, and, thus, misclassification bias may have resulted. However, because advertisements were targeted to MySpace user profiles based on selected demographic profile information, respondents would have had to consistently provide incorrect information both in their profile and in the online survey to participate.

Upon completion of the study, technological faults in the survey were discovered. In one question of the baseline survey, if a participant pushed the Enter key instead of clicking on the Next button on the survey screen, they were mistakenly taken to the end of the survey and were not permitted to go back to the previous screen. Therefore, these men, who may have been eligible for the follow-up survey, were never asked if they were willing to participate in the follow-up survey. As previously noted, 21 men were offered incentives at multiple levels and were not included in this analysis. While this represented a small proportion of the sample, it is still a technological error that needs to be addressed in future studies. Additionally, there was a technological difficulty in the emails sent to participants with the link to the follow-up survey. These emails were generated from the non-Emory server (SurveyGizmo) hosting the survey, but the return email address was from the emory.edu domain. Because many email account spam filters filter emails that are generated from a server that is contradictory to the email extension of the sending email address, it is possible that a large portion of the reminder emails were either not delivered to participants or were delivered to a quarantine or spam box, where they were less likely to be recovered by participants. This may have contributed to our low overall retention. However, there is no mechanism with which to track the number of emails that were filtered away from participants' in-boxes, so the magnitude of the effect of this error on survey retention remains unclear. In future research, requesting participants to add the sending email address to their list of acceptable senders may increase the likelihood of participants receiving study-related emails.

Although this analysis has a number of limitations, the results from this study may provide guidance in the development of strategies to increase retention in online HIV prevention research with MSM. First, we identified that, consistent with similar online studies of MSM, black and Hispanic men are less likely to be retained in the research study. Since a possible explanation for the decreased retention in these groups is access to high speed Internet, participants should be asked about their access to high speed Internet in follow-up studies to further understand this issue. However, we must also explore new avenues to contact these participants at follow-up time points. Recent research suggests that black and Hispanic Americans have higher mobile phone ownership than white Americans, and that 79% of black and 83% of English-speaking Hispanic adults report sending and receiving text messaging on a typical day, compared with 68% of white adults [35]. Therefore, the use of mobile phones, specifically short messaging system (SMS or text messaging), may be an alternative to the Internet for data collection in future studies, since its accessibility for black and Hispanic MSM may be helpful in increasing the retention of these study participants.

Second, we found that offering an incentive increased the likelihood of participants being retained in the study. Although offering a US $20 incentive was significantly associated with linking into the follow-up survey while US $5 and US $10 were not, we are cautious to conclude that US $20 is the minimum amount to offer participants in order to increase retention. It has been suggested that offering incentives may lead to participants enrolling in studies multiple times in order to claim more than one incentive. This study included two elements to ensure that this would not occur. First, men could only access the survey by clicking on the banner advertisement posted on their profile. Because the advertisements were displayed to random users at varying times of day, it is unlikely that a participant would have seen an advertisement twice. Second, multiple surveys could not be completed from the same Internet protocol (IP) address, so a participant would have had to change his IP address in order to take a second survey. Therefore we believe that we were successful in paying incentives only once to participants and that it is feasible to offer incentives in future online research. Finally, the analysis of email characteristics of participants provides guidance for ways in which to solicit participant email addresses. Future studies may see increased retention by prompting participants to provide an email address that they check at least once per day or to provide an email address that is important to them (ie, one they use for financial management). This may increase the likelihood of participants viewing an email request to complete follow-up surveys, for example. Additionally, although not included as part of this analysis, it may be of interest in future studies to examine whether economic considerations other than providing an incentive (ie, paying money for sex) differentiates groups that participate in online longitudinal research. Further, examining the effect of geographic region (eg, rural versus urban or western versus eastern United States) on participation in longitudinal Internet-based studies may be informative.

As the HIV epidemic among MSM in the United States persists and MSM continue to meet sex partners online, it is essential to create effective methods for the proper creation and testing of online HIV prevention interventions. This study identified factors that may predict enrollment or retention in online research so that mitigation strategies to decrease participant attrition can be developed. Only by creating ways in which to retain participants in online research can critical outcome measures of interventions be appropriately assessed.

Acknowledgments

The authors disclose receipt of the following financial support for the research and/or authorship of this article: Emory Center for AIDS Research (P30 AI050409) and the National Center for Minority Health and Health Disparities (1RC1MD004370-01, 5RC1MD00437-02).

Conflicts of Interest

None declared

References

1. Sullivan P, Wolitski R. HIV infection among gay and bisexual men. In: Wolitski RJ, Stall R, Valdiserri RO, editors. Unequal Opportunity: Health Disparities Affecting Gay and Bisexual Men in the United States. New York, NY: Oxford University Press; 2008:220-247.
2. Centers for Disease Control and Prevention. Diagnosis of HIV and AIDS in the United States and Dependent Areas, 2008.: HIV Surveillance Report, Vol 20; 2010. URL: http://www.cdc.gov/hiv/surveillance/resources/reports/2008report/index.htm [accessed 2010-12-17] [WebCite Cache ID 5v3FSBy7N]
3. Centers for Disease Control and Prevention (CDC). HIV prevalence estimates--United States, 2006. MMWR Morb Mortal Wkly Rep 2008 Oct 3;57(39):1073-1076 [FREE Full text] [Medline: 18830210]
4. US Census Bureau. Population Estimates: National - Characteristics. 2010. Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States: April 1, 2000 to July 1, 2009 URL: http://www.census.gov/popest/national/asrh/NC-EST2009-srh.html [accessed 2010-12-17] [WebCite Cache ID 5v3BbHk4o]
5. Centers for Disease Control and Prevention (CDC). Trends in HIV/AIDS diagnoses among men who have sex with men--33 states, 2001-2006. MMWR Morb Mortal Wkly Rep 2008 Jun 27;57(25):681-686 [FREE Full text] [Medline: 18583954]
6. Sullivan PS, Hamouda O, Delpech V, Geduld JE, Prejean J, Semaille C, Annecy MSM Epidemiology Study Group. Reemergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996-2005. Ann Epidemiol 2009 Jun;19(6):423-431. [doi: 10.1016/j.annepidem.2009.03.004] [Medline: 19460672]
7. Fenton KA, Imrie J. Increasing rates of sexually transmitted diseases in homosexual men in Western europe and the United States: why? Infect Dis Clin North Am 2005 Jun;19(2):311-331. [doi: 10.1016/j.idc.2005.04.004] [Medline: 15963874]
8. Holmes WC, Pace JL. HIV-seropositive individuals' optimistic beliefs about prognosis and relation to medication and safe sex adherence. J Gen Intern Med 2002 Sep;17(9):677-683. [Medline: 12220363]
9. Hart GJ, Elford J. Sexual risk behaviour of men who have sex with men: emerging patterns and new challenges. Curr Opin Infect Dis 2010 Feb;23(1):39-44. [doi: 10.1097/QCO.0b013e328334feb1] [Medline: 19949328]
10. Parsons JT, Schrimshaw EW, Wolitski RJ, Halkitis PN, Purcell DW, Hoff CC, et al. Sexual harm reduction practices of HIV-seropositive gay and bisexual men: serosorting, strategic positioning, and withdrawal before ejaculation. AIDS 2005 Apr;19(Suppl 1):S13-S25. [Medline: 15838191]
11. Hanson J, Posner S, Hassig S, Rice J, Farley TA. Assessment of sexually transmitted diseases as risk factors for HIV seroconversion in a New Orleans sexually transmitted disease clinic, 1990-1998. Ann Epidemiol 2005 Jan;15(1):13-20. [doi: 10.1016/j.annepidem.2004.05.001] [Medline: 15571989]
12. Dougan S, Elford J, Chadborn TR, Brown AE, Roy K, Murphy G, et al. Does the recent increase in HIV diagnoses among men who have sex with men in the UK reflect a rise in HIV incidence or increased uptake of HIV testing? Sex Transm Infect 2007 Apr;83(2):120-125. [doi: 10.1136/sti.2006.021428] [Medline: 17090569]
13. Rudy ET, Shoptaw S, Lazzar M, Bolan RK, Tilekar SD, Kerndt PR. Methamphetamine use and other club drug use differ in relation to HIV status and risk behavior among gay and bisexual men. Sex Transm Dis 2009 Nov;36(11):693-695. [doi: 10.1097/OLQ.0b013e3181ad54a3] [Medline: 19704399]

14. Wong W, Chaw JK, Kent CK, Klausner JD. Risk factors for early syphilis among gay and bisexual men seen in an STD clinic: San Francisco, 2002-2003. Sex Transm Dis 2005 Jul;32(7):458-463. [Medline: 15976605]
15. Elford J, Bolding G, Sherr L, Hart G. High-risk sexual behaviour among London gay men: no longer increasing. AIDS 2005 Dec 2;19(18):2171-2174. [Medline: 16284470]
16. Bolding G, Davis M, Hart G, Sherr L, Elford J. Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners? AIDS 2005 Jun 10;19(9):961-968. [Medline: 15905678]
17. Bolding G, Davis M, Hart G, Sherr L, Elford J. Where young MSM meet their first sexual partner: the role of the Internet. AIDS Behav 2007 Jul;11(4):522-526. [doi: 10.1007/s10461-007-9224-9] [Medline: 17347876]
18. Liau A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. Sex Transm Dis 2006 Sep;33(9):576-584. [doi: 10.1097/01.olq.0000204710.35332.c5] [Medline: 16540884]
19. McFarlane M, Bull SS, Rietmeijer CA. The Internet as a newly emerging risk environment for sexually transmitted diseases. JAMA 2000 Jul 26;284(4):443-446 [FREE Full text] [Medline: 10904506]
20. Kim AA, Kent C, McFarland W, Klausner JD. Cruising on the Internet highway. J Acquir Immune Defic Syndr 2001 Sep 1;28(1):89-93. [Medline: 11579282]
21. Horvath KJ, Rosser BR, Remafedi G. Sexual risk taking among young internet-using men who have sex with men. Am J Public Health 2008 Jun;98(6):1059-1067 [FREE Full text] [doi: 10.2105/AJPH.2007.111070] [Medline: 18445804]
22. Benotsch EG, Kalichman S, Cage M. Men who have met sex partners via the Internet: prevalence, predictors, and implications for HIV prevention. Arch Sex Behav 2002 Apr;31(2):177-183. [Medline: 11974643]
23. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. J Acquir Immune Defic Syndr 2004 Apr 1;35(4):421-426. [Medline: 15097159]
24. Chiasson MA, Shaw FS, Humberstone M, Hirshfield S, Hartel D. Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). AIDS Care 2009 Sep;21(9):1081-1089. [doi: 10.1080/09540120902730013] [Medline: 20024766]
25. de Wit J, Adam P. Reducing non-premeditated risk-taking in MSM: a new intervention protocol to increase vigilance and control tested for efficacy in a prospective RTC. 2008 Presented at: XVII International AIDS Conference; August 3-8, 2008; Mexico City, Mexico.
26. Centers for Disease Control and Prevention. HIV Risk Reduction: Efficacy Review: Efficacy Criteria.: Department of Health and Human Services; 2010. URL: http://www.cdc.gov/hiv/topics/research/prs/efficacy_criteria.htm [accessed 2011-04-18] [WebCite Cache ID 5y2OHVNyZ]
27. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. AIDS Care 2004 Nov;16(8):931-943. [doi: 10.1080/09540120412331292507] [Medline: 15511725]
28. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
29. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
30. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008;10(5):e40 [FREE Full text] [doi: 10.2196/jmir.993] [Medline: 19017582]
31. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
32. US Census Bureau. Computer and Internet Use. 2009. Internet Use in the United States: 2009 - Detailed Tables URL: http://www.census.gov/population/www/socdemo/computer/2009.html [accessed 2010-11-15] [WebCite Cache ID 5uGBWJgpZ]
33. US Department of Commerce. A Nation Online: How Americans Are Expanding Their Use of the Internet.: National Telecommunications and Information Administration; 2002. URL: http://www.ntia.doc.gov/ntiahome/dn/anationonline2.pdf [accessed 2010-11-15] [WebCite Cache ID 5uGBnQgA1]
34. Fogel J, Ribisl KM, Morgan PD, Humphreys K, Lyons EJ. Underrepresentation of African Americans in online cancer support groups. J Natl Med Assoc 2008 Jun;100(6):705-712. [Medline: 18595573]
35. Smith A. Mobile Access 2010. Washington, DC: Pew Internet & American Life Project; 2010 Jul 07. URL: http://www.pewinternet.org/Reports/2010/Mobile-Access-2010.aspx [accessed 2011-01-25] [WebCite Cache ID 5w0Y095d0]

Abbreviations

CDC: Centers for Disease Control and Prevention
CI: confidence interval

CNI: condition index
GED: general equivalency diploma
HIV: human immunodeficiency virus
IP: Internet protocol
IRB: institutional review board
MSM: men who have sex with men
OR: odds ratio
SMS: short messaging system
UAI: unprotected anal intercourse

Edited by G Eysenbach; submitted 21.12.10; peer-reviewed by B Piper; comments to author 13.04.11; revised version received 18.04.11; accepted 19.04.11; published 11.07.11.

Please cite as:
Khosropour CM, Sullivan PS
Predictors of Retention in an Online Follow-up Study of Men Who Have Sex With Men
J Med Internet Res 2011;13(3):e47
URL: http://www.jmir.org/2011/3/e47/
doi:10.2196/jmir.1717
PMID:21745792

Original Paper

A Web-Based Computerized Adaptive Testing (CAT) to Assess Patient Perception in Hospitalization

Tsair-Wei Chien[1], MBA; Wen-Chung Wang[2], PhD; Sheng-Yun Huang[2], SRA; Wen-Pin Lai[3], MD; Julie Chi Chow[4], MD

[1]Department of Management, Chi Mei Medical Center, Yungkang, Taiwan
[2]Assessment Research Center, The Hong Kong Institute of Education, Hong Kong, China
[3]Department of Emergency, Chi Mei Medical Center, Yungkang, Taiwan
[4]Department of Paediatrics, Chi Mei Medical Center, Yungkang, Taiwan

Corresponding Author:
Julie Chi Chow, MD
Department of Paediatrics
Chi Mei Medical Center
No. 901 Junghua Rd.
Yungkang, 710
Taiwan
Phone: 886 62812811 ext 52903
Fax: 886 62820534
Email: jchow@mail.chimei.org.tw

Abstract

Background: Many hospitals have adopted mobile nursing carts that can be easily rolled up to a patient's bedside to access charts and help nurses perform their rounds. However, few papers have reported data regarding the use of wireless computers on wheels (COW) at patients' bedsides to collect questionnaire-based information of their perception of hospitalization on discharge from the hospital.

Objective: The purpose of this study was to evaluate the relative efficiency of computerized adaptive testing (CAT) and the precision of CAT-based measures of perceptions of hospitalized patients, as compared with those of nonadaptive testing (NAT). An Excel module of our CAT multicategory assessment is provided as an example.

Method: A total of 200 patients who were discharged from the hospital responded to the CAT-based 18-item inpatient perception questionnaire on COW. The numbers of question administrated were recorded and the responses were calibrated using the Rasch model. They were compared with those from NAT to show the advantage of CAT over NAT.

Results: Patient measures derived from CAT and NAT were highly correlated ($r = 0.98$) and their measurement precisions were not statistically different ($P = .14$). CAT required fewer questions than NAT (an efficiency gain of 42%), suggesting a reduced burden for patients. There were no significant differences between groups in terms of gender and other demographic characteristics.

Conclusions: CAT-based administration of surveys of patient perception substantially reduced patient burden without compromising the precision of measuring patients' perceptions of hospitalization. The Excel module of animation-CAT on the wireless COW that we developed is recommended for use in hospitals.

(J Med Internet Res 2011;13(3):e61) doi:10.2196/jmir.1785

KEYWORDS

Computerized adaptive testing; computer on wheels; classic test theory; IRT; item response theory; nonadaptive testing

Introduction

As computer technology and health care become more integrated, many hospitals have adopted mobile nursing carts that can be easily rolled up to a patient's bedside to access charts and help nurses perform their rounds [1-3]. Besides increasing efficiency by including basic functions such as billing records and decreasing the number of trips nurses need to take to the medication room [3], the carts can reduce patient burden by allowing them to answer questions on activities of daily living using computerized adaptive testing (CAT) [1]. However, few papers have reported data regarding the bedside use of wireless computers on wheels (COW) to collect questionnaire-based information on their perception of hospitalization on discharge from the hospital. Collecting patients' feedback on their perspectives has become an important part of patient involvement and participation for health caregivers; thus, this question is important [4-6].

Gathering Feedback Efficiently From Patients

Two new modes of survey administration have been reported to make surveys more easily accessible to those who cannot read or write [7]. These include using automated telephone technology through an interactive voice response system and using Internet-like visualizations to complete questionnaires online. In medical practice, hospital staff usually hand a questionnaire to patients at the end of their visit and ask them to complete it prior to leaving hospital. At the Picker Institute Europe [5], questionnaires are sent annually to a randomized list of eligible patients who had been discharged from the hospital. Both of these methods are less prompt and efficient than using wireless COW to collect data on patients' perspectives on being discharged from the hospital.

Computer Assessment and Computer-Adaptive Testing

There is no doubt that using wireless COW at a patient's bedside is an efficient way of instantly gathering feedback from patients. Traditional paper-and-pencil or computer-based devices (nonadaptive testing [NAT]) impose a large respondent burden because patients are required to answer all the questions. In contrast, CAT-based tests developed using item response theory (IRT) [8] can achieve a similar degree of measurement precision to NAT using only about half the test length [1,9-11]. Most studies investigating IRT- and CAT-based tests have evaluated both efficiency and precision for CAT-based tests with dichotomous items. Whether CAT-based tests with polytomously scored items (CAT as defined in this study) can be incorporated with wireless COW in hospitals for gathering feedback from patients should be investigated.

Rasch Analysis

In classical test theory, raw scores (or linear transformation scores, eg, T score) are usually adopted as respondent measures. However, subsequent parametric statistical analyses, such as computing mean, variance, correlation coefficient, and Cronbach alpha [12,13], would be incorrect because raw scores are not on an additive interval scale [14].

To overcome this obstacle, the IRT-based Rasch model [15], a probabilistic relationship between a person's level of a latent trait (commonly referred to as ability or measure) and an item's property (difficulty or threshold), was developed. Both person ability and item difficulty (calibrated in terms of log odds or logits) are located along the same continuum. A useful scale (or questionnaire) is usually examined by 3 important criteria for the Rasch model, namely, unidimensionality, item fit, and item invariance (or so-called differential item functioning [16]). These criteria are detailed in Smith et al [17]. There are many published papers [1,18-21] of studies using the Rasch model to develop CAT in clinical settings, but none of them have incorporated the Internet-based polytomously scored CAT to gather feedback from patients in hospitals.

Objectives

The purpose of this study was to evaluate the relative efficiency of an Internet-based polytomously scored CAT and the precision of CAT-based measures of perceptions of hospitalized patients, as compared with those measured by NAT. An Excel (Microsoft Corporation, Redmond, WA, USA) module of our CAT multicategory assessment is provided as an example.

Methods

Data collection

Participants

The study sample was recruited from inpatients at a 1333-bed medical center in southern Taiwan. Patients who had been discharged were selected randomly by the digit code of their invoice number during each morning and afternoon interval from Monday through Friday in summer 2010.

Procedure

As an incentive for participation, patients were offered a gift card for US $3.20 good for purchases at 7-11 convenience stores. A total of 200 patients either completed the questionnaire on COW themselves or were helped by a trained volunteer (if they were unable to personally complete the questionnaire); proxies were allowed because most of those helping patients carry out their discharge procedure were the patients' family members or friends. Time spent by each patient was recorded in Excel after they completed the questionnaire. This study was approved and monitored by the Research and Ethical Review Board of the Chi-Mei Medical Center, Tainan, Taiwan.

Tool: CAT-Format Questionnaire

We designed the 18-item CAT questionnaire in Excel based on an 18-item inpatient perception questionnaire (IPQ-18) [5]; see Table 1). Unidimensionality, local independence, item fit, and differential item functioning using the Rasch model to investigate these criteria have been previously reported [5].

Data collected from the patients included demographic characteristics (gender, treatment department, age, and person completing survey, ie, proxy or patient); perception measure in a logit unit; number of items needed to be completed; and mean square errors (MNSQ) of infit and outfit (indicators of response patterns for the IPQ-18 scale [5]) (see Table 1, Multimedia Appendix 1, and Multimedia Appendix 2).

Table 1. Items of the 18-item scale ordered by item overall difficulties

	Scale content		Difficulty				
Item number	Category[a]	Item	Overall	Step1	Step2	Step3	Step4
39	L	Did staff tell you about medication side effects when going home?	3.78	0.02	1.87	5.35	7.89
41	L	Did doctors or nurses give your family information needed to help you?	2.76	–1.00	0.85	4.33	6.87
27	N	Did hospital staff talk about your worries and fears?	2.22	–1.54	0.31	3.79	6.33
11	W	Were you ever bothered by noise at night from other patients?	1.58	–2.18	–0.33	3.15	5.69
24	N	Were you involved in decisions about your care and treatment?	0.67	–3.09	–1.24	2.24	4.78
30	N	How long was it after using the call button before you got the help you needed?	0.42	–3.34	–1.49	1.99	4.53
42	L	Did staff tell you how to contact them if worries arose after leaving?	–0.3	–4.06	–2.21	1.27	3.81
9	A	Did you feel you waited a long time to get to a bed on a ward?	–0.63	–4.39	–2.54	0.94	3.48
44	O	How would you rate how well the doctors and nurses worked together?	–0.71	–4.47	–2.62	0.86	3.4
2	A	How organized was the care you received in the emergency room?	–0.95	–4.71	–2.86	0.62	3.16
5	A	Were you given enough notice of your date of admission?	–1.08	–4.84	–2.99	0.49	3.03
12	W	Were you bothered by noise at night from hospital staff?	–1.1	–4.86	–3.01	0.47	3.01
17	D	Did you have confidence and trust in the doctors treating you?	–1.1	–4.86	–3.01	0.47	3.01
23	N	Did staff say one thing and something quite different happened to you?	–1.1	–4.86	–3.01	0.47	3.01
38	L	Did staff explain the purpose of the medicines so that you could understand?	–1.1	–4.86	–3.01	0.47	3.01
18	D	Did doctors talk in front of you as if you weren't there?	–1.12	–4.88	–3.03	0.45	2.99
19	N	Did you get answers that you could understand from a nurse?	–1.12	–4.88	–3.03	0.45	2.99
34	P	Did hospital staff do everything they could to help you control your pain?	–1.12	–4.88	–3.03	0.45	2.99

[a] Categories are A: admission to hospital; D: doctors; L: leaving hospital; N: nurses; O: overall; P: pain; W: hospital and ward.

CAT Procedure

Outfit and Infit Statistics

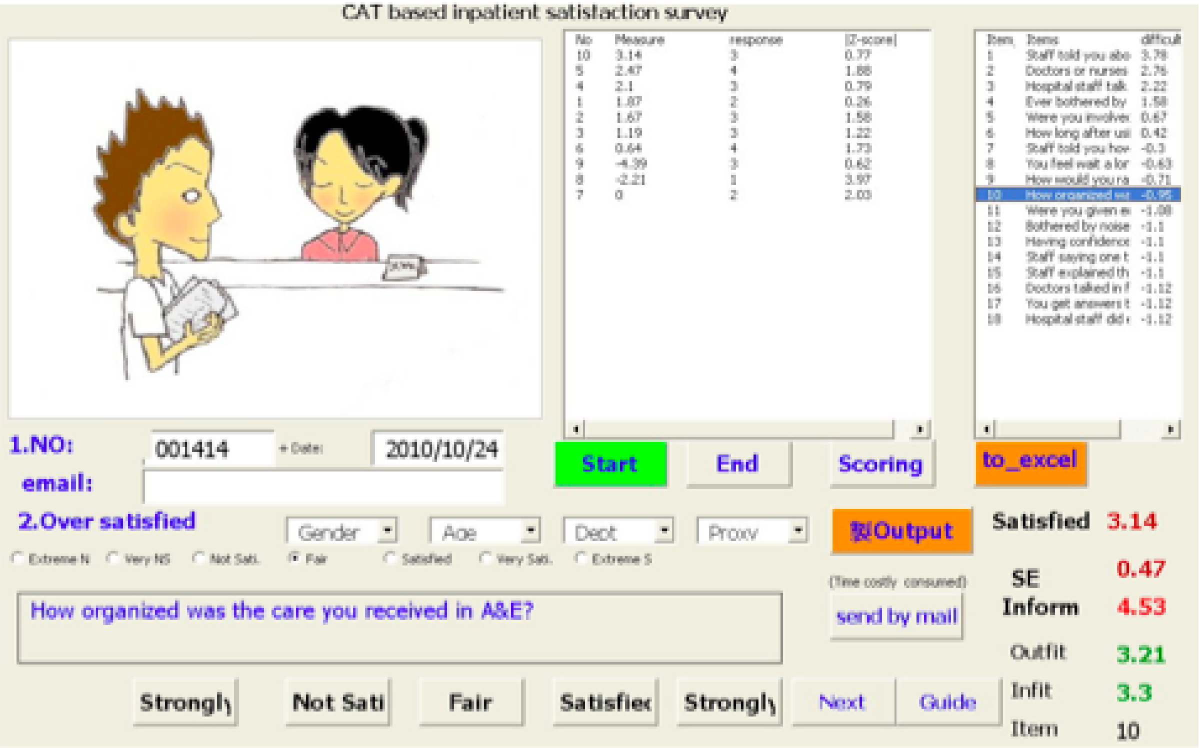
Outfit statistics are based on unweighted sum of squared standardized residuals and are sensitive to unexpected outlying, off-target, and low-information responses; whereas infit statistics are based on weighted sum of squared standardized residuals and are sensitive to unexpected inlying patterns among informative and on-target observations [22]. Smith [23] found that Rasch outfit MNSQ approaching 1.0 [24] demonstrates a higher power than its counterpart of infit MNSQ. Outfit MNSQ of 2.0 or greater for a patient indicate a possibly aberrant response pattern [24].

CAT Procedures and Features

We programmed a Visual Basic for Applications (VBA) module in Microsoft Excel and on the Internet (http://www.healthup.org.tw/cat.asp, http://www.webcitation.org/60xWv6p6d) complying with several rules and criteria for CAT-based testing on COW (Figure 1, Figure 2). The person separation reliability (similar to Cronbach alpha) calculated from the original paper [5] was 0.94 (mean 2.64, SD 2.09). Based on this number, the CAT stop rule for measurement of standardized error was determined to be 0.51(SD × sqrt(1 – alpha) = 2.09 × sqrt(1 – 0.94)).

Figure 1. Using a wireless computer on wheels (COW) to collect data on patients' perspectives on hospitalization

Figure 2. Snapshot of computerized adaptive testing (CAT)-based inpatient perception questionnaire for patients



CAT based inpatient satisfaction survey

No	Measure	response	[Z-score]
10	3.14	3	0.77
5	2.47	4	1.88
4	2.1	3	0.79
1	1.87	2	0.26
2	1.67	3	1.58
3	1.19	3	1.22
6	0.64	4	1.73
9	-4.39	3	0.62
8	-2.21	1	3.97
7	0	2	2.03

Item	Items	difficult
1	Staff told you abo	3.78
2	Doctors or nurses	2.76
3	Hospital staff talk	2.22
4	Ever bothered by	1.58
5	Were you involve	0.67
6	How long after usi	0.42
7	Staff told you how	-0.3
8	You feel wait a lor	-0.63
9	How would you ra	-0.71
10	How organized wa	-0.95
11	Were you given e	-1.08
12	Bothered by noise	-1.1
13	Having confidence	-1.1
14	Staff saying one t	-1.1
15	Staff explained th	-1.1
16	Doctors talked in f	-1.12
17	You get answers t	-1.12
18	Hospital staff did	-1.12

1.NO: 001414 Date: 2010/10/24

email:

2.Over satisfied

○ Extreme N ○ Very NS ○ Not Sati. ● Fair ○ Satisfied ○ Very Sati. ○ Extreme S

How organized was the care you received in A&E?

Satisfied 3.14

SE 0.47

Inform 4.53

Outfit 3.21

Infit 3.3

Item 10

We also set another stop rule so that the minimum number of questions required for completion was 10 items (10/18, 56%), because CAT achieves a similar measurement precision to NAT with only about half the test length [1,9-11]. The initial question was selected from the pool of 18 items according the patient's overall perception of satisfaction in their hospitalization. The provisional person measure is estimated by the maximum of the log-likelihood function using an iterative Newton-Raphson procedure [1,10] (Multimedia Appendix 2) after 3 items were responded to, without all answers corresponding to either 0 or 4. The next question selected was the one with the highest information value among the remaining unanswered questions weighted against the provisional person measure. The details of CAT procedures are shown in Multimedia Appendix 2 and Multimedia Appendix 3.

Comparison of Efficiency Between NAT and CAT

Two indicators used to examine CAT efficiency in this study are (1) whether the number of questions needed was significantly less than for NAT (18 questions) and (2) whether the precision of person measures was less than that for NAT. We used paired *t* tests to make these two statistical inferences.

Accordingly, the perception measure based on NAT should be estimated in advance for each patient who was assumed to have answered all 18 items. The following steps were thus followed: (1) we used a standard item response-generation method [25-29] to generate responses to 18 questions for each patient with given question difficulty parameters (in Table 1) and a patient perception measure (by CAT), and (2) we applied the rectangle metric of 18 questions × 200 persons to re-estimate NAT measures for each patient using WINSTEPS software (WINSTEPS version 3.72.0: Winsteps.com, Chicago, IL, USA) (the 18-question difficulties are anchored in WINSTEPS with iafile shown in Multimedia Appendix 2).

Statistical Analysis

SPSS software for Windows (Version 12, SPSS, Chicago, IL) was used for all statistical analysis.

Descriptive Statistics

Data on patient gender, age, treatment department, and proxy respondent were collected. Noncontinuous variables were reported as frequency and percentages, and were examined by chi-square tests.

Analytic statistics

For continuous variables, CAT and NAT measures were compared using the Pearson correlation coefficient. Patient perception measures obtained by CAT were compared between groups using *t* tests or analysis of variance (ANOVA). Time spent by patients was averaged by using logarithmic transformation and reported as mean (SD) by exponential function. All analyses were considered to be statistically significant at the .05 alpha level.

Results

As seen in Table 2, there were no significant associations between gender and other demographic characteristics (ie, treatment department, age, and participant). Among inpatients we approached, 13% (31/231) were unwilling to respond to the CAT questions due to personal reasons, despite the incentive we offered. CAT and NAT measures were highly correlated (*r* = 0.98).

Table 2. Demographic characteristics of the study population (N = 200)

Variable	Male		Female			$\chi^2_{(r-1)*(c-1)}$	
	n	%	n	%	Total	Test	*P* value
Respondent						0.6	.45
Willing to participate	100	50	100	50	200		
Unwilling to participate	13	42	18	58	31		
Age (years)						0.9	.82
≤16	31	31	25	25	56		
17–40	27	27	30	30	57		
41–70	25	25	27	27	52		
>70	17	17	18	18	35		
Department						3.9	.42
Internal medicine	44	44	41	41	85		
Surgery	28	28	22	22	50		
Obstetrics and gynecology	8	8	14	14	22		
Pediatrics	11	11	7	7	18		
Other	12	12	16	16	28		
Participant /proxy						1.1	.57
Family	75	75	81	81	156		
Friend	15	15	12	12	27		
Patient	10	10	7	7	17		

Mean time spent by patients in CAT was 54.91 seconds (SD 8.00; maximum 76; minimum 33). As shown in Table 3, CAT required substantially fewer questions than NAT (*P* < .001). NAT required all of the 200 patients to respond to all 18 questions, and thus yielded a total of 3600 responses. In CAT, a total of 2083 responses were required, meaning that on average a patient answered 10.42 questions. Thus, as compared with NAT, CAT received an efficiency gain in test length of 0.42 (defined as 1 – ratio of total responses by CAT and NAT: 1 – 2084/3600).

Table 3. Comparison of computerized adaptive testing (CAT) versus nonadaptive testing (NAT) (all questions having to be answered) in efficiencya as assessed by paired t test

	Mean	Variance	Response	Maximum	Minimum	Paired t_{199} test	*P* value
Test length (number of questions answered)							
NAT	18	0.00	3600[b]	18	18	–476.72	<.001
CAT	10.42	0.25	2084[b]	12	10		
Estimated measures(mean and variance)							
NAT	0.69	2.66	3600	4.16	–2.69	1.10	.14
CAT	0.71	2.62	2084	4.00	–2.56		
Time spent (seconds)							
CAT	54.91[c]	64.04[c]	2084	76[3]	33[3]		

[a] Efficiency = (1 – 2084/3600) = 0.58.

[b] 3600 = 200 × 18; 2084 = 200 × 10.42.

[c] On second unit.

Regarding precision of measurement, person measures from CAT did not statistically differ from those from NAT (*P* = .14). ANOVA revealed that patient perception measures did not differ between groups across elements; *t* test analyses showed that they also did not differ between genders (Table 4).

Table 4. Comparison of inpatient perception by demographic characteristic

Variable	Male		Female		ANOVA[a]	
	Mean	SD	Mean	SD	Test	*P* value
Proportion	0.77	1.59	0.65	1.66	$t_{398} = 0.55$	.59
Age (years)					$F_{3,196} = 0.71$	.55
≤16	0.77	1.72	0.83	1.81	–0.12	.89
17–40	1.23	1.54	0.58	1.40	1.68	.09
41–70	0.72	1.48	0.53	1.74	0.42	.67
>70	0.13	1.45	0.69	1.83	–1.00	.32
Department					$F_{4,195} = 0.92$	.45
Internal medicine	0.65	1.53	0.49	1.48	0.47	.63
Surgery	0.61	1.56	0.9	1.77	–0.77	.44
Obstetrics and gynecology	1.00	1.91	0.77	1.70	0.28	.77
Pediatrics	0.45	1.79	0.19	2.00	0.30	.78
Other	1.73	1.29	0.68	1.85	1.67	.11
P articipant/p roxy					$F_{2,197} = 0.36$	.69
Family	0.90	1.58	0.60	1.62	1.14	.25
Friend	0.58	1.60	0.93	2.10	–0.49	.62
Patient	0.16	1.62	0.72	1.43	–0.73	.47

[a] Analysis of variance.

Total person mean 0.71 logits (SD 1.62); median 0.59; coefficient of skewness 0.103 (*P* = .54); coefficient of kurtosis –0.89 (*P* = .03); D'Agostino-Pearson test for normal distribution accept normality (*P* = .09).

Discussion

Key Finding

The results from this study indicate that CAT-based testing using COW can reduce patient (or proxy) burdens. It is up to 42% more efficient in answering questions and achieves a very similar degree of measurement precision to NAT.

What This Adds to What Was Known

Consistent with the literature [1,9-11,30], the efficiency of CAT was supported. We confirmed that the CAT-based IPQ-18 on COW requires significantly fewer questions to measure patient perception than NAT, but does not compromise precision of measurement.

What is the Implication, What Should be Changed

Using an Excel module of animation for CAT on COW as a tool that can help hospital staff efficiently and precisely gather feedback from patients is technically feasible. Outfit MNSQ of 2.0 or greater can be used to examine whether patient responses are distorted or abnormal—that is, many more responses unexpectedly did not fit the model's requirement and were deemed to be very likely to be careless, mistaken, or awkward [1,5,6,24]. Thus, CAT makes detection of problematic responses easier—normally, inspecting problematic feedback from patients using classical test theory is rather difficult.

Strength of This Study

There are 2 major forms of standardized assessments in clinical settings [31]: (1) a lengthy questionnaire to achieve a precise assessment that requires significant amounts of time and training to administer, and (2) a rapid short-form scale that briefly screens for the most common symptoms using cut-off points to identify degrees of impairment [32,33]. CAT has the advantages of both forms: precision and efficiency. This paper reports several achievements, including using the Rasch rating scale model [34] (instead of dichotomy) to design CAT in a perception survey, applying CAT on a COW, and offering an Excel module with an animation prototype (demonstrated in Multimedia Appendix 2 or http://www.healthup.org.tw/cat.asp). This module and available files can complement the limited uses for interactive voice response or Internet-like visualization online [7] to complete questionnaires and put them into clinical practice.

We conducted an actual CAT-based survey instead of CAT with simulations. This study demonstrates the whole procedure of a CAT-based survey, from the beginning of data collection (Figure 2and Multimedia Appendix 3) through the end of the evaluation report (Table 4), and fulfills the goal of creating a Web-CAT with graphs and animations, as stated in our previous paper [35].

Limitations of the Study

Several issues should be considered more thoroughly in further studies. First, a total of 200 patients were surveyed on the IPQ-18. The generalizability of this study needs to be investigated with more studies on different samples and different instruments. Second, there is a potential sampling bias in this study. Those who completed the IPQ-18 CAT on COW tended to be younger; and proxies were used to represent patients to complete the discharge procedure from hospital, because they were selected randomly by the digit code of their invoice number on the patient's discharge. The proportion of proxies, who are assumed to be healthier and more capable of filling out a questionnaire, was very high (183/200, 91.5%; see Table 2). This sample therefore does not reflect mostly the patients' perspective on hospitalization, which possibly affects the study results shown in Table 4. Third, the patient burden was determined by the number of questions administered in this study. Other indices may be collected where available, such as time and effort required for test administration, and accessibility of the hospital [33, 34].

In addition, we set at least 10 items in CAT to be completed as one of the stop rules, which might inflate the test length to some extent. As a result, the test length of CAT was about 58% that of NAT, a little higher than in previous studies with about half the test length [1,9-11].

Applications

A large variety of behavior-change techniques and other methods to promote exposure to interventions have been used [36]. There are concerns about how to entice patients (or proxies) to complete surveys before they are discharged from the hospital. Offering reward points or coupons good for credits toward another service is recommended because perception surveys are not similar to other clinical scales conducted by clinicians, where patients themselves consider the benefits to their health.

A telephone survey with CAT-based administration or patient self-report on the Internet (demonstrated at http://www.healthup.org.tw/cat.asp) can be combined with the CAT on COW for gathering feedback from patients easily, quickly, and efficiently.

There are many issues that should be addressed in the future, including studies that address the limitations noted above. For example, using CAT on COW at patients' bedsides to gather their feedback before discharge from the hospital can solve the problem of sampling bias (eg, when proxies constitute a high proportion of respondents) and warrants further study. Surveying perceptions of hospital service via the Internet by CAT-type telephone or self-report is encouraged to complement CAT on COW and questionnaires delivered by mail to discharged patients, such as the Picker Institute Europe's annual survey.

One of the important advantages of CAT scoring is that the item pool can be expanded without changing the metric [37]. CAT administrators may expand the IPQ-18 item pool or replace items with other kinds of questions as presented in the Excel spreadsheet example. It must be noted that (1) overall item and step (threshold) difficulties of the questionnaire must be calibrated in advance using Rasch analysis (eg, the IPQ-18 of this study was examined by Rasch analysis in a previous paper [5]), and (2) picture and voice files for each question should be well prepared in an appropriate folder that can be shown simultaneously with the corresponding question in an animation module of CAT.

Conclusion

CAT-based administration of surveys of patient perception reduces patient burden without compromising measurement precision. The Excel module for animation-CAT on COW connected to a mainframe computer is recommended for assessing patients' perceptions of their experience in the hospital.

Acknowledgments

This study was supported by Grant 98cm-kmu-18 from the Chi Mei Medical Center, Taiwan.

Conflicts of Interest

None declared

Authors' Contributions

Chien,Lai and Chou collected all data, generated the database, designed and performed the statistical analysis and wrote the manuscript. Wang and Huang contributed to the development of the study design and advised on the performance of the statistical analysis. The analysis and results were discussed by all authors together. Chien contributed to the Excel programming, helped to interpret the results and helped to draft the manuscript. All authors read and approved the final manuscript.

Multimedia Appendix 1

Excel VBA module for CAT delivering results to the website through an Internet address

[ZIP file (Zip Archive), 1141 KB - jmir_v13i3e61_app1.zip]

Multimedia Appendix 2

Comprehensive overview of Rasch models and the CAT process

[PDF file (Adobe PDF File), 405 KB - jmir_v13i3e61_app2.pdf]

Multimedia Appendix 3

Screenshot of the module with an animation-CAT design

[WMV file (Windows Media Video File), 16,264 KB - jmir_v13i3e61_app3.wmv]

References

1. Chien TW, Wu HM, Wang WC, Castillo RV, Chou W. Reduction in patient burdens with graphical computerized adaptive testing on the ADL scale: tool development and simulation. Health Qual Life Outcomes 2009;7:39 [FREE Full text] [doi: 10.1186/1477-7525-7-39] [Medline: 19416521]
2. Briggs B. Point of care on a roll. Health Data Manag 2003 Dec;11(12):48-50. [Medline: 14682255]
3. Lavin M, Sierzega G, Pucklavage D, Kleinbach D, Gogal C, Bokovoy J. Carts and care. Roll out safer medication delivery and smoother workflow with mobile technology. Nurs Manage 2007 Nov;Suppl Pharmacy:16-18. [doi: 10.1097/01.NUMA.0000298279.92200.8c] [Medline: 18176103]
4. Davies AR, Ware JE. Involving consumers in quality of care assessment. Health Aff (Millwood) 1988;7(1):33-48 [FREE Full text] [Medline: 3360392]
5. Chien TW, Wang WC, Wang HY, Lin HJ. Online assessment of patients' views on hospital performances using Rasch model's KIDMAP diagram. BMC Health Serv Res 2009;9:135 [FREE Full text] [doi: 10.1186/1472-6963-9-135] [Medline: 19646267]
6. Chien TW, Wang WC, Lin SB, Lin CY, Guo HR, Su SB. KIDMAP, a web based system for gathering patients' feedback on their doctors. BMC Med Res Methodol 2009;9:38 [FREE Full text] [doi: 10.1186/1471-2288-9-38] [Medline: 19534773]
7. 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] [doi: 10.2196/jmir.6.3.e29] [Medline: 15471755]
8. Lord FM. Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum; 1980.
9. Wainer HW, Dorans NJ, Flaugher R, Green BF, Mislevy RJ, Steinberg L, et al. Computerized adaptive testing: a primer. Hillsdale, N.J.: L. Erlbaum Associates; 1990.
10. Embretson S, Reise S. Item response theory for psychologists. Mahwah, N.J.: L. Erlbaum Associates; 2000:158-186.
11. Weiss DJ. Improvement measurement quality and efficiency with adaptive testing. Applied Psychological Measurement 1982;6:473-492. [doi: 10.1177/014662168200600408]
12. Crocker L Algina J. Introduction to classical and modern test theory. New York: Holt, Rinehart, and Winston; 1986.
13. Nunnally JC. Bernstein .IH. Psychometric theory. New York: McGraw-Hill; 1994.
14. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989 Nov;70(12):857-860. [Medline: 2818162]
15. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press; 1980.
16. Holland PW, Wainer H. Differential item functioning. Hillsdale: Lawrence Erlbaum Associates; 1993.
17. Smith AB, Wright P, Selby PJ, Velikova G. A Rasch and factor analysis of the Functional Assessment of Cancer Therapy-General (FACT-G). Health Qual Life Outcomes 2007;5:19 [FREE Full text] [doi: 10.1186/1477-7525-5-19] [Medline: 17448239]
18. Velozo CA, Wang Y, Lehman L, Wang JH. Utilizing Rasch measurement models to develop a computer adaptive self-report of walking, climbing, and running. Disabil Rehabil 2008;30(6):458-467. [doi: 10.1080/09638280701617317] [Medline: 18297500]
19. Lehman LA, Woodbury M, Shechtman O, Wang YC, Pomeranz J, Gray DB, et al. Development of an item bank for a computerised adaptive test of upper-extremity function. Disabil Rehabil 2011 Mar 14. [doi: 10.3109/09638288.2011.560336] [Medline: 21401332]
20. Öztuna D, Elhan AH, Küçükdeveci AA, Kutlay S, Tennant A. An application of computerised adaptive testing for measuring health status in patients with knee osteoarthritis. Disabil Rehabil 2010;32(23):1928-1938. [doi: 10.3109/09638281003777572] [Medline: 20384449]
21. Elhan AH, Oztuna D, Kutlay S, Küçükdeveci AA, Tennant A. An initial application of computerized adaptive testing (CAT) for measuring disability in patients with low back pain. BMC Musculoskelet Disord 2008;9:166 [FREE Full text] [doi: 10.1186/1471-2474-9-166] [Medline: 19094219]
22. Linacre JM, Wright BD. Chi-square fit statistics. Rasch Meas Trans 1994;8(2):360.
23. Smith RM. Fit analysis in latent trait measurement models. J Appl Meas 2000;1(2):199-218. [Medline: 12029178]
24. Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas 2002;3(1):85-106. [Medline: 11997586]
25. Kieffer KM, Reese RJ. A reliabilty generalization study of the ceriatric scale. Educational and Psychological Measurement 2002;62(6):969-994. [doi: 10.1177/0013164402238085]
26. Harwell M, Stone CA, Hsu TC, Kirisci L. Monte Carlo studies in item response theory. Applied Psychological Measurement 1996, 20, 101-125 1996. [doi: 10.1177/014662169602000201]
27. Macdonald P, Paunonen SV. A monte carlo comparison of item and person statistics based on item response theory versus classical test theory. Educational and Psychological Measurement 2002;62:921-943. [doi: 10.1177/0013164402238082]

28. Chien TW, Lin SJ, Wang WC, Leung HW, Lai WP, Chan AL. Reliability of 95% confidence interval revealed by expected quality-of-life scores: an example of nasopharyngeal carcinoma patients after radiotherapy using EORTC QLQ-C 30. Health Qual Life Outcomes 2010;8:68 [FREE Full text] [doi: 10.1186/1477-7525-8-68] [Medline: 20626903]
29. Linacre JM. How to simulate Rasch data. Rasch Meas Trans 2007;21(3):1125.
30. Ware JE, Kosinski M, Bjorner JB, Bayliss MS, Batenhorst A, Dahlöf CG, et al. Applications of computerized adaptive testing (CAT) to the assessment of headache impact. Qual Life Res 2003 Dec;12(8):935-952. [Medline: 14651413]
31. Eack SM, Singer JB, Greeno CG. Screening for anxiety and depression in community mental health: the beck anxiety and depression inventories. Community Ment Health J 2008 Dec;44(6):465-474. [doi: 10.1007/s10597-008-9150-y] [Medline: 18516678]
32. Ramirez Basco M, Bostic JQ, Davies D, Rush AJ, Witte B, Hendrickse W, et al. Methods to improve diagnostic accuracy in a community mental health setting. Am J Psychiatry 2000 Oct;157(10):1599-1605 [FREE Full text] [Medline: 11007713]
33. Shear MK, Greeno C, Kang J, Ludewig D, Frank E, Swartz HA, et al. Diagnosis of nonpsychotic patients in community clinics. Am J Psychiatry 2000 Apr;157(4):581-587 [FREE Full text] [Medline: 10739417]
34. Andrich D. A rating formulation for ordered response categories. Psychometrika 1978;43:561-573.
35. Chien TW, Lai WP, Lu CW, Wang WC, Chen SC, Wang HY, et al. Web-based computer adaptive assessment of individual perceptions of job satisfaction for hospital workplace employees. BMC Med Res Methodol 2011;11:47 [FREE Full text] [doi: 10.1186/1471-2288-11-47] [Medline: 21496311]
36. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
37. Aday LA. Designing and conducting health surveys: a comprehensive guide. San Francisco: Jossey-Bass Publishers; 1996.

Abbreviations

ANOVA: analysis of variance
CAT: computerized adaptive testing
COW: computers on wheels
IPQ: inpatient perception questionnaire
IRT: item response theory
MNSQ: mean square errors
NAT: nonadaptive testing
VBA: Visual Basic for Applications

Edited by G Eysenbach; submitted 26.02.11; peer-reviewed by T Bond, CL Hsieh, A Smith; comments to author 13.04.11; revised version received 05.05.11; accepted 09.05.11; published 15.08.11.

Please cite as:
Chien TW, Wang WC, Huang SY, Lai WP, Chow JC
A Web-Based Computerized Adaptive Testing (CAT) to Assess Patient Perception in Hospitalization
J Med Internet Res 2011;13(3):e61
URL: http://www.jmir.org/2011/3/e61/
doi:10.2196/jmir.1785
PMID:21844001

Original Paper

Development of a Web-Based Survey for Monitoring Daily Health and its Application in an Epidemiological Survey

Hiroaki Sugiura[1], MD; Yasushi Ohkusa[2], PhD; Manabu Akahane[1], MD, PhD; Tomomi Sano[1], MD; Nobuhiko Okabe[2], MD, PhD; Tomoaki Imamura[1], MD, PhD

[1]Department of Public Health, Health Management and Policy, Nara Medical University School of Medicine, Kashihara, Japan
[2]Infectious Disease Surveillance Center, National Institute of Infectious Disease, Tokyo, Japan

Corresponding Author:
Hiroaki Sugiura, MD
Department of Public Health, Health Management and Policy
Nara Medical University School of Medicine
840 Shijo-cho
Kashihara, 634-8521
Japan
Phone: 81 744 22 3051
Fax: 81 744 25 7657
Email: tomomarie@smn.enjoy.ne.jp

Abstract

Background: Early detection of symptoms arising from exposure to pathogens, harmful substances, or environmental changes is required for timely intervention. The administration of Web-based questionnaires is a potential method for collecting information from a sample population.

Objective: The objective of our study was to develop a Web-based daily questionnaire for health (WDQH) for symptomatic surveillance.

Methods: We adopted two different survey methods to develop the WDQH: an Internet panel survey, which included participants already registered with an Internet survey company, and the Tokyo Consumers' Co-operative Union (TCCU) Internet survey, in cooperation with the Japanese Consumers' Co-operative Union, which recruited participants by website advertising. The Internet panel survey participants were given a fee every day for providing answers, and the survey was repeated twice with modified surveys and collection methods: Internet Panel Survey I was conducted every day, and Internet Panel Survey II was conducted every 3 days to reduce costs. We examined whether the survey remained valid by reporting health conditions on day 1 over a 3-day period, and whether the response rate would vary among groups with different incentives. In the TCCU survey, participants were given a fee only for initially registering, and health information was provided in return for survey completion. The WDQH included the demographic details of participants and prompted them to answer questions about the presence of various symptoms by email. Health information collected by the WDQH was then used for the syndromic surveillance of infection.

Results: Response rates averaged 47.3% for Internet Panel Survey I, 42.7% for Internet Panel Survey II, and 40.1% for the TCCU survey. During a seasonal influenza epidemic, the WDQH detected a rapid increase in the number of participants with fever through the early aberration reporting system.

Conclusions: We developed a health observation method based on self-reporting by participants via the Internet. We validated the usefulness of the WDQH by its practical use in syndromic surveillance.

(J Med Internet Res 2011;13(3):e66) doi:10.2196/jmir.1872

KEYWORDS

Web-based survey; syndromic surveillance; long-term operation

Introduction

The collection of health crisis information has been an important task in every country since the 2005 implementation of the World Health Organization's International Health Regulations to prevent the global spread of illness [1]. Early detection of health events related to exposure to various pathogens, harmful substances, or environmental changes is indispensable for timely intervention to minimize health crises.

Syndromic surveillance is a method used to investigate epidemics of infections [2-5]. Unlike sentinel surveillance, which uses a traditional definitive diagnosis and pathogen identification, this method encompasses the surveillance of symptoms. For example, this type of surveillance has been used at medical institutions to determine the number of patients with fever, cough, diarrhea, or vomiting, and changes in the number of absentees from school or the workplace, sales of commercial drugs, and prescriptions [6-8]. Syndromic surveillance is important as a means of gathering information during the early stages of an epidemic, and it has practical application in many countries. Thus, an effective means of collecting daily health information from people directly and quickly is desirable.

Use of the Web to perform an epidemiological survey was reported in 1996 [9]. This method has since been applied to national-scale surveys in various countries where residents voluntarily input information on influenza-like symptoms directly into a dedicated website so that epidemiologists can gain an understanding of the influenza epidemic [10-13].

To broaden the range of such a survey in terms of contributors and infectious diseases, we developed and conducted a daily health survey of the general population using the Internet and named this survey the Web-based daily questionnaire for health (WDQH) [14]. We report the methodology of data collection and processing of the WDQH and clarify its use in syndromic surveillance. We performed this study with participants in panels registered at Internet survey companies. In addition, we investigated a method for the long-term operation of the survey by reducing the cost of each individual survey.

Methods

Recruitment

Two different methods were used to recruit participants for the WDQH. First, Internet panel surveys comprised people who were already registered with an established Internet survey company. Second, the Tokyo Consumers' Co-operative Union (TCCU) Internet survey comprised members of the TCCU, in cooperation with the Japanese Consumers' Co-operative Union (JCCU), who were invited to participate via advertising on the company website.

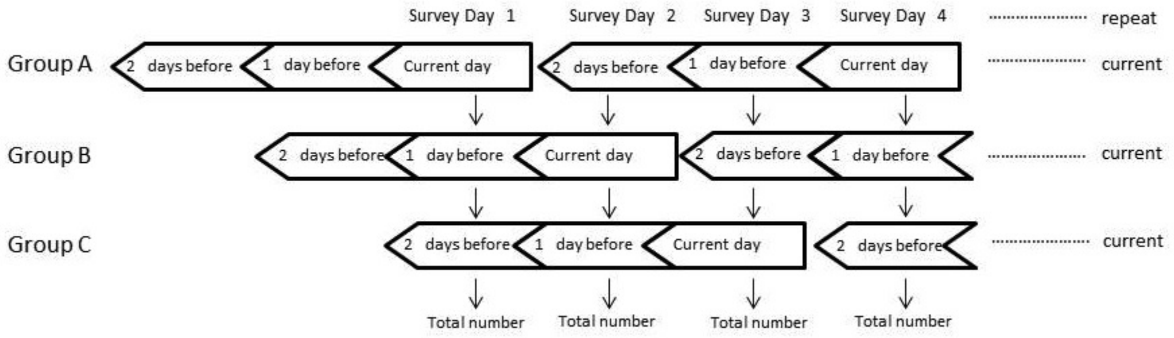
An Internet survey company conducts questionnaire surveys via the Internet. For survey participants registered in advance, questionnaires and a response column are displayed on the website for the respondents to complete and transmit their responses. Additionally, the Internet panel survey was repeated twice with different survey and collection methods (Internet panel surveys I and II). In the Internet panel surveys, the respondents were registered as panel members with the company and were residents of Izumo City (150,000 inhabitants) in western Japan, which had 89.5% Internet coverage in 2008. The youngest respondent was 16 years old. The respondents also provided information regarding symptoms in family members. Internet Panel Survey I was conducted daily between December 1, 2007, and March 28, 2008, among 245 respondents who were paid 60 yen (US $0.75;US $1.00 = 80 yen at the time of writing) per survey completed. A reminder email was sent daily to those who agreed to participate. Internet Panel Survey I included 702 registrants. Respondents were those who completed the survey, and all family members included in the survey were considered to be participants.

In Internet Panel Survey II, conducted between January 8 and March 13, 2009, we examined changes in the data acquisition method to reduce survey costs. We investigated whether reporting health conditions once every 3 days could reduce survey costs. Internet Panel Survey II was conducted continuously with 264 respondents and included 716 registrants. The respondents were divided into groups A, B, and C, and each group was surveyed by shifting the survey date by 1 day to determine the applicability of recall for 1-in-3-day reporting. Thus, groups A, B, and C received the questionnaire on days 1, 2, and 3, respectively. For example, the data on survey day 1 included symptoms that were experienced on the current day by group A, on the day before the survey day by group B, and 2 days before the survey day by group C. Thus, on survey day 1, responses were obtained from all the respondents (Figure 1). Each group was divided randomly. A reminder email was sent to each group on the survey day.

We also investigated whether the response rate varied according to the incentive. Each group was further divided into three subgroups in which the members were given a reward of 40 (US $0.50), 60 (US $0.75), or 80 yen (US $1.00). The response rates were then investigated.

In the TCCU survey, we examined methods of collecting health information from the city's residents without the use of an Internet survey company. The respondents were those who accessed the website of the TCCU's home delivery services and applied to participate in the survey, which was advertised with an onscreen banner. Participants were recruited between January 15 and January 31, 2009, and any applicant could participate. There were 427 respondents from Tokyo, which had 95.2% Internet coverage in 2008. They were given 100 yen (US $1.25) for registering. No fee was paid for each survey, but health information was provided to the participants in the relevant residential areas based on survey results. The TCCU survey was conducted in cooperation with the JCCU, which has 1 million members in Tokyo among a population of 12.3 million. The TCCU has a strong corporate philosophy regarding food safety and understanding the health concerns of consumers. Many respondents were homemakers, as the proportion of female respondents was 97.6%. They provided information about themselves and family members, providing 1453 participants, who were 49.5% male and 50.5% female.

Figure 1. Data collection method for Internet Panel Survey II.



Response Method

On the day of the survey, the survey administrator sent a reminder email to all those recruited. Respondents accessed the password-protected website designated in the email and responded to the questions. The questionnaire ascertained whether respondents or their family members had any symptoms. The gender and age (in 5-year intervals) of those who developed symptoms as well as their specific symptoms (Table 1) were noted. In Internet Panel Survey I, 6 symptoms associated with diseases of infection and bioterrorism were selected. In Internet Panel Survey II and in the TCCU survey, 12 symptoms associated with seasonal allergic diseases and changes in body conditions were added, and "fever" was divided into "slight fever" and "high fever."

Table 1. Items in the three surveys

Internet Panel Survey I	Internet Panel Survey II (TCCU[a] survey)
Fever	Slight fever
Cough	High fever
Diarrhea	Runny nose
Vomiting	Cough
Eruption	Diarrhea
Convulsion	Vomiting
	Convulsion
	Eye itch
	Eruption
	Diagnosis of influenza
	Diagnosis of gastroenteritis
	Arthritic pain
	Muscle pain
	Shoulder stiffness
	Sneeze
	Skin itch
	Rough hands
	Sleeplessness
	Decreased concentration

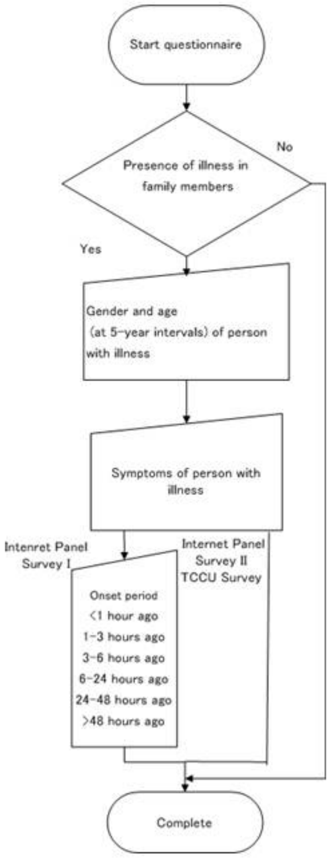
[a] Tokyo Consumers' Co-operative Union.

These surveys were conducted with varied symptoms to examine whether the WDQH could be applied in the surveillance of various diseases according to symptoms.

Reports by asymptomatic people are essential for calculating prevalence rates and an analysis of variance by the presence of symptoms. The symptoms quoted in this survey are common, particularly during the acute phases of diseases caused by infection and environmental factors. The time of symptom onset was determined (<1 hour ago, 1–3 hours ago, 3–6 hours ago, 6–24 hours ago, 24–48 hours ago, and >48 hours ago) (Figure 2) in Internet Panel Survey I.

In Internet Panel Survey II and the TCCU survey, only cases with an onset on the current day were reported. If a respondent tried to exit the survey without answering all the questions, the system would alert the respondent in order to prevent invalid responses.

Figure 2. Flowchart for all three surveys (TCCU = Tokyo Consumers' Co-operative Union).

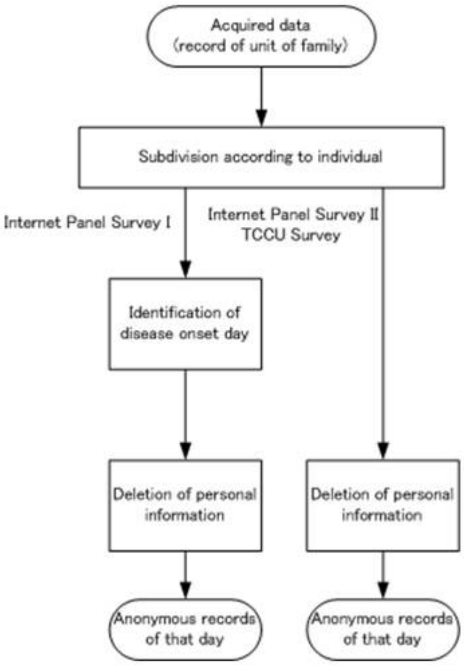


Data Processing

The survey also included details such as survey date, presence of illness in family members, presence of symptoms in family members, and time of symptom onset in a family member. Data entered in the WDQH were then transmitted to a server managed by a researcher. Subsequently, the records for each household were subdivided by family member. In Internet Panel Survey I, the symptom onset dates for participants were determined from the time elapsed between symptom onset and reporting. Those with a symptom onset >48 hours before the survey were excluded. This was intended to include only people with symptoms at an acute stage. Personal information was then deleted (Figure 3).

The final participant records consisted of survey date, presence of illness and presence of symptoms by participant, and symptom onset date. Cross-correlation was used during syndromic surveillance with the collected data to determine the number of participants by symptom and date.

Figure 3. Data processing flowchart for all three surveys (TCCU = Tokyo Consumers' Co-operative Union).



Examples of Using the Data in Syndromic Surveillance

Symptoms were cross-tabulated to determine the symptom onset dates and number of participants who developed a particular symptom. The results were used to prepare time-series graphs by symptom, with the prevalence of symptoms plotted against the date. Subsequently, alerts by symptom were reported on the day when the number of participants who developed the symptom increased rapidly compared with the baseline of the previous 10 days using the early aberration reporting system (EARS) algorithm recommended by the US Centers for Disease Control and Prevention [15,16].

This study was approved by the Ethical Committee of Nara Medical University (Authorization Code: 220).

Results

Respondents' Demographic Characteristics and Response Rates

Table 2 presents the number of respondents, gender, age distribution, number of participants including families of the respondents, and daily mean response rates for the three surveys. The numbers of respondents (total number of participants) in Internet Panel Survey I, Internet Panel Survey II, and the TCCU survey were 245 (702), 264 (716), and 427 (1453), respectively.

Table 2. Demographics and response rates of participants in the three surveys

	Internet Panel Survey I	Internet Panel Survey II	TCCU[a] survey
Number of respondents	245	264	427
Men	44.5%	52.7%	2.6%
Women	55.5%	47.3%	97.6%
Age distribution of respondents (years)			
≤29	26.5%	26.9%	4.6%
30–39	43.3%	41.7%	35.3%
40–49	21.9%	21.9%	39.4%
50–59	5.8%	7.9%	16.3%
≥60	2.5%	1.6%	4.4%
Daily mean response rate	47.3%	42.7%	40.1%
Total number of participants	702	716	1453

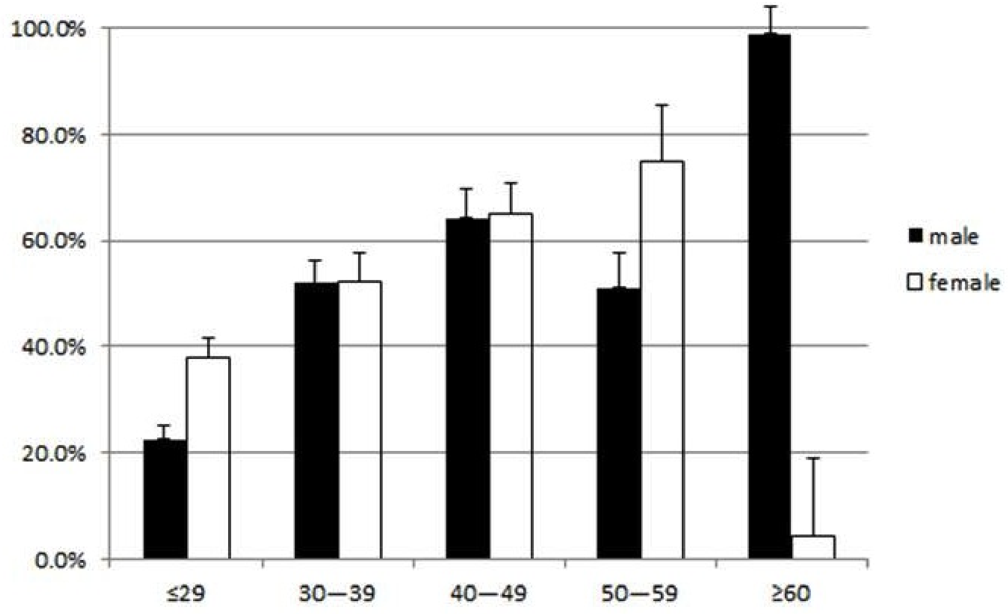
[a] Tokyo Consumers' Co-operative Union.

In Internet Panel Survey I, a constant response rate was observed from the initiation to the end of the survey, and there was no tendency to respond when only 1 symptom was present. The response rate was 48.7% on weekdays and 44.4% on the weekend, indicating a significantly higher rate on weekdays (*P* < .001). In addition, the percentage of respondents with a 100% response rate was 3.2%, whereas the percentage of those with no responses was 34.5%. The response of "presence of fever" was given by 184 participants, including family members of the respondent. Among these, data for 2 participants were given for the first time only when the symptom was present. The daily mean response rate was determined for different ages and genders. When those aged ≥60 years were excluded from the analysis, the lowest response rate was 22.6% for males aged ≤29 years, and the highest response rate was 74.9% for women aged 50–59 years (Figure 4).

In Internet Panel Survey II, the response rate was constant from the initiation of the survey to its end. The response rate was 44.0% on weekdays and 39.9% on the weekend. The percentage of respondents with a 100% response rate was 6%, and the percentage of those without a response was 36.2%.

In the TCCU survey, the response rate decreased gradually from the first to the final day. The response rate was 41.5% on weekdays and 38.9% on weekends. No significant differences were observed among the groups. The percentage of respondents with a 100% response rate was 3.3%, and the percentage of those without a response was 5.9%.

Figure 4. Population distribution by age in the surveys. Data are presented as the mean and standard deviation, which is indicated by error bars. Age is given in years.



Elapsed Time From the Development of Symptoms to a Report

In Internet Panel Survey I, the appearance of symptoms was as follows: >48 hours ago (59%), 6–24 hours ago (13%), 24–48 hours ago (12%), 3–6 hours ago (3%), 1–3 hours ago (1%), and <1 hour ago (1%). Of all responses, the daily average reporting rates by symptom were as follows: cough (8%), fever (3%), diarrhea (2%), vomiting (1%), rash (0%), and convulsion (0%).

Examples of Using the Data in Syndromic Surveillance

Figure 5 presents a graph for fever in Internet Panel Survey I.

The number of participants with fever was made a parameter. When the number of persons with fever was 3 or more standard deviations above the mean of the previous week, EARS provided an alert [15]. The number of those who reported febrile symptoms at history-taking in the outpatient section was made a parameter. The outpatient symptomatic surveillance reported 8 alerts, whereas the WDQH reported 16 alerts. Because of the time factor in an epidemic, the presence of an alert within the gold standard (3 days before to 3 days after symptom onset) was examined. The sensitivity was 0.43, and the specificity was 0.88. For cough, 8 alerts were reported during outpatient symptomatic surveillance. In Internet Panel Survey I, 19 alerts were reported. Similarly, for diarrhea, there were 30 alerts in outpatient symptomatic surveillance and 25 alerts in Internet Panel Survey I. For vomiting, there were 24 alerts in outpatient symptomatic surveillance and 22 alerts in Internet Panel Survey I. For rash, there was 1 alert in outpatient symptomatic surveillance and 10 alerts in Internet Panel Survey I. For convulsions, there was 1 alert in outpatient symptomatic surveillance and 7 alerts in Internet Panel Survey I.

Figure 6 shows the results of syndromic surveillance in Internet Panel Survey II: the number of participants with influenza in the area (published by the trend of symptom onset), those who reported cough, and those with fever. During the survey, 9 alerts each were reported for cough and fever.

Figure 5. Results of syndromic surveillance conducted in Internet Panel Survey I. Circles: alerts reported by early aberration reporting system (EARS). Diamonds: alert occurrence dates coincident with the regional outpatient symptomatic surveillance in medical institutions.

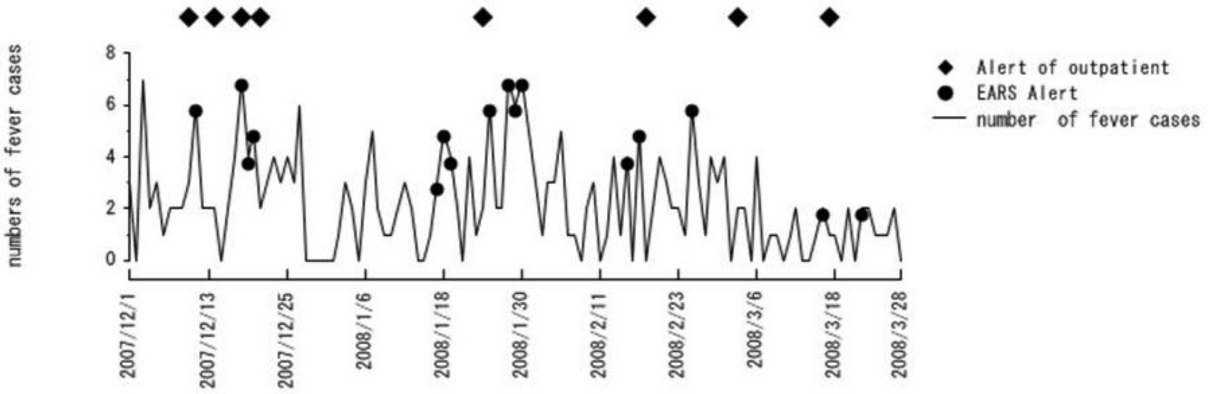
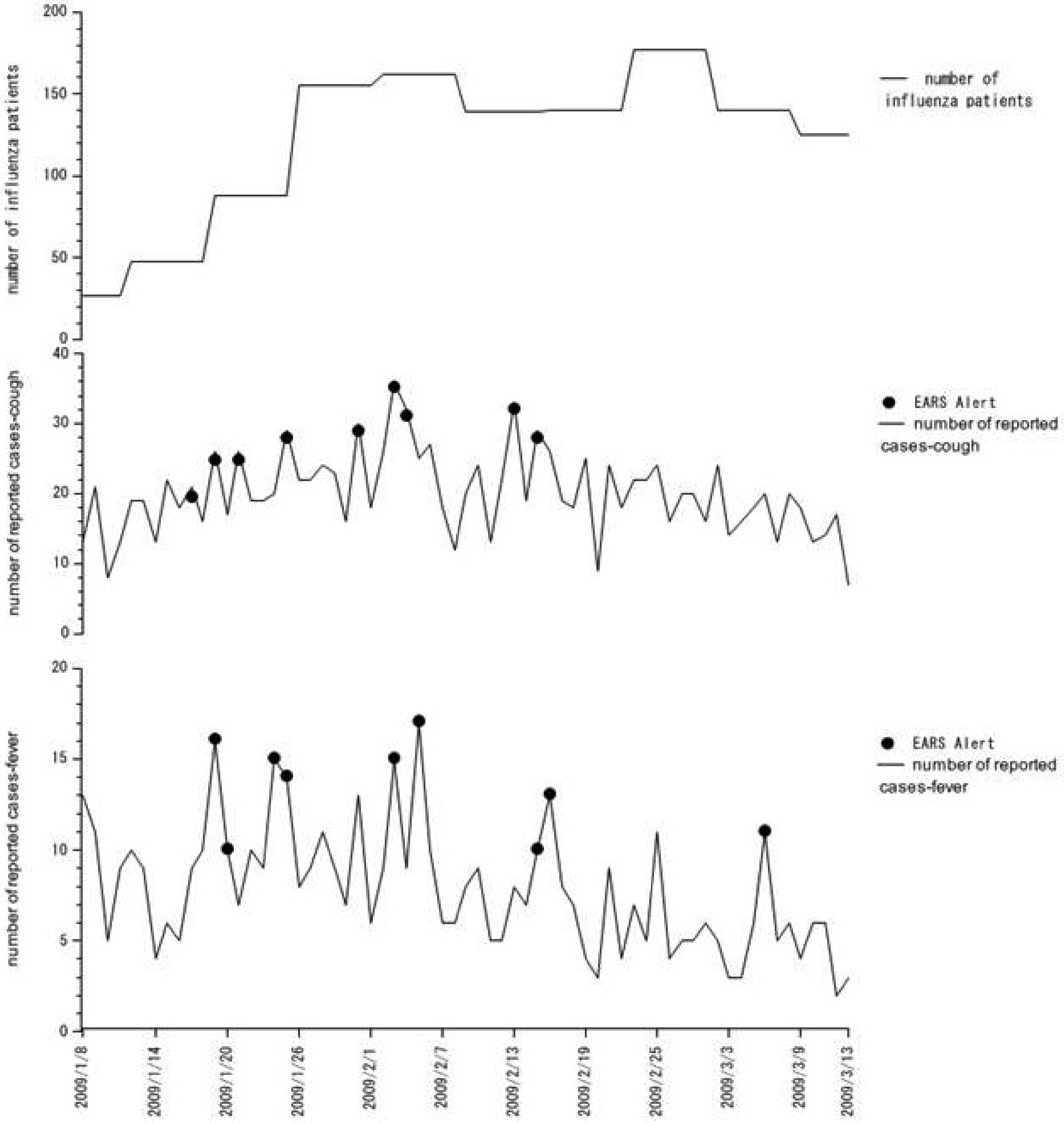


Figure 6. Results of syndromic surveillance conducted in Internet Panel Survey II. Circles: alerts reported by early aberration reporting system (EARS).



Examination of Methods to Reduce Survey Costs

Table 3 shows the fixed, variable, and total costs for the three surveys. Initial costs were very low only for screening the questions for panel research. Variable costs consisted of the investigation days and the number of investigation panels. In the TCCU survey, the fixed costs for development were the highest. Variable costs were only for incentives paid when participation was declared.

Table 3. Survey expenses for the three surveys

	Internet Panel Survey I	Internet Panel Survey II	TCCU[a] survey
Fixed cost			
Yen	20,000	20,000	2,457,000
US $	250	250	30,712.5
Variable cost			
Yen	8,260,000	2,480,000	43,000
US $	103,250	31.000	537.5
Total cost			
Yen	8,280,000	2,500,000	2,500,000
US $	103,500	31,250	31,250

[a] Tokyo Consumers' Co-operative Union.

There were 3 respondents in Internet Panel Survey II. All respondents gave their answers regarding symptoms that presented on the same day, providing responses on the day, the day after, and 2 days after the sentinel day. The response rates were 42.4%, 43.1%, and 42.7% in groups A, B, and C, respectively. No significant difference was found in the response rate among the groups.

The response rates by fee paid for a single response were as follows: 46.7% (40 yen), 39.7% (60 yen), and 41.6% (80 yen). A 1-way analysis of variance revealed a significant difference; thus, a multiple comparison test was conducted. Significant differences were observed in the average response rates between the 40-yen and 60-yen groups and between the 40-yen and 80-yen groups, with a greater response rate in the 40-yen group.

Discussion

We developed and validated a health observation method based on self-reporting by participants via the Internet. We clarified the usefulness of the WDQH by its practical use for syndromic surveillance.

Conventional paper-based surveys can be conducted at a low cost in a small population, and these surveys do not incur major initial expenses for the system. Moreover, combining Web- and paper-based surveys improves the response rate [17]. However, although requiring greater up-front costs, the WDQH allows daily inexpensive repetitive surveys to be conducted in a large number of participants, illustrating the advantage of a Web-based survey [18]. Furthermore, Web-based surveys permit a more efficient statistical analysis of data by computer. Thus, cost-effective and rapid surveys of a large number of participants, with high data precision, have become possible.

Previous studies have reported Internet surveys of asthma and diet, for example, in specific groups and patients [19,20]. Various countries have been using a method to understand an influenza epidemic in which residents voluntarily input information on influenza-like symptoms directly into a website. Thus, this method has been verified with an actual influenza epidemic and its usefulness has been demonstrated [10,13].

Although reports are available on the surveillance of symptoms in volunteers, no reports are available on the surveillance of symptoms of people identified by an Internet survey company. To promote the robustness of data gathering, the WDQH was conducted among registered members of an Internet survey company who were more likely than anonymous respondents to provide reliable data. Additionally, because the respondents were recruited from among registered members, only a short time was required from the decision making at the initiation of the survey to actual data collection. Thus, this survey provides value in this regard.

Rates of 52.6% [21] and 50% [22] have been reported in surveys that ended after a single investigation. In a meta-analysis conducted on 68 response rates of sampling surveys, the average response rate for Web-based surveys was 39.6% [23]. During surveillance of symptoms in volunteers, some participants who initially did not respond to the survey responded only when a symptom was present [13]. In our study, respondents to the Internet panel surveys who completed the questionnaire the first time tended to always cooperate with the survey. This finding indicates that these surveys are a useful method for reporting the appearance of symptoms.

There are problems with previous surveillance methods, such as the length of time required, indirect data collection, and no data collection during holidays. However, we developed the WDQH with the objective of acquiring data immediately after symptom onset. In addition, the WDQH allowed data collection on Saturdays, Sundays, and public holidays. Thus, we were able to conduct consistent daily surveillance. Furthermore, we used preventive measures, such as a branched and stepwise-structured questionnaire, to eliminate mistakes and discrepancies in responses [24].

Internet Panel Survey I confirmed that participants' health information could be collected daily via the Internet. However, the survey cost was 8.28 million yen (US $103,500), which was considered too expensive over a long period. We thus conducted both Internet Panel Survey II, which is economical for a panel survey, and the TCCU survey, without using the Internet survey company. In Internet Panel Survey II, two surveys were conducted. The first was used to reduce the frequency of surveys to once every 3 days. Changes in the actual number of participants with influenza corresponded with the changes determined by Internet Panel Survey II, which was conducted for syndromic surveillance without impairing data precision. This method allowed the implementation of a survey 3 times as long for the same cost as one conducted daily. The cost of Internet Panel Survey II was approximately one-half that of Internet Panel Survey I.

The second survey in Internet Panel Survey II investigated cash incentives. The Internet survey company that we used typically paid a fee of 60 yen for a single response. Surprisingly, the response rate was highest when the fee was set at the lowest level of 40 yen. Generally, higher fees act as an incentive for recruitment, but this study found that the offer of a higher reward did not result in a higher response rate. This point has been supported by a previous study [25]. As there was a sufficient response with no payment for each TCCU survey, any cost-associated restrictions on the survey period were eliminated. A fee was paid to the members of the TCCU only for survey registration, and information about the results was provided to the respondents. The response rate for the TCCU survey was lower than that for Internet panel surveys I and II. However, even at this lower rate, a large number of members were included because membership in the JCCU numbers at least 24 million throughout Japan, including 300,000 registered to its website.

From the WDQH data, we used EARS as an alert so that a level measured on the current day that was greater than 3 standard deviations different from the mean observed level for the previous week was reported as abnormal. If data are accumulated for several years, the number of participants can be estimated by multivariate analyses, where the number of participants, number of weeks, day of the week, holidays, and day after holidays are considered dummy variables. However, in this study, data were not continuously accumulated for 1 year or longer; thus, a multivariate analysis was not performed.

Syndromic surveillance could be implemented as a result of these validations. The Internet survey company in this study used an existing survey panel. The time required from planning to implementation was short with the use of this company, which already had its registered members as recruited participants. Thus, an urgent surveillance can be conducted within 3 days regardless of the location in Japan. During syndromic surveillance using the WDQH, measures against a health crisis can be readily put in place.

Removing selection bias is difficult in Internet surveys. The population tended to be biased toward young people because Internet surveys require respondents to have computer skills. Introducing an easy system to increase the response rate of older people could reduce this bias. However, because we believe that the increase or decrease in symptoms is reliable regardless of bias, we used EARS for all methods. Cough was often excluded from previous surveys because most cases of cough were present >48 hours before the survey, and cough probably requires a longer time to be recognized as bothersome to the same degree as other symptoms, such as fever and vomiting. To use the WDQH for syndromic surveillance, questions to respondents should be limited to those regarding acute symptoms, and a system that allows easy reporting within 24 hours should be established in the future.

In this study, we conducted the surveys with the same respondents. We think that it would be difficult for the respondents to maintain their interest every day for several years. For long-term operation of the survey, we consider that new respondents should be recruited after a certain period.

There are two further applications for the surveys other than the surveillance of symptoms. First, when environmental data published later by public institutions, such as average temperature and atmospheric pressure, are linked to the records by participant and locality on the same date, a cross-correlation survey of symptoms and environmental factors can be implemented. In the future, various daily surveys can be conducted, such as those for mean air temperature and the presence or absence of fever. These are topics to be investigated in the future.

Another application for the WDQH could be in postmarketing surveillance of food similar to that conducted for pharmaceuticals. Food safety is more widely expected by consumers today than it was in the past. To date, postmarketing surveillance of food has been conducted in only a single instance for a food additive [26]. A cross-correlation survey using a consumer database to identify the relationship between daily symptom data of the respondents obtained by the WDQH and consumed foods based on sales records may allow the reporting of adverse events when certain symptoms are associated with specific products.

Conclusions

We developed a health observation method via the Internet using self-reporting by respondents and validated the method for its application in syndromic surveillance. The Internet allows quick, cost-effective epidemiological surveys to be conducted that would be difficult to conduct by conventional methods.

Acknowledgments

Financial support for this study was provided by a grant from the Ministry of Health, Labour and Welfare, Japan.

Conflicts of Interest

None declared

Authors' Contributions

All authors contributed equally to this work.

Multimedia Appendix 1

CHERRIES checklist for Internet panel surveys I and II and the Tokyo Consumers' Co-operative Union (TCCU) survey.

[XLS File (Microsoft Excel File), 42KB - jmir_v13i3e66_app1.xls]

References

1. World Health Organization. International Health Regulations (2005). 2nd edition. Geneva: WHO Press; 2008.
2. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. J Am Med Inform Assoc 2004;11(2):141-150. [doi: 10.1197/jamia.M1356] [Medline: 14633933]
3. Henning KJ. What is syndromic surveillance? MMWR Morb Mortal Wkly Rep 2004 Sep 24;53 Suppl:5-11 [FREE Full text] [Medline: 15714620]
4. Buehler JW, Sonricker A, Paladini M, Soper P, Mostashari F. Syndromic surveillance practice in the United States: findings from a survey of state, territorial, and selected local health departments. Adv Dis Surveill 2008;6(3):1-16 [FREE Full text]
5. Bravata DM, McDonald KM, Smith WM, Rydzak C, Szeto H, Buckeridge DL, et al. Systematic review: surveillance systems for early detection of bioterrorism-related diseases. Ann Intern Med 2004 Jun 1;140(11):910-922. [Medline: 15172906]
6. Besculides M, Heffernan R, Mostashari F, Weiss D. Evaluation of school absenteeism data for early outbreak detection, New York City. BMC Public Health 2005;5:105 [FREE Full text] [doi: 10.1186/1471-2458-5-105] [Medline: 16212669]
7. Ohkusa Y, Shigematsu M, Taniguchi K, Okabe N. Experimental surveillance using data on sales of over-the-counter medications--Japan, November 2003-April 2004. MMWR Morb Mortal Wkly Rep 2005 Aug 26;54 Suppl:47-52 [FREE Full text] [Medline: 16177693]
8. Najmi AH, Magruder SF. An adaptive prediction and detection algorithm for multistream syndromic surveillance. BMC Med Inform Decis Mak 2005;5:33 [FREE Full text] [doi: 10.1186/1472-6947-5-33] [Medline: 16221308]
9. Bell DS, Kahn CE. Health status assessment via the World Wide Web. Proc AMIA Annu Fall Symp 1996:338-342. [Medline: 8947684]
10. Marquet RL, Bartelds AI, van Noort SP, Koppeschaar CE, Paget J, Schellevis FG, et al. Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003-2004 influenza season. BMC Public Health 2006;6:242 [FREE Full text] [doi: 10.1186/1471-2458-6-242] [Medline: 17018161]
11. van Noort SP, Muehlen M, Rebelo de Andrade H, Koppeschaar C, Lima Lourenço JM, Gomes MG. Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. Euro Surveill 2007 Jul;12(7):E5-E6. [Medline: 17991409]
12. Carlson SJ, Dalton CB, Tuyl FA, Durrheim DN, Fejsa J, Muscatello DJ, et al. Flutracking surveillance: comparing 2007 New South Wales results with laboratory confirmed influenza notifications. Commun Dis Intell 2009 Sep;33(3):323-327. [Medline: 20043603]
13. Tilston NL, Eames KT, Paolotti D, Ealden T, Edmunds WJ. Internet-based surveillance of Influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic. BMC Public Health 2010;10:650 [FREE Full text] [doi: 10.1186/1471-2458-10-650] [Medline: 20979640]
14. Sugiura H, Ohkusa Y, Akahane M, Sugahara T, Okabe N, Imamura T. Construction of syndromic surveillance using a web-based daily questionnaire for health and its application at the G8 Hokkaido Toyako Summit meeting. Epidemiol Infect 2010 Oct;138(10):1493-1502. [doi: 10.1017/S095026880999149X] [Medline: 20067657]
15. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). J Urban Health 2003 Jun;80(2 Suppl 1):i89-i96. [Medline: 12791783]
16. Centers for Disease Control and Prevention. CDC Emergency Risk Communication Branch (ERCB), Division of Emergency Operations (DEO), Office of Public Health Preparedness and Response (OPHPR). 2010 Jan 21. Early Aberration Reporting System URL: http://www.bt.cdc.gov/surveillance/ears/ [accessed 2011-09-03] [WebCite Cache ID 61QUALJLu]
17. Kroth PJ, McPherson L, Leverence R, Pace W, Daniels E, Rhyne RL, Prime Net Consortium. Combining web-based and mail surveys improves response rates: a PBRN study from PRIME Net. Ann Fam Med 2009;7(3):245-248 [FREE Full text] [doi: 10.1370/afm.944] [Medline: 19433842]
18. Schleyer TK, Forrest JL. Methods for the design and administration of web-based surveys. J Am Med Inform Assoc 2000;7(4):416-425. [Medline: 10887169]
19. Weisel CP, Weiss SH, Tasslimi A, Alimokhtari S, Belby K. Development of a Web-based questionnaire to collect exposure and symptom data in children and adolescents with asthma. Ann Allergy Asthma Immunol 2008 Feb;100(2):112-119. [doi: 10.1016/S1081-1206(10)60419-X] [Medline: 18320912]

20. Lu C, Pearson M, Renker S, Myerburg S, Farino C. A novel system for collecting longitudinal self-reported dietary consumption information: the internet data logger (iDL). J Expo Sci Environ Epidemiol 2006 Sep;16(5):427-433. [doi: 10.1038/sj.jes.7500479] [Medline: 16552428]
21. Greenlaw C, Brown-Welty S. A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. Eval Rev 2009 Oct;33(5):464-480. [doi: 10.1177/0193841X09340214] [Medline: 19605623]
22. Converse PD, Wolfe ED, Oswald FL. Response rates for mixed-mode surveys using mail and e-mail/web. American Journal of Evaluation 2008;29(1):99-107. [doi: 10.1177/1098214007313228]
23. Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or Internet-based surveys. Educ Psychol Meas 2000;60:821-836. [doi: 10.1177/00131640021970934]
24. Austin TM, Richter RR, Reinking MF. A primer on web surveys. J Allied Health 2008;37(3):180-186. [Medline: 18847116]
25. Kypri K, Gallagher SJ. Incentives to increase participation in an Internet survey of alcohol use: a controlled experiment. Alcohol Alcohol 2003;38(5):437-441 [FREE Full text] [Medline: 12915520]
26. Slough CL, Miday RK, Zorich NL, Jones JK. Postmarketing surveillance of new food ingredients: design and implementation of the program for the fat replacer olestra. Regul Toxicol Pharmacol 2001 Apr;33(2):218-223. [doi: 10.1006/rtph.2001.1461] [Medline: 11350204]

Abbreviations

EARS: early aberration reporting system
JCCU: Japanese Consumers’ Co-operative Union
TCCU: Tokyo Consumers’ Co-operative Union
WDQH: Web-based daily questionnaire for health

Edited by G Eysenbach; submitted 23.06.11; peer-reviewed by C Freifeld, N Tilston; comments to author 18.07.11; revised version received 13.08.11; accepted 25.08.11; published 23.09.11.

Please cite as:
Sugiura H, Ohkusa Y, Akahane M, Sano T, Okabe N, Imamura T
Development of a Web-Based Survey for Monitoring Daily Health and its Application in an Epidemiological Survey
J Med Internet Res 2011;13(3):e66
URL: http://www.jmir.org/2011/3/e66/
doi:10.2196/jmir.1872
PMID:21946004

Original Paper

A Comparison of a Postal Survey and Mixed-Mode Survey Using a Questionnaire on Patients' Experiences With Breast Care

Marloes Zuidgeest[1,2], MSc; Michelle Hendriks[2], PhD; Laura Koopman[2], PhD; Peter Spreeuwenberg[2], MA; Jany Rademakers[2], PhD

[1]Tranzo, Academic Research Centre for Health and Social Care, Tilburg University, Tilburg, Netherlands

[2]Netherlands Institute for Health Services Research (NIVEL), Utrecht, Netherlands

Corresponding Author:
Marloes Zuidgeest, MSc
Tranzo, Academic Research Centre for Health and Social Care
Tilburg University
PO Box 90153
Tilburg, 5000 LE
Netherlands
Phone: 31 134663627
Fax: 31 134663637
Email: m.zuidgeest@uvt.nl

Abstract

Background: The Internet is increasingly considered to be an efficient medium for assessing the quality of health care seen from the patients' perspective. Potential benefits of Internet surveys such as time efficiency, reduced effort, and lower costs should be balanced against potential weaknesses such as low response rates and accessibility for only a subset of potential participants. Combining an Internet questionnaire with a traditional paper follow-up questionnaire (mixed-mode survey) can possibly compensate for these weaknesses and provide an alternative to a postal survey.

Objective: To examine whether there are differences between a mixed-mode survey and a postal survey in terms of respondent characteristics, response rate and time, quality of data, costs, and global ratings of health care or health care providers (general practitioner, hospital care in the diagnostic phase, surgeon, nurses, radiotherapy, chemotherapy, and hospital care in general).

Methods: Differences between the two surveys were examined in a sample of breast care patients using the Consumer Quality Index Breast Care questionnaire. We selected 800 breast care patients from the reimbursement files of Dutch health insurance companies. We asked 400 patients to fill out the questionnaire online followed by a paper reminder (mixed-mode survey) and 400 patients, matched by age and gender, received the questionnaire by mail only (postal survey). Both groups received three reminders.

Results: The respondents to the two surveys did not differ in age, gender, level of education, or self-reported physical and psychological health (all *Ps* > .05). In the postal survey, the questionnaires were returned 20 days earlier than in the mixed-mode survey (median 12 and 32 days, respectively; $P < .001$), whereas the response rate did not differ significantly (256/400, 64.0% versus 242/400, 60.5%, respectively; $P = .30$). The costs were lower for the mixed-mode survey (€2 per questionnaire). Moreover, there were fewer missing items (3.4% versus 4.4%, $P = .002$) and fewer invalid answers (3.2% versus 6.2%, $P < .001$) in the mixed-mode survey than in the postal survey. The answers of the two respondent groups on the global ratings did not differ. Within the mixed-mode survey, 52.9% (128/242) of the respondents filled out the questionnaire online. Respondents who filled out the questionnaire online were significantly younger ($P < .001$), were more often highly educated ($P = .002$), and reported better psychological health ($P = .02$) than respondents who filled out the paper questionnaire. Respondents to the paper questionnaire rated the nurses significantly more positively than respondents to the online questionnaire (score 9.2 versus 8.4, respectively; $\chi^2_1 = 5.6$).

Conclusions: Mixed-mode surveys are an alternative method to postal surveys that yield comparable response rates and groups of respondents, at lower costs. Moreover, quality of health care was not rated differently by respondents to the mixed-mode or postal survey. Researchers should consider using mixed-mode surveys instead of postal surveys, especially when investigating younger or more highly educated populations.

(J Med Internet Res 2011;13(3):e68) doi:10.2196/jmir.1241

KEYWORDS

Data collection; health care quality; consumer satisfaction; breast cancer; patient preferences; health care quality indicator

Introduction

In the Netherlands, health care policy stresses regulated competition between health care providers [1]. Efforts are made to enhance the transparency of health care quality, to stimulate informed decision making among consumers, and to improve the performance of health care providers. Comparative information about the performance of health care providers is needed for consumers to make informed decisions. This comparative information can be gathered in different ways. One possibility is to ask a sample of patients about their actual experiences concerning quality of care provided by health care providers.

Measuring the quality of care from the patients' perspective has been standardized in the Netherlands since 2006, using a new instrument called the Consumer Quality Index (CQ-index or CQI) [2]. CQI questionnaires are usually self-administered paper questionnaires (eg, CQI Rheumatoid Arthritis [3], CQI Breast Care [4]). Individual structured interviews are conducted in cases where a self-administered paper questionnaire is not feasible because of respondents' visual, physical, or cognitive limitations (CQI Care for the Disabled [5], CQI Long-Term Care [6]). Postal surveys (with multiple reminders) and interviews are relatively expensive and time consuming. It would therefore be interesting to know whether other data collection methods can be applied in this field.

The Internet is increasingly considered to be an efficient medium for assessing quality of care from a patient's perspective. In populations that already use the Internet, Internet surveys have been found to be a useful means of conducting research [7-9]. Efficiency gains are found in shorter response times and field costs reductions (50%–80%) [10-12]. In contrast to paper questionnaires, Internet questionnaires can contain various interactive features that allow complex skip patterns that are invisible to respondents, and the Internet allows validation of responses by using an instant feedback function while respondents are still online [12,13]. Consequently, the quality of data collected with an Internet survey is higher. Some Internet surveys have shown promising response rates (up to 94% in Web forums) [10,12,14]. The extreme response in Web forums can be explained by a probable selection bias in these studies. Those who participate in Web forums are most likely people who are familiar with the Internet and frequently use the Internet, leading to a higher response rate to Internet questionnaires. This high response rate has not been realized in other studies; response rates ranged from 17% to 70% [15]. In CQI research, the response rate to paper questionnaires varied from 20% to 79% with an average response rate of 55% [16]. One CQI study compared an Internet questionnaire with a paper questionnaire. The response rate to the Internet questionnaire (8%) was considerably lower than to the paper questionnaire (35%) [17,18]. To increase the response rate one can send a prenotification or reminders, give an incentive, or use short questionnaires. A salient subject of a questionnaire also increases the response rate [19].

The potential of Internet surveys should, however, be balanced against an equally large weakness. The Netherlands has the largest percentage of households with Internet access in the European Union, but there are still 1.2 million Dutch people (7.3% of the population) with no Internet access at home and 0.5 million Dutch people (3.1% of the population) who do not use the Internet [20]. People who use the Internet are more affluent, better educated, more often male, and younger than people who do not use Internet. Only a part of the population can thus be reached through the Internet [10,11,21]. To compensate for the selection of people in an Internet survey, a combination of data collection methods can be used such as combining an Internet questionnaire with a more traditional postal follow-up [19].

It is known that the way questionnaire are administered has an effect on answers of respondents (so-called mode effects). For example, telephone respondents were found to be more likely to rate health care positively and their own health status negatively than postal respondents [22,23]. This finding is similar to a study where telephone respondents provide more positive ratings than Web respondents [24]. Another example is that students who completed a Web-based questionnaire responded more favorably on different scales (such as college challenge and learning, education, and personal and social gains) than students who filled out a paper questionnaire [25]. It is suggested that computer anxiety affects participants' responses. Moreover, biases could occur in the way people perceive and process questions presented on screen versus on paper. A study that tested the difference in test–retest reliability and internal consistency between Internet and paper versions of the SF-36, however, found little or no evidence for mode effects [26]. Knowing that these mode effects exist, it is important to investigate whether the answers of respondents in a postal and mixed-mode survey differ.

To examine whether a mixed-mode survey can be an alternative to postal survey, our research question is "What are the differences between a mixed-mode survey (Internet with paper follow-up) and a more traditional postal survey in terms of respondent characteristics, response rates and time, quality of data, costs, and mode effects?" The differences were examined within a sample of breast care patients who reported their experiences with health care using the CQI Breast Care questionnaire.

Methods

Sample

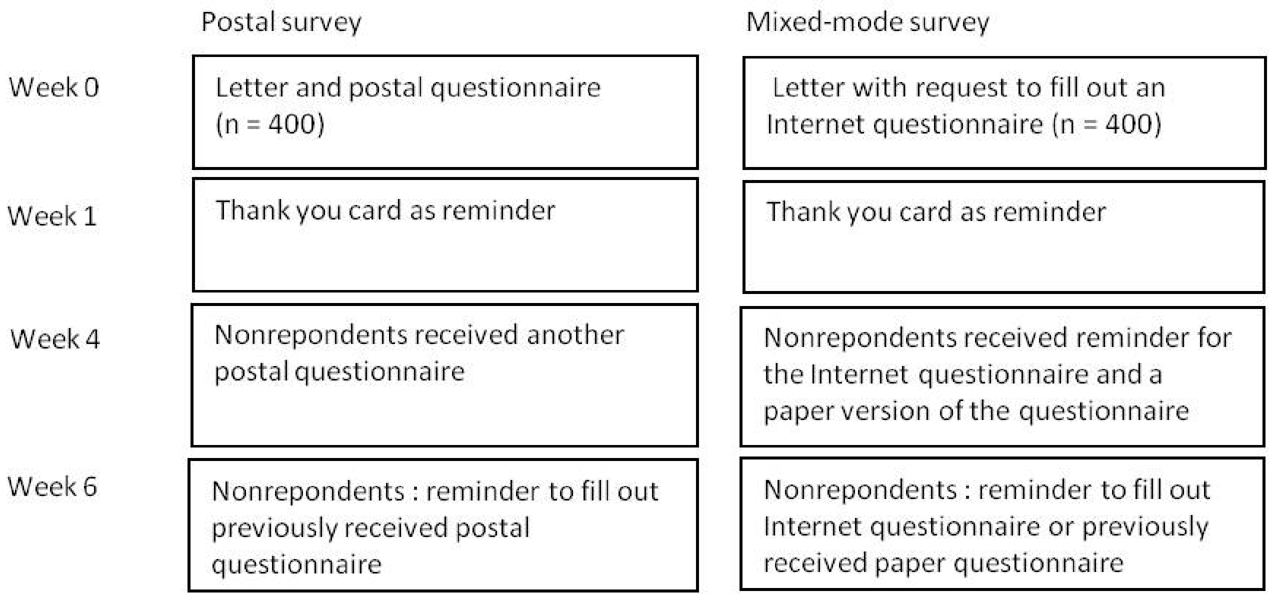
Data were collected within a larger study assessing the usability of CQI Breast Care [27]. For the mixed-mode survey, 200 patients with a benign abnormality and 200 patients with breast cancer were selected from the reimbursement files of seven

Dutch health insurance companies. Inclusion criteria were (1) being older than 18 years and (2) having received breast care in the last 24 months. We used the same procedure to select 3955 patients who received the questionnaire by mail only as part of another study. Of these 3955 patients, we selected 400 patients (200 with breast cancer and 200 with benign abnormalities) for the comparison of the two surveys. These 400 patients were not randomly selected, but were matched by age and gender to the respondents in the mixed-mode survey.

Data Collection

Patients received a letter from their health insurance company with the request to fill out a paper questionnaire (postal survey) or an Internet questionnaire with unique username and password (mixed-mode survey). A total of three reminders were sent and in both surveys nonrespondents received a paper version of the questionnaire in the third mail-shot. This data protocol was based on Dillman et al [28]. (See Figure 1 for detailed information on the mail-shots.) The data were collected in the Netherlands in the spring of 2008.

Figure 1. Mail-shots sent to the patients.



	Postal survey	Mixed-mode survey
Week 0	Letter and postal questionnaire (n = 400)	Letter with request to fill out an Internet questionnaire (n = 400)
Week 1	Thank you card as reminder	Thank you card as reminder
Week 4	Nonrepondents received another postal questionnaire	Nonrepondents received reminder for the Internet questionnaire and a paper version of the questionnaire
Week 6	Nonrepondents : reminder to fill out previously received postal questionnaire	Nonrepondents : reminder to fill out Internet questionnaire or previously received paper questionnaire

Questionnaire

The CQI Breast Care contains items measuring the actual experiences of patients with breast examinations, surgery for breast cancer, other treatment, subsequent treatment, cooperation between health care providers, continuity of care, accessibility of care, and expertise of health care providers [4]. There are two versions of the CQ-index: one for patients with breast cancer (151 items) and one for patients with a benign abnormality (60 items). The questionnaire for patients with a benign abnormality is the same as the questionnaire for breast cancer, except that it does not contain questions about surgery and treatments. Both questionnaires have three scales in common, and the questionnaire for patients with breast cancer consists of 11 extra scales. Cronbach alpha for these scales varied between 0.74 and 0.93. Example items are presented in Table 1. The questionnaires additionally contain items on respondents characteristics (eg, age, education, ethnicity, and patient’s self-assessed physical and psychological health) and global ratings of health care providers general practitioner, hospital care in the diagnostic phase, surgeon, nurses, radiotherapy, chemotherapy, and hospital care in general). In the present study, we focused on the global ratings of the health care providers. These ratings ranged from 0 to 10, with a score of 0 indicating the worst possible health care or provider and a score of 10 indicating the best possible health care or provider. The respondents were asked to report their experiences in the last 24 months.

Table 1. Scales in the Consumer Quality Index Breast Care, their reliability (Cronbach alpha for internal consistency), and example items

Scale		Number of items	$Alpha_1$[a]	$Alpha_2$[b]	Total	Example of item
1	Conduct of professionals during breast examination	7	.90	.91	.90	How often did caregivers listen to you carefully?
2	Conduct of general practitioner	4	.91	.88	.90	How often did your general practitioner take you seriously?
3	Conduct of nurses	5	–	.87	.87	How often did nurses pay personal attention to you?
4	Conduct of surgeon	4	–	.89	.89	How often did the surgeon spend enough time with you?
5	Autonomy regarding treatment	4	–	.82	.82	How often did you get the chance to decide about your treatment?
6	Autonomy regarding follow-up treatment	2	–	.93	.93	How often were your specific wishes about follow-up treatment taken into account?
7	Conduct of professionals during radiotherapy	5	–	.88	.88	How often did you get the opportunity to ask questions about radiotherapy?
8	Information about radiotherapy	2	–	.78	.78	How often did you get enough information about radiotherapy?
9	Conduct of professionals during chemotherapy	4	–	.92	.92	How often did caregivers listen carefully to you?
10	Information about chemotherapy	4	–	.80	.80	How often did caregivers explain aspects of chemotherapy in a way that was easy to understand?
11	Cooperation	5	.91	.87	.89	How often did caregivers make good arrangements with each other?
12	Continuity of psychosocial care	3	–	.84	.84	Were you informed about the options for psychosocial care?
13	Continuity of physiotherapy	3	–	.74	.74	Were you assisted with a referral to physiotherapy?
14	Continuity of rehabilitation	3	–	.80	.80	Did you have as rapid access to a rehabilitation program as you wanted?

[a] Questionnaire for patients with benign abnormality.

[b] Questionnaire for patients with breast cancer.

Statistical Analyses

Respondent Characteristics

To check whether our matching procedure was successful, we compared the selected patients within the two surveys on age and gender. Respondents were compared concerning age, level of education, self-reported physical and psychological health (Mann-Whitney test), and gender (χ^2 test).

Response Rate and Time

Response rates were calculated as the number of valid received questionnaires divided by the number of patients in the starting sample. The response time was calculated as the number of days between the first letter (January 31, 2008) and the return date of the valid questionnaire. For the mixed-mode survey, the number of days between sending the paper questionnaire (February 28, 2008) and receiving the valid paper questionnaire was also calculated. The closing date of the data collection was April 1, 2008. A chi-square test was used to examine the differences in response rates between the two surveys because of the dichotomous variable (respondent/ nonrespondent). The differences in response time were determined using a Mann-Whitney test because the response time is a continuous variable.

Quality of Data

The percentage of items that were skipped while they needed to be answered (missing items) and the percentage of the items that were answered while they needed to be skipped (invalid answers) were calculated. The percentages were compared between the two surveys using a Mann-Whitney test because these percentages are continuous variables.

Total Costs

Expenses considered in cost calculations included setup costs (document layout, programming and testing of the questionnaire for each survey, and mailing supplies), field costs (postage, technical support, and project management staff), and scanning data costs (data entry of paper questionnaires). The costs per valid questionnaire received were calculated by dividing the total costs by the number of valid questionnaires received.

Mode Effects

We performed multilevel regression analyses to examine the mode effects. Multilevel regression analyses take into account the hierarchical structure of our data: individual patients (level 1) are nested within hospitals (level 2). The analyses were conducted using MLwiN version 2.02 software package (Centre for Multilevel Modelling, University of Bristol, Bristol, UK). Mode effects were examined by comparing the estimated mean scores on seven global ratings of general practitioner, hospital care in the diagnostic phase, surgeon, nurses, radiotherapy, chemotherapy, and hospital care in general using a chi-square test ($P < .05$ if $\chi^2 > 3.8$ and $P < .001$ if $\chi^2 > 6.6$). The mean scores were adjusted for the influence of age, education level, and self-reported health status of respondents.

In addition, within the mixed-mode survey, we examined the differences in respondent characteristics, differences in response rates, and time and mode effects for respondents who filled out the Internet questionnaire and the paper questionnaire.

Results

Respondent Characteristics

Characteristics of the sample are presented in Table 2. Our matching procedure was successful since age and gender of the selected patients did not differ between the postal and mixed-mode survey. Patients with benign abnormalities were younger then patients with breast cancer ($P < .001$).

Table 2. Sample characteristics

	Postal survey	Mixed-mode survey	Mean difference	95% CI	*P* value
Overall (n)	400	400			
Mean age (SD) years	55.5 (14.5)	55.5 (14.8)	–0.1	–2.1 to 1.9	.93
Female	97.3% (389/400)	97.3% (389/400)	0.0	0.0	1.00
Breast cancer (n)	200	200			
Mean age (SD) years	61.3 (12.7)	61.8 (12.9)	–0.4	–2.9 to 2.1	.77
Female	99.0% (198/200)	99.0% (198/200)	0.0	0.0	1.00
Benign abnormalities (n)	200	200			
Mean age (SD) years	49.5 (13.7)	49.3 (14.0)	0.2	–2.5 to 2.9	.89
Female	95.5% (191/200)	95.5% (191/200)	0.0	0.0	1.00

Table 3 shows that also the characteristics of the respondents did not differ between the postal and mixed-mode survey.

Table 3. Respondents' age, gender, level of education, and self-reported physical and psychological health

	Postal survey	Mixed-mode survey	Mean difference	95% CI	*P* value
Overall (n)	256	242			
Mean age (SD), years	55.8 (13.5)	57.0 (13.6)	–1.2	–3.6 to 1.2	.32
Female	97.7% (250/256)	97.5% (236/242)	0.1	–2.6 to 2.8	1.00
Breast cancer (n)	134	132			
Mean age (SD), years	60.2 (12.4)	62.1 (12.4)	–1.9	–5.0 to 1.0	.26
Female	98.5% (132/134)	99.2% (131/132)	0.7	–3.3 to 3.3	1.00
Benign abnormalities (n)	122	110			
Mean age (SD), years	50.8 (13.1)	50.8 (12.4)	0.1	–3.2 to 3.4	.29
Female	96.7% (118/122)	95.5% (105/110)	1.2	–3.7 to 6.2	.74
Education level (n)	251	232			
Mean (SD)	4.4 (1.9)	4.6 (1.8)	–0.2	–0.6 to 0.1	.09
Less than high school	41.1% (103/251)	31.0% (72/232)			
High school graduated	20.3% (51/251)	25.0% (58/232)			
Higher education	31.6% (79/251)	39.3% (91/232)			
College degree	4.8% (12/251)	2.2% (5/232)			
Other	2.4% (6/251)	2.6% (6/232)			
Self-reported physical health (n)	254	239			
Mean (SD)	2.9 (0.8)	2.8 (0.9)	0.1	–0.04 to 0.3	.29
Excellent	5.1% (13/254)	11.3% (27/239)			
Very good	20.1% (51/254)	16.7% (40/239)			
Good	55.5% (141/254)	55.6% (133/239)			
Fair	17.3% (44/254)	13.4% (32/239)			
Poor	2.0% (5/254)	2.9% (7/239)			
Self-reported psychological health (n)	255	239			
Mean (SD)	2.6 (1.0)	2.6 (1.0)	0.1	–0.1 to 0.2	.40
Excellent	16.9% (43/255)	18.4% (44/239)			
Very good	18.4% (47/255)	23.0% (55/239)			
Good	51.8% (129/255)	44.4% (106/239)			
Fair	11.4% (32/255)	13.4% (32/239)			
Poor	1.6% (4/255)	0.8% (7/239)			

Within the mixed-mode survey, differences were found between those who filled out the Internet questionnaire and those who filled out the paper questionnaire. Internet respondents were younger, were more often highly educated, and reported better psychological health compared with respondents who filled out the paper questionnaire (Table 4). Also, both paper and Internet respondents with benign abnormalities were younger than their counterparts with breast cancer (*Ps* < .001; not in table).

Table 4. Respondents' characteristics within the mixed-mode survey: age, gender, level of education, and self-reported physical and psychological health

	Postal	Internet	Mean difference	95% CI	*P* value
Overall (n)	114	128			
Mean age (SD), years	61.8 (14.0)	52.7 (11.6)	–9.1	–12.3 to –5.8	<.001
Female	99.1% (113/114)	96.1% (123/128)	–3.0	–6.9 to 0.9	.22
Breast cancer (n)	68	64			
Mean age (SD), years	67.9 (11.2)	56.0 (10.5)	–12.0	–15.7 to 8.2	1.00
Female	98.5% (67/68)	100% (64/64)	1.5	–1.5 to 4.5	1.00
Benign abnormalities (n)	46	64			
Mean age (SD), years	52.7 (12.9)	49.5 (11.9)	–3.2	–7.9 to 1.5	.18
Female	100% (46/46)	92.2 % (59/64)	–7.8	–15.7 to 1.0	.07
Education level (n)	105	127			
Mean (SD)	4.2 (1.8)	4.9 (1.8)	0.7	0.3 to 1.2	.002
Less than high school	38.1% (40/105)	25.2% (32/127)			
High school graduated	41.0% (43/105)	38.6% (49/127)			
Higher education	18.1% (19/105)	29.9% (38/127)			
University degree	0.0% (0/105)	3.9% (5/12)			
Other	2.9% (1/105)	2.4% (3/127)			
Self-reported physical health (n)	112	127			
Mean (SD)	2.9 (0.9)	2.7 (0.9)	–0.1	–0.4 to 0.1	.14
Excellent	12.5% (14/112)	10.2% (13/127)			
Very good	11.6% (13/112)	21.3% (27/127)			
Good	55.4% (62/112)	55.9% (71/127)			
Fair	17.9% (20/112)	9.4% (12/127)			
Poor	2.7% (3/112)	3.1% (4/127)			
Self-reported psychological health (n)	112	127			
Mean (SD)	2.7 (1.0)	2.4 (0.9)	–0.3	–0.5 to 0.03	.02
Excellent	17.0% (19/112)	19.7% (25/127)			
Very good	17.0% (19/112)	28.3% (36/127)			
Good	46.4% (52/112)	42.5% (54/127)			
Fair	18.8% (21/112)	8.7% (11/127)			
Poor	0.9% (1/112)	0.8% (1/127)			

Response Rates and Times

The response rate did not differ between the two surveys and was 64.0% (256/400 patients) for the postal survey and 60.5% (242/400 patients) for the mixed-mode survey (*P* = .31; Table 5). While the response rates of patients with breast cancer and of patients with benign abnormalities did not differ in the postal survey (134/200, 67.0% versus 122/200, 61.0%, respectively; *P* = .21), the response rate was significantly higher for patients with breast cancer than for patients with benign abnormalities in the mixed-mode survey (132/200, 66.0% versus 110/200, 55.0%, respectively; *P* = .02).

In the mixed-mode survey, 52.9% (128/242) of the respondents filled out the questionnaire online. The percentage of patients with benign abnormalities who filled out the questionnaire online was higher (64/110, 58%) than the percentage of patients with breast cancer (64/134, 49%). However, this difference was not significant (*P*= .13).

Table 5. Response rates for each survey and for patients with breast cancer or benign abnormalities

	Postal survey		Mixed-mode survey		Mean difference	95% CI	*P* value
Overall response	64.0%	256/400	60.5%	242/400	3.4%	–3.2% to 10.2%	.32
Breast cancer	67.0%	134/200	66.0%	132/200	1.0%	–8.3% to 10.3%	.83
Benign abnormality	61.0%	122/200	55.0%	110/200	6.0%	–3.7% to 15.7%	.23

Figure 2 and Figure 3 show the cumulative percentage of questionnaires received by days after the first mail-shot. The vertical lines in the graphs represent the reminders that were sent. In the postal survey, questionnaires were returned 20 days earlier than in the mixed-mode survey ($z = –3.59$, $P < .001$). The median number of days expired before the questionnaire was returned was 12 days (range 4–60 days) in the postal survey and 32 days (range 2–61 days) in the mixed-mode survey.

In the mixed-mode survey, the paper questionnaires were sent in week 4 (second reminder). The median number of days expired before these paper questionnaires were returned was 7 days (range 4–33 days). The median number of days expired before online questionnaires were filled out was 9 days (range 2–59 days). In other words, the longer response time in the mixed-mode survey was mainly caused by the group who did not respond using the Internet.

Figure 2. Percentage of received questionnaires by days after first mail-shot for the postal and mixed-mode surveys.

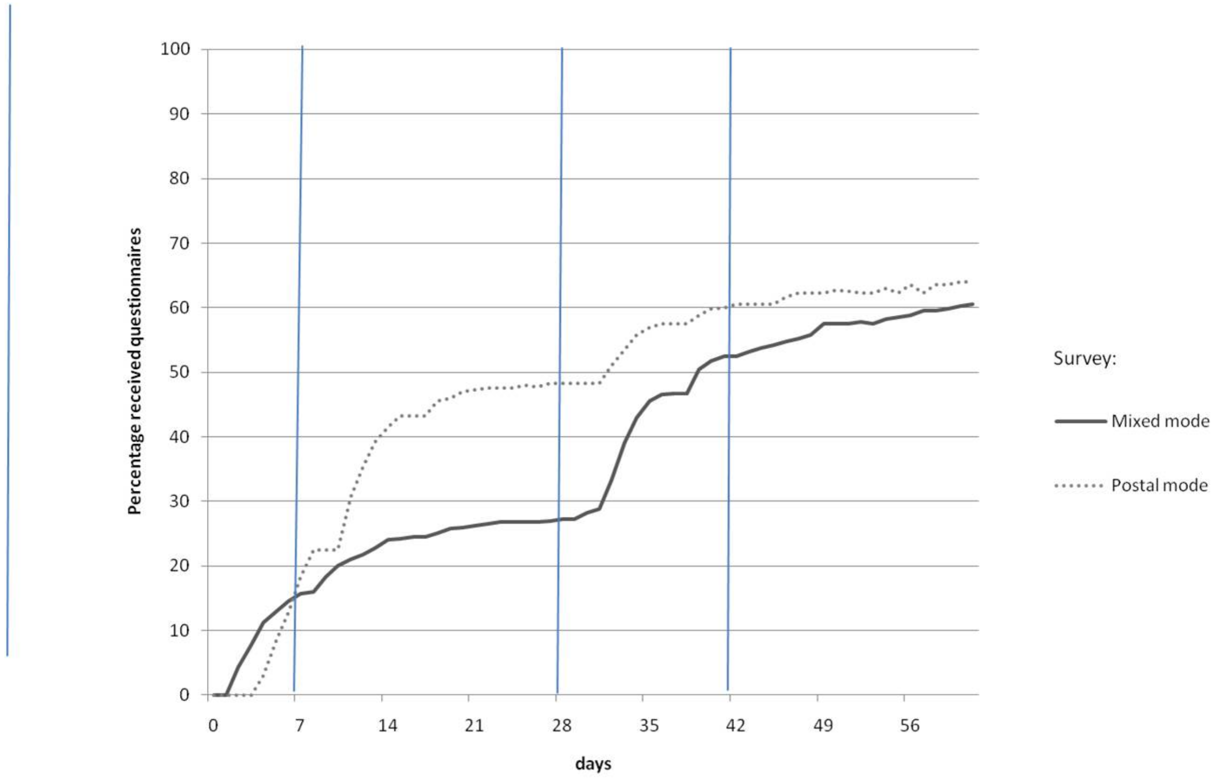
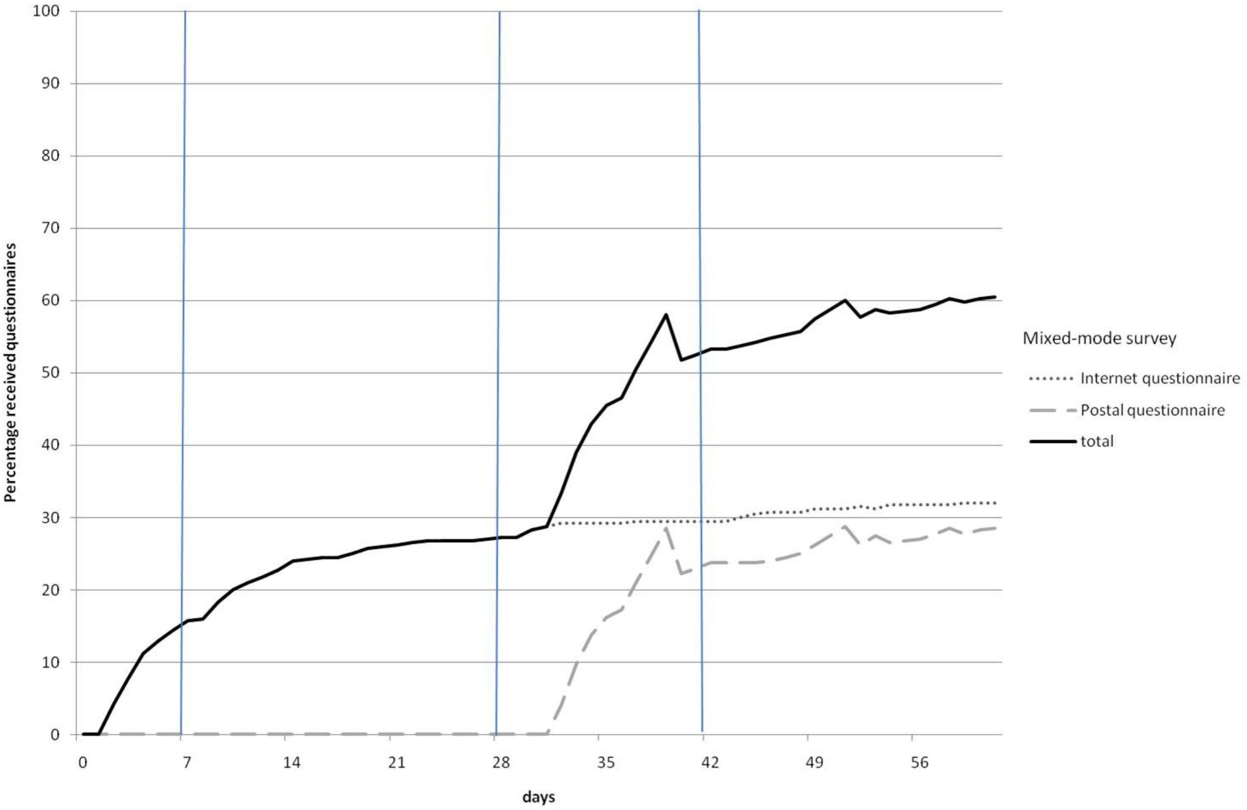


Figure 3. Percentage of received questionnaires by days after first mail-shot for the Internet and paper questionnaires within the mixed-mode survey.



Quality of Data

The mean percentage of missing items per question differed significantly between the two surveys ($z = –3.08$, $P = .002$): the mean percentage of missing items was lower in the mixed-mode survey than in the postal survey (5.04/150, 3.4% versus 6.60/150, 4.4%, respectively). In addition, the mean percentage of invalid answers was twice as high in the postal survey as in the mixed-mode survey (4.99/81, 6.2% versus 2.50/81, 3.2%, respectively; $z = –3.68$, $P < .001$).

Costs

The costs per valid questionnaire returned were higher in the postal survey than in the mixed-mode survey (€25.8 versus €23.9 per valid questionnaire returned, respectively). Compared with the postal survey, the variable costs were reduced by 17% of the total costs in the mixed-mode survey, but the fixed costs were raised by 17% (Table 6).

Table 6. Fixed and variable costs per valid questionnaire returned

Costs	Postal survey		Mixed-mode survey	
%		Cost (€)	%	Cost (€)
Fixed costs	41.5	10.7	58.4	14.0
General	3.6	0.9	2.1	0.5
Information technology (programming software, scanning, Internet questionnaire design)	35.0	9.0	53.8	12.9
Processing results and making a data file	2.9	0.8	2.5	0.6
Variable costs	58.5	15.1	41.6	9.9
Material (paper, envelopes)	1.4	0.4	0.9	0.2
Printing (letters, survey, reminders)	26.3	6.8	21.6	5.2
Preparing tasks (folding forms/questionnaires, thank you cards, filling envelopes)	4.2	1.1	4.6	1.1
Response processing (opening envelope, checking, scanning data)	8.6	2.2	4.6	1.1
Postal costs	17.9	4.6	9.9	2.4
Total	100.0	25.8	100.0	23.9

Mode Effects

In Table 7, the mean scores on seven global ratings of different health care providers are presented. These mean scores have been corrected for hospital, age, level of education, and self-reported health status. The scores are relatively high, ranging from 8.3 to 9.0. Respondents in the postal survey gave the radiotherapist a score of 9.0 and the total care in the hospital a score of 8.3. The respondents in the mixed-mode survey rated the general practitioner and chemotherapy care the highest (score = 8.8) and gave care at the hospital in the diagnosis phase and hospital care a score of 8.4. We found no significant differences in global ratings between the two surveys.

Table 7. Mean scores on global ratings of different health care providers (corrected for hospital, age, education, and self-reported health status) for respondents to the postal survey and mixed-mode survey

		Postal survey			Mixed-mode survey			
Ratings of health care providers		n	Mean[a]	SE	N	Mean[a]	SE	χ^2_1
1	General practitioner	105	8.5	0.21	114	8.8	0.22	2.1
2	Hospital care in diagnostic phase	240	8.4	0.11	220	8.4	0.11	0.0
3	Surgeon[b]	85	8.9	0.19	102	8.5	0.20	2.6
4	Nurses[b]	75	8.7	0.19	77	8.7	0.20	0.1
5	Radiotherapy[b]	68	9.0	0.19	80	8.7	0.22	2.1
6	Chemotherapy[b]	41	8.9	0.24	50	8.8	0.23	0.2
7	Hospital care in general	239	8.3	0.11	222	8.4	0.14	0.5

[a] Measured on an 11-point scale from 0 (worst possible) to 10 (best possible).

[b] Only in breast cancer questionnaire.

Table 8 shows the differences in global ratings given by respondents to the paper and Internet questionnaires within the mixed-mode survey. The global rating of nurses differed significantly between these two groups: respondents filling out the paper questionnaire rated the nurses significantly more positively than respondents filling out the questionnaire online (score 9.2 versus 8.4, respectively; $c^2 > 3.8$).

Table 8. Mean scores on global ratings of different health care providers (corrected for hospital, age, education, and self-reported health status) for respondents to the postal or Internet questionnaire within the mixed-mode survey

Ratings of health care providers		Paper questionnaire			Internet questionnaire			
		n	Mean[a]	SE	n	Mean[a]	SE	χ^2_1
1	General practitioner	48	8.7	0.25	66	8.7	0.23	0.0
2	Hospital care during diagnosis phase	96	8.4	0.16	124	8.3	0.16	0.3
3	Surgeon[b]	49	8.7	0.26	56	8.2	0.25	3.0
4	Nurses[b]	35	9.2	0.33	42	8.4	0.28	5.6
5	Radiotherapy[b]	38	8.7	0.28	43	8.5	0.26	0.3
6	Chemotherapy[b]	21	9.0	0.34	30	8.7	0.28	1.4
7	Care at hospital	98	8.4	0.16	124	8.3	0.16	0.3

[a] Measured on an 11-point scale from 0 (worst possible) to 10 (best possible).

[b] Only in breast cancer questionnaire.

Discussion

This study examined whether a mixed-mode survey (Internet questionnaire with paper follow-up) is an alternative to the more traditional postal survey. The results showed that combining an Internet questionnaire with a paper follow-up improved the quality of data and was less expensive than a postal survey. However, the time before questionnaires were received was longer in the mixed-mode survey. No differences between the mixed-mode survey and postal survey were found concerning respondent characteristics, response rates, and global ratings of different health care providers.

The findings showed that the characteristics of the respondents were the same for the two surveys. This means that mixed-mode surveys attract the same population as postal surveys. In total, 53% of respondents in the mixed-mode survey (128/242) filled out the questionnaire online. It appeared that in the mixed-mode survey Internet respondents were younger and more often highly educated and that they reported better psychological health than paper respondents. The younger people probably were more familiar with the Internet and were more likely to have access to the Internet than older people [11,21]. To overcome the problem of possible exclusion of the elderly and less highly educated, a mixed-mode survey should be chosen rather than an Internet survey [11,29].

The response rate was relatively high for both surveys (over 60%). In other CQI surveys, the response rates varied between 20% and 79% [16]. Perhaps the relatively high response rate is due the subject under study, namely abnormality of the breast. The response rate among women referred for mammography in another study was comparably high, both for the Internet (64%) and for the paper questionnaire (77%) [26]. Breast abnormality is a disease that has a huge impact on the emotional and physical quality of life of patients [30]. A review showed that saliency of the subject of questionnaires yields higher response rates [19]. Our results confirm the result of that review. In the mixed-mode survey, the response rate for patients with breast cancer was higher than the response rate for patients with benign abnormalities, even though the questionnaire for breast cancer was longer.

The response time for the questionnaires to be returned was longer in the mixed-mode survey than in the postal survey. This effect was unexpected because using the Internet can reduce the time taken to return a questionnaire [10-12]. Both groups in the mixed-mode survey (paper and Internet respondents) responded relatively quickly (median number of days 7 and 9 days, respectively), but respondents with no access to or interest in the Internet questionnaire only responded after 4 weeks when the paper questionnaire was sent. The relatively quick response by postal respondents in the mixed-mode survey could be explained by the fact that respondents had already been informed about the study. Use of prenotification has been shown to shorten response times [19,31]. Another method to reduce the return time is sending the paper questionnaire out earlier.

Research has shown that an Internet survey results in more complete data compared with a postal survey [32]. This conclusion is confirmed in our study; the quality of data was higher in the mixed-mode survey than in the postal survey. One of the advantages of using the Internet for survey research is the technique of designing questionnaires so that complex skip patterns are invisible to respondents. As a consequence, the online questionnaire resulted in zero missing items and zero invalid answers (eg, answers to questions that had to be skipped). However, given the fact that some groups of people are underrepresented on the Internet (for instance, the elderly), conducting surveys through the Internet alone is not (yet) possible [11,21].

One of the key potential advantages of using the Internet over paper questionnaires is reducing costs. This study showed that the costs per returned questionnaire was €2 lower in the mixed-mode survey than in the postal survey. In the present study, the information technology costs were, however, relatively high for the mixed-mode survey. This was due to the need to program two applications, one for scanning the paper questionnaires and one for the Internet questionnaires. In the future, more costs can possibly be saved by using one and the same program for the different data collection methods within a mixed-mode survey. In addition, the variable costs per questionnaire were lower and the fixed costs per questionnaire were higher in the mixed-method survey than in the postal survey. Fixed costs per questionnaire can be reduced if a larger sample is taken, because the fixed activities are divided over the number of returned questionnaires. In other words, the larger the sample, the more money can be saved by using a mixed-mode survey.

Our study was the first to examine so-called mode effects between a mixed-method survey (Internet with paper follow-up) and postal survey. We found no differences between the two surveys concerning global ratings respondents gave to different health care providers. This is beneficial, because it implies that there is no bias in the scores that is a function of the manner of data collection. Other studies did find mode effects between the answers of telephone respondents and postal respondents [23], Internet respondents and telephone respondents [24], and Internet respondents and postal respondents [25,28,33]. One study investigated the differences between a postal and an Internet questionnaire, where a subset of the participants filled out also the alternative version (Internet and paper questionnaire, respectively). They found little or no evidence for a difference in test–retest reliability and internal consistency when they compared the Internet and paper versions of the questionnaire [26].

We did not ask why respondents in the mixed-mode survey did not fill out the questionnaire online. In one study among nonrespondents of an Internet questionnaire, the nonrespondents indicated that they did not have a computer or access to the Internet. Other reasons were having no experience with the Internet or not trusting the Internet [31]. This corresponds with findings by other researchers, who showed that factors influencing response times are privacy concerns and computer anxiety [19,28,33

Acknowledgments

This study was financial supported by the Netherlands Organization for Health Research and Development (ZonMw).

Conflicts of Interest

None declared

References

1. Schut FT, Van de Ven WP. Rationing and competition in the Dutch health-care system. Health Econ 2005 Sep;14(Suppl 1):S59-S74. [doi: 10.1002/hec.1036] [Medline: 16161190]
2. Delnoij DM, ten Asbroek G, Arah OA, de Koning JS, Stam P, Poll A, et al. Made in the USA: the import of American Consumer Assessment of Health Plan Surveys (CAHPS) into the Dutch social insurance system. Eur J Public Health 2006 Dec;16(6):652-659 [FREE Full text] [doi: 10.1093/eurpub/ckl023] [Medline: 16524940]
3. Zuidgeest M, Sixma H, Rademakers J. Measuring patients' experiences with rheumatic care: the consumer quality index rheumatoid arthritis. Rheumatol Int 2009 Apr 16;30(2):159-167. [doi: 10.1007/s00296-009-0926-3] [Medline: 19370351]
4. Damman OC, Hendriks M, Sixma HJ. Towards more patient centred healthcare: a new Consumer Quality Index instrument to assess patients' experiences with breast care. Eur J Cancer 2009 Jun;45(9):1569-1577. [doi: 10.1016/j.ejca.2008.12.011] [Medline: 19167212]
5. Brandt H, Zuidgeest M, Sixma H. Development of CQ-index Care for the Handicapped: measuring the quality of care from a client perspective (in Dutch). Utrecht: NIVEL; 2007. URL: http://www.nivel.nl/pdf/Pilot-ontwikkeling-CQ-index-Gehandicaptenzorg-2007.pdf [accessed 2010-12-06] [WebCite Cache ID 5ulsU7mfM]
6. Wiegers TA, Stubbe JH, Triemstra M. Development of the CQ-Index Nursing Homes and Homecare: Quality of Life for Residents, Representatives and Clients. Utrecht: NIVEL; 2007.
7. Couper MP. Web survey design and administration. Public Opin Q 2001;65(2):230-253. [Medline: 11420757]
8. Kaplowitz MD, Hadlock TD, Levine R. A comparison of web and mail survey response rates. Public Opin Q 2004;68(1):94-101.
9. Sills SJ, Song CY. Innovations in survey research: an application of web-based surveys. Soc Sci Comput Rev 2002;20(1):22-30.
10. Kiesler S, Sproull LS. Response effects in the electronic survey. Public Opin Q 1986;50(3):402-413.
11. Kwak N, Radler R. A comparison between mail and web surveys: response pattern, respondent profile, and data quality. J Off Stat 2002;18(2):257-73.
12. Schaefer DR, Dillman DA. Development of a standard e-mail methodology. Public Opin Q 1998;62(3):378-397.
13. Schmidt CB. Casting the Net: surveying an Internet population. J Comput Mediat Commun 1997;3(1).
14. Tse ACB. Comparing the response rate, response speed, and response quality of two methods of sending questionnaires: e-mail vs. mail. J Market Res Soc 1998;40(4):354-361.
15. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison (2). J Med Internet Res 2004 Oct 29;6(4):e39 [FREE Full text] [doi: 10.2196/jmir.6.4.e39] [Medline: 15631963]
16. Zuidgeest M, Hendriks M, Spreeuwenberg J, Rademakers J. Usefulness of the CQ index: Report Volume 1: Methods of Data Collection for CQI Research (in Dutch). Utrecht: NIVEL; 2008.
17. Boer de D, Hendriks M, Damman OC, Spreeuwenberg P, Rademakers J, Delnoij DJM, et al. Experiences of Insured With Care and Health insurer: CQ-Index Health and Health Insurance, Measurement 2007 (in Dutch). Utrecht: NIVEL; 2007.
18. Slijkhuis R. Is Dillman Absolute? Research Into the Influence of Layout and Contact Process of CQI Surveys on the Quantitative Response, Qualitative Response and Cost of Data Collection (in Dutch) [Bachelors thesis]. Twente, Sweden: University Twente; 2008.
19. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Methods to increase response rates to postal questionnaires. Cochrane Database Syst Rev 2007(2):MR000008. [doi: 10.1002/14651858.MR000008.pub3] [Medline: 17443629]
20. Hoksbergen M. Centraal Bureau voor de Statistiek. 2008 Dec 17. 1.2 Million Dutch People Have no Internet Access at Home URL: http://www.cbs.nl/nl-NL/menu/themas/vrije-tijd-cultuur/publicaties/artikelen/archief/2008/2008-2641-wm.htm [accessed 2011-05-11] [WebCite Cache ID 5ybU734dq]
21. Madden M. Internet Penetration and Impact. Washington, DC: Pew Internet & American Life Project; 2009 Mar 24. URL: http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Internet_Impact.pdf.pdf [accessed 2009-03-03] [WebCite Cache ID 5fViAM9uE]
22. van Campen C, Sixma H, Kerssens JJ, Peters L. Comparisons of the costs and quality of patient data collection by mail versus telephone versus in-person interviews. Eur J Public Health 1998:66-70.
23. de Vries H, Elliott MN, Hepner KA, Keller SD, Hays RD. Equivalence of mail and telephone responses to the CAHPS Hospital Survey. Health Serv Res 2005 Dec;40(6 Pt 2):2120-2139. [doi: 10.1111/j.1475-6773.2005.00479.x] [Medline: 16316441]
24. Christian LM, Dillman DA, Smyth JD. The Effects of Mode and Format on Answers to Scalar Questions in Telephone and Web Surveys. In: Lepkowski JM, Tucker C, Leeuw EDD, Japec L, Lavrakas P, Link M, et al, editors. Advances in Telephone Survey Methodology. Hoboken, NJ: Wiley-Interscience; 2007:250-275.
25. Carini RM, Hayek JC, Kuh GD, Kennedy JM, Ouimet JA. College student responses to web and paper surveys: does mode matter? Res Higher Educ 2003;44(1):1-19 [FREE Full text]
26. Basnov M, Kongsved SM, Bech P, Hjollund NH. Reliability of short form-36 in an Internet- and a pen-and-paper version. Inform Health Soc Care 2009 Jan;34(1):53-58. [doi: 10.1080/17538150902779527] [Medline: 19306199]

27. Koopman L, Rademakers J. CQ-index Breast Care: Research Into the Discriminative Power: Quality of Care as Perceived by People With a Breast Abnormality From the Patients' Perspective (in Dutch). Utrecht: NIVEL; 2008. URL: http://www.nivel.nl/pdf/Rapport-CQI-Mammacare.pdf [accessed 2010-12-06] [WebCite Cache ID 5ultiv6lD]
28. Dillman DA, Smuth JD, Cristian LM. Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: Wiley; 2009.
29. Dickerson MD, Gentry JW. Characteristics of adopters and non-adopters of home computers. J Consumer Re 1983 Sep;10(2):225-235.
30. Slijkhuis R. The Internet: Valid or Flawed: The Internet: Valid or Flawed: Study Into the Validity of the Internet as a Mode of Data Collection in a Survey Among the General Dutch Public [Masters thesis]. Twente, Sweden: University Twente; 2008.
31. 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] [doi: 10.2196/jmir.6.3.e29] [Medline: 15471755]
32. Larsson J, Sandelin K, Forsberg C. Health-related quality of life and healthcare experiences in breast cancer patients in a study of Swedish women. Cancer Nurs 2010;33(2):164-170. [doi: 10.1097/NCC.0b013e3181bb0d05] [Medline: 20142743]
33. 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 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]

Edited by G Eysenbach; submitted 24.03.09; peer-reviewed by NH Hjollund, P Hoonakker, D Dillman; comments to author 28.04.09; revised version received 20.04.10; accepted 20.08.10; published 27.09.11.

Please cite as:
Zuidgeest M, Hendriks M, Koopman L, Spreeuwenberg P, Rademakers J
A Comparison of a Postal Survey and Mixed-Mode Survey Using a Questionnaire on Patients' Experiences With Breast Care
J Med Internet Res 2011;13(3):e68
URL: http://www.jmir.org/2011/3/e68/
doi:10.2196/jmir.1241
PMID:21946048

Original Paper

Using Web-Based and Paper-Based Questionnaires for Collecting Data on Fertility Issues Among Female Childhood Cancer Survivors: Differences in Response Characteristics

Marleen H van den Berg[1], PhD; Annelies Overbeek[1,2], MD, ; Helena J van der Pal[3], MD; A. Birgitta Versluys[4], MD; Dorine Bresters[5], MD, PhD; Flora E van Leeuwen[6], PhD; Cornelis B Lambalk[2], MD, PhD; Gertjan J.L Kaspers[1], MD, PhD; Eline van Dulmen-den Broeder[1], PhD

[1]VU University Medical Center Amsterdam, Department of Pediatrics, division of Oncology-Hematology, Amsterdam, Netherlands

[2]VU University Medical Center, Department of Obstetrics & Gynaecology, Amsterdam, Netherlands

[3]Emma Children's Hospital/ Academic Medical Center, Department of Pediatric Oncology and Medical Oncology, Amsterdam, Netherlands

[4]University Medical Center Utrecht/Wilhelmina Children's Hospital, Department of Pediatric Hematology/Immunology and Bone Marrow Transplantation, Utrecht, Netherlands

[5]Leiden University Medical Center, Willem-Alexander Children's Hospital, Leiden, Netherlands

[6]Netherlands Cancer Institute, Department of Epidemiology, Amsterdam, Netherlands

Corresponding Author:
Marleen H van den Berg, PhD
VU University Medical Center Amsterdam
Department of Pediatrics, division of Oncology-Hematology
P.O. box 7057
Amsterdam, 1007 MB
Netherlands
Phone: 31 20 4446021
Fax: 31 20 4445122
Email: mh.vandenberg@vumc.nl

Abstract

Background: Web-based questionnaires have become increasingly popular in health research. However, reported response rates vary and response bias may be introduced.

Objective: The aim of this study was to evaluate whether sending a mixed invitation (paper-based together with Web-based questionnaire) rather than a Web-only invitation (Web-based questionnaire only) results in higher response and participation rates for female childhood cancer survivors filling out a questionnaire on fertility issues. In addition, differences in type of response and characteristics of the responders and nonresponders were investigated. Moreover, factors influencing preferences for either the Web- or paper-based version of the questionnaire were examined.

Methods: This study is part of a nationwide study on reproductive function, ovarian reserve, and risk of premature menopause in female childhood cancer survivors. The Web-based version of the questionnaire was available for participants through the Internet by means of a personalized user name and password. Participants were randomly selected to receive either a mixed invitation (paper-based questionnaire together with log-in details for Web-based questionnaire, n = 137) or a Web-only invitation (log-in details only, n = 140). Furthermore, the latter group could request a paper-based version of the questionnaire by filling out a form.

Results: Overall response rates were comparable in both randomization groups (83% mixed invitation group vs 89% in Web-only invitation group, *P* = .20). In addition, participation rates appeared not to differ (66% or 90/137, mixed invitation group vs 59% or 83/140, Web-only invitation group, *P* =.27). However, in the mixed invitation group, significantly more respondents filled out the paper-based questionnaire compared with the Web-only invitation group (83% or 75/90 and 65% or 54/83, respectively, *P* = .01). The 44 women who filled out the Web-based version of the questionnaire had a higher educational level than the 129 women who filled out the paper-based version (*P* = .01). Furthermore, the probability of filling out the Web-based questionnaire appeared to be greater for women who were allocated to the Web-only invitation group (OR = 2.85, 95% CI 1.31 - 6.21), were older (OR

= 1.08, 95% CI 1.02 - 1.15), had a higher educational level (OR high vs low = 0.06, 95% CI 0.01 - 0.52), or were students (OR employed vs student = 3.25, 95% CI 1.00 - 10.56).

Conclusions: Although overall response as well as participation rates to both types of invitations were similar, adding a paper version of a questionnaire to a Web-only invitation resulted in more respondents filling out the paper-based version. In addition, women who were older, had a higher level of education, or were students, were more likely to have filled out the Web-based version of the questionnaire. Given the many advantages of Web-based over paper-based questionnaires, researchers should strongly consider using Web-based questionnaires, although possible response bias when using these types of questionnaires should be taken into account.

Trial Registration: Nederlands Trial Register NTR2922; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2922 (Archived by WebCite at http://www.webcitation.org/5zRRdMrDv)

(J Med Internet Res 2011;13(3):e76) doi:10.2196/jmir.1707

KEYWORDS

Paper; Internet; questionnaire; response; fertility; cancer survivors

Introduction

The number of Internet users worldwide has doubled in the past 5 years, and it is estimated there are over 2 billion users in 2010 [1]. In the Netherlands, there were up to 15 million Internet users in 2010, representing 85.6% of the Dutch population [2]. Not surprisingly, within research settings, the Internet is increasingly being used as a tool for collecting data by means of Web-based questionnaires. The use of this type of questionnaire is less time-consuming and less costly compared with the use of paper-based questionnaires. For Web-based questionnaires, no printing and mailing costs are involved and the time spent by a researcher on data entry is minimal since the returned data are already in an electronic format. In addition, studies have reported that Web-based questionnaires have fewer response errors and fewer socially desirable responses, while no differences have been found in the accuracy of the reported information between the two types of questionnaires [3-5]. Thus, Web-based questionnaires might serve as an attractive alternative to paper-based questionnaires, especially when the target study population primarily consists of relatively young respondents [6]. However, important technical and methodological issues have been raised that should be carefully considered when using Web-based questionnaires [7,8]. An important issue is obtaining representative samples of the study population with adequate response rates to secure external validity. If response rates to a Web-based questionnaire appear to be low or seem to come from a selective group, response bias is introduced and the results might be misleading [6,9].

In the past decade, several studies have investigated response rates of Web-based versus paper-based questionnaires in many different populations and in many different settings. Response rates appear to vary widely and seem to be more dependent on the population sampled than on any other factor [10-12]. Although conflicting results have been published, recent studies have demonstrated an increase in response rates to Web-based compared with paper-based questionnaires [10,11,13,14]. Employing a mixed-mode strategy, enabling patients to fill out either a Web-based or a paper-based questionnaire, seems to enhance response rates even further [15,16]. These two types of questionnaires can be offered simultaneously or sequentially, a factor that also seems to influence response rates [17].

The current study is part of a Dutch nationwide study. This study, which was initiated in 2007, examines the effects of childhood cancer and its treatment on reproductive function, ovarian reserve, and risk of premature menopause in female childhood cancer survivors. The questionnaire used in this study, of which both a paper-based version as well as a Web-based version were available, contains questions about several fertility-related issues. So far, no studies are available comparing response rates to a Web- and paper-based version of a questionnaire on fertility issues among young adult female cancer survivors. Indeed, previous studies among female childhood cancer survivors have predominantly used paper-based questionnaires, telephone interviews, and face-to-face interviews to collect data [18]. In the US Childhood Cancer Survivor Study (CCSS) as well as in the UK British Childhood Cancer Survivor Study (BCCSS), paper-based questionnaires have been sent to large cohorts of childhood cancer survivors. These questionnaires contained questions about sociodemographic items, adverse health outcomes, use of medications, lifestyle behavior, pregnancy history, and family history. Reported response rates were 82% (CCSS) and 71% (BCCSS), respectively [19,20].

In addition, studies investigating response rates to Web-based questionnaires among survivors of childhood cancer are scarce. Thompson et al [21] used a Web-based questionnaire to investigate difficulties regarding romantic relationships in childhood cancer survivors. For this purpose, 603 survivors were sent a letter by postal mail inviting them to participate. Only 60 survivors (10%) agreed to participate and filled out the Web-based questionnaire. Low response rates were also reported by Cantrell et al [22] in their study of health-related quality of life following childhood cancer. A Web-based survey was used, which was brought to the attention of potential respondents by posting a link on six different websites intended for use by survivors of childhood cancer. Although exact response rates could not be calculated, the authors reported the response rate to be low and the time needed to collect data to be long. In another study, childhood cancer survivors were recruited for a Web-based survey on physical activity through advertisements posted on cancer survivor–related websites and newsletters [23]. Since that study also used a reactive recruitment method, no true response rates could be calculated. However, the authors stated that they realize that the generalizability of their study was limited, as the recruitment method used probably had led to a specific group of survivors responding to the study invitation.

In conclusion, it is not known what response and participation rates can be expected when inviting female childhood cancer survivors to fill out a Web-based or a paper-based questionnaire on fertility issues. More specifically, no information is available on the impact of adding a paper-based questionnaire to an invitation to fill out the Web-based questionnaire. Therefore, we aimed to evaluate whether sending a mixed invitation (paper-based together with Web-based questionnaire) rather than a Web-only invitation (Web-based questionnaire only) results in higher response and participation rates for female childhood cancer survivors filling out a questionnaire on fertility issues. Furthermore, in order to identify possible response bias, differences in type of response and characteristics of the responders and nonresponders were investigated. Moreover, factors influencing preferences for either the Web- or paper-based version of the questionnaire were examined.

Methods

Eligible survivors for the nationwide study were selected from a cohort of patients treated for childhood cancer at one of the five Dutch pediatric oncology centers or one of the two stem cell transplant centers between 1965 and 2002. Within the collaborative Dutch Childhood Oncology Late Effects Group, an electronic database has been set up in each center that includes patient and treatment details of all patients treated for cancer before the age of 18 years. Inclusion criteria for the nationwide study and the current study were identical and were defined as: female sex, having been treated for a malignancy or central nervous system tumor before the age of 18, having survived for at least 5 years, being alive, and being at least 18 years of age at study entry. Patients were excluded if they were not able to speak or read Dutch or if they had severe sequelae related to mental health.

The nationwide study consists of three components: a questionnaire, the provision of a blood sample, and a transvaginal ultrasound measurement of the reproductive organs. The last two of these components require a hospital visit. Patients can either refuse to participate or take part in one, two, or all three components of the study. For the purpose of the current report, only the questionnaire component was taken into account.

Questionnaire and Procedures for Distribution

The questionnaire used in the study is an adaptation of a well-tested questionnaire used by the Department of Epidemiology of the Netherlands Cancer Institute in a large-scale Dutch cohort study on long-term effects of ovarian stimulation for in vitro fertilization [24]. It addresses the following issues: sociodemographic characteristics, medical history, menstrual and reproductive history, pregnancy outcomes, menopausal symptoms and menopause, and family history of cancer and family history of subfertility or infertility.

The paper- and Web-based version of the questionnaire were identical in terms of the questions asked, their wording, and their order of presentation. In the Web-based version, however, questions not relevant to the participant were automatically skipped. The Web-based version of the questionnaire was accessible for participants through a website which was specially designed for the nationwide study.

The study population for this study consisted of 277 female childhood cancer survivors from three participating centers of the nationwide study (Emma Children’s Hospital/ Academic Medical Center Amsterdam, Leiden University Medical Center, University Medical Center Utrecht/Wilhelmina Children's Hospital). These women were randomly allocated to two groups: the mixed invitation group and the Web-only invitations group.

The Mixed Invitation Group

Participants in the mixed invitation group received an invitation that contained a paper-based questionnaire together with an instruction sheet for the Web-based questionnaire. This instruction sheet contained a personalized username, the name of the website, and a log-in code allowing them to log in to a secured part of the website and fill out the questionnaire.

The Web-only Invitation Group

Participants in the Web-only group received the above-mentioned instruction sheet containing the name of the website and the log-in details alone. However, a paper-based questionnaire could be acquired by ticking this option on the informed consent form.

For practical and logistical reasons, invitations for the nationwide study (and thus for the current study) were sent out consecutively in batches consisting of invitations to 30 to 50 women. The calculation of the target sample size was based on the expected proportions of participants in both randomization groups filling out the paper-based questionnaire. Based on a previous study by Quigley et al [25], in which similar randomization groups were used, it was estimated that 77% of participants in the mixed invitation group and 27% of the participants in the Web-only invitation group would complete and return the paper-based version of the questionnaire. With 95% power and a significance level of .05, it was estimated that a minimum of 26 participants would be required in each group [26]. However, it was decided to include all women who were invited for the nationwide study during a fixed period of time (ie, January 1, 2009, through May 31, 2010), thereby assuring that the target sample size would be met.

Randomization occurred by sorting the survivors alphabetically based on the street name of their address, after which the first half of the survivors was allocated to the mixed invitation group and the second half to the Web-only invitation group.

All eligible female survivors received a study information package by postal mail consisting of an informed consent form, a refusal form, a postage-paid reply envelope, and an instruction sheet with personalized log-in details. Depending on the allocated randomization group, a paper-based questionnaire was added to this study information package. The envelope containing the study information package was sealed and put in another envelop together with a cover letter, signed by the head of the relevant pediatric oncology department, in which the study was explained very briefly. This was done in order to give survivors the chance to return the entire study information package without having to open the envelope containing this package and without having to read the extensive study information. Thus, survivors could respond to the study information package that was sent in four different ways. These were: (1) sending back a filled-out questionnaire (either Web-based or paper-based) together with a filled-out informed consent form, (2) sending back a filled-out informed consent form only in cases where the potential participant was not willing to fill out the questionnaire but was willing to take part in other parts of the study, (3) sending back a filled-out refusal form, (4) sending back the entire study information package marked return to sender. For the purpose of this study, survivors were categorized as being responders if they chose one of the four above-mentioned response options, otherwise they were categorized as nonresponders.

Participants in both groups were assured that all information provided both by the paper-based as well as the Web-based questionnaire was confidential. Moreover, it was mentioned that data provided via the Web-based version were transmitted over a secure Internet connection and could not be viewed by unauthorized persons.

Follow up and Reminders

If an envelope appeared undeliverable because of an incorrect or nonexistent postal address, the online telephone directory was used to find the correct address. If this proved unhelpful, vital status and current address were checked by means of the Gemeentelijke Basis Administratie (Dutch Municipal Population Register).

If the questionnaire was not returned within 3 weeks, a reminder was sent by postal mail. For participants in the mixed invitation group, this reminder consisted of a letter in which the relevance of the study was again stressed and in which the individual was asked to respond. For participants in the Web-only invitation group, a paper-based version of the questionnaire was added to this reminder letter. When, after 3 weeks, still no response was received, patients in both groups were contacted by telephone and were asked to respond.

For the purpose of the current study, response time is defined as the time (number of days) elapsed between the day the envelope with the study information package was sent and the day a response was received.

Data Analysis

Data were analyzed using SPSS for Windows, version 15.0 (SPSS Inc, Chicago, IL, USA). Descriptive statistics were used to describe differences between (1) participants allocated to the mixed invitation group and the Web-only invitation group, (2) respondents filling out the paper-based questionnaire and the Web-based questionnaire, and (3) responders and nonresponders. Independent samples t tests and Pearson chi-square tests were used to test whether these differences were statistically significant. A P value of less than .05 was considered to be statistically significant.

Multivariable logistic regression analysis was used to predict the probability of filling out the Web-based version of the questionnaire by reporting odds ratios (ORs) and 95% confidence intervals (CIs). A prediction model was developed using a backward selection procedure with a P value of .10 as the criterion for exclusion of variables.

Results

General Response Characteristics

Included in this study were 277 women. Table 1 outlines the response characteristics of the participants allocated to the two randomization groups. Response rates were comparable in both groups (83% in mixed invitation group vs 89% in Web-only invitation group, $P = .20$). In addition, participation rates—defined as the number of women who were willing to fill out the questionnaire—did not differ significantly between the mixed invitation group and the Web-only invitation group (66% or 90/137 vs 59% or 83/140, respectively, $P = .27$). Moreover, median response time was comparable in both groups. In the mixed invitation group, significantly more respondents filled out the paper-based questionnaire compared with the Web-only invitation group (83% or 75/90 and 65% or 54/83, respectively, $P = .01$).

Table 1. Response characteristics of the participants receiving the mixed invitation and the Web-only invitation (n = 277)

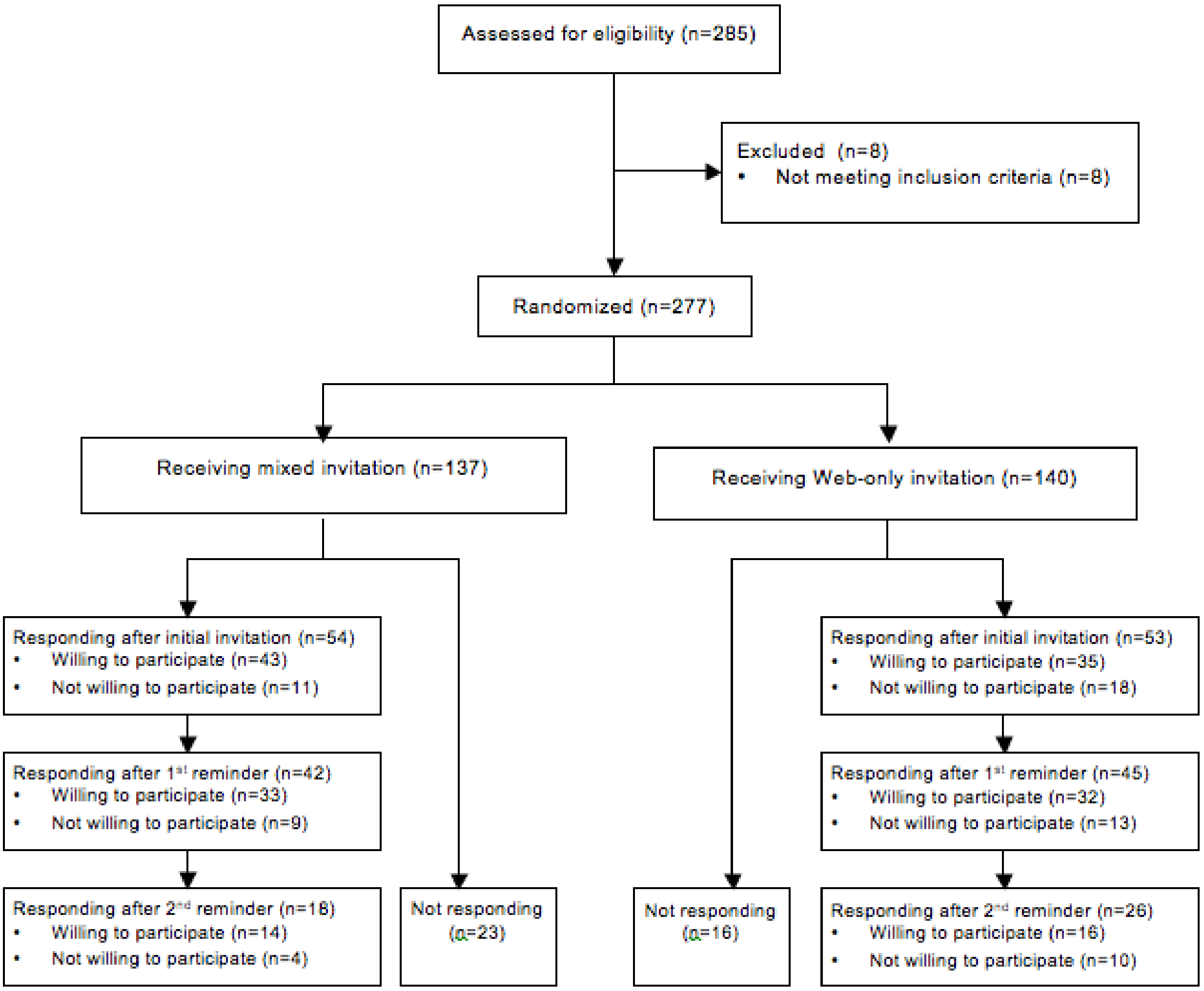
	Mixed Invitation Group (n=137)	Web-Only Invitation Group (n=140)	*P* Value
Number of responders, n (%)	114 (83)	124 (89)	.20
Response time in days, median (interquartile range [IQR])	32.5 (54.5)	34.5 (44.8)	.51
Reminders sent, n (%)			
By mail	88 (64)	93 (66)	.70
By telephone	29 (21)	31 (22)	.84
Timing of response, n (%)			
After initial invitation	54 (47)	53 (43)	
After 1st reminder (by mail)	42 (37)	45 (36)	.56
After 2nd reminder (by telephone)	18 (16)	26 (21)	
Type of response, n (%)			
Returned envelope to sender	6 (5)	10 (8)	
Refused (sent back refusal form)	18 (16)	31 (25)	.12
Willing to participate	90 (79)	83 (67)	
Type of questionnaire filled out, n (%)			
Paper-based	75 (83)	54 (65)	.01
Web-based	15 (17)	29 (35)	

Figure 1 summarizes the participant flow. Overall there were no differences between the responders in the mixed invitation group and the Web-only invitation group with respect to the number of women responding after the initial invitation, after the first reminder, and after the second reminder. However, when the results regarding the timing of the response are related to the type of questionnaire filled out by the respondents, some differences between the two groups appear. Among the group of responders who sent back either type of the questionnaire before the first postal reminder was sent, that is, the "fast responders," the proportion of responders filling out the paper-based version of the questionnaire was significantly larger in the mixed invitation compared with the Web-only invitation group (74% or 32/43 vs 51% or 18/35, respectively, *P* = .04). This difference in type of response remained after the first postal reminder (to which a paper-based version of the questionnaire was added) was sent, but it was no longer statistically significant: 91% (30/33) in the mixed invitation group compared with 75% (24/32) in the Web-only invitation group filled out the paper-based version of the questionnaire after being reminded by postal mail (*P* = .09).

When the 238 women who responded to the study invitation by sending back the questionnaire, the informed consent form, the refusal form, or the entire study information package were compared with the 39 women who did not respond at all, it appeared that these two groups did not significantly differ regarding age, age at diagnosis, or type of diagnosis. The nonresponder group included 20 survivors (12 of 137 or 9% in the mixed invitation group and 8 of 140 or 6% in the Web-only invitation group) whose postal address could not be verified and who could not be contacted by telephone either. It was decided to consider these survivors to be nonresponders. However, they might not be "true" nonresponders since it is not known whether they indeed received the study information package and the postal reminder.

Comparing the 173 women who participated in this study with the 104 women who did not participate (ie, women indicating they refused to participate and women who did not respond) did not reveal significant differences regarding current age or age at diagnosis. However, it appeared that the proportion of women with leukemia was significantly higher in the participant group compared with the nonparticipant group (52% or 90/173 vs 38% or 39/104, respectively, *P* = .02) (data not shown).

Figure 1. Flow diagram of participants.



Characteristics of Questionnaire Respondents

Table 2 shows the characteristics of the 173 women who returned a questionnaire. It appeared that the 44 women who filled out the Web-based version of the questionnaire were more likely to have had a higher educational level than the 129 women who filled out the paper-based version ($P = .01$). No differences were found regarding age, type of diagnosis, age at diagnosis, employment status, or marital status.

Table 2. Characteristics of respondents who filled out the paper-based and the Web-based questionnaire (n = 173)

	Paper-Based Questionnaire (n=129)	Web-Based Questionnaire (n=44)	*P* Value
Age in years, mean ± SD (range)	29.7 ± 7.9 (18.8-52.3)	30.9 ± 8.6 (19.4-52.1)	.40
Age at diagnosis in years, mean ± SD (range)	7.4 ± 4.7 (0.4-19.5)	8.9 ± 4.6 (0.6-15.9)	.07
Type of diagnosis, n (%)			
Leukemias	70 (54)	20 (45)	
Lymphomas	17 (13)	9 (21)	
Brain and central nervous system cancers	5 (4)	3 (7)	
Bone tumors	10 (8)	2 (5)	.73
Neuroblastomas	6 (5)	1 (2)	
Germ cell tumors	3 (2)	2 (5)	
Nephroblastomas	5 (4)	3 (7)	
Other	13 (10)	4 (9)	
Educational level, n (%)[a]			
High	30 (24)	19 (43)	
Medium	78 (62)	24 (55)	.01
Low	18 (14)	1 (2)	
Employment status, n (%)			
Unemployed	28 (22)	4 (9)	
Student	15 (12)	8 (18)	.12
Employed	82 (66)	32 (73)	
Marital status, n (%)			
Never married	86 (67)	29 (66)	
Married	39 (31)	15 (34)	.56
Divorced	3 (2)	0	

[a] Categorized as low, up to and including lower technical and vocational training; medium, up to and including secondary technical and vocational training; and high, up to and including higher technical and vocational training and university

Among the group of respondents filling out the questionnaire, it was investigated which factors influenced the probability of filling out either the paper- or the Web-based version of the questionnaire. Table 3 shows the odds ratios and 95% confidence intervals (CI) for the variables in the final model of the logistic regression analysis. Age, educational level, employment status, and randomization group were significant factors influencing the probability of filling out the Web-based questionnaire. More specifically, the probability of filling out the Web-based questionnaire was higher for participants allocated to the Web-based invitation group for participants who were older, and for participants with a higher educational level. Finally, students appeared to have a higher probability of filling out the Web-based questionnaire compared with participants who were employed.

Table 3. Factors associated with the probability of filling out the Web-based version of the questionnaire: results of logistic regression[a]

	P value	OR	95% CI	
			Lower	Upper
Age	.01	1.08	1.02	1.15
Randomization group (reference group: mixed invitation group)	.01	2.85	1.31	6.21
Educational level (reference group: high level)	.04			
Medium		0.65	0.28	1.53
Low		0.06	0.01	0.52
Employment status (reference group: employed)	.03			
Student		3.25	1.00	10.56
Unemployed		0.35	0.10	1.29

[a] Nagelkerke pseudo R^2 = 0.21

Discussion

Statement of Principal Findings

In the present study, we examined differences in response between female childhood cancer survivors who received either a mixed invitation (paper-based questionnaire together with log-in details for Web-based questionnaire) or a Web-only invitation (log-in details only). The results show that survivors receiving the mixed invitation preferred filling out the paper-based version instead of the Web-based questionnaire as compared with the survivors receiving the Web-only invitation. Thus, when a paper-based version of the questionnaire was added to an invitation in which also the possibility of filling out the Web-based version was mentioned, the survivors were more likely to choose the paper-based option. Moreover, when the results regarding the timing of the response are taken into account this finding is endorsed since a large proportion (75%, 24/32) of females who initially received the log-in details only responded by filling out the paper-based questionnaire after they received a postal reminder (3 weeks later) to which a paper-based version of the questionnaire was added. This proportion is comparable to the proportion of females filling out the paper-based questionnaire immediately after the invitation (ie, before the postal reminder) among those who initially received the log-in details together with the paper-based version of the questionnaire (74%, 24/32).

Comparison With Other Studies

To our knowledge, no studies are available that have compared response rates to a Web- and paper-based version of a questionnaire on reproductive and fertility issues among young adult women. However, a few studies are available that have evaluated these issues by means of a Web-based questionnaire only. In a group of female survivors of breast cancer, the response rate to this type of questionnaire was 51% [27,28] whereas in a group of women aged 17 to 21 years, this rate was 72% [29]. However, no information was provided regarding characteristics of the nonresponder group.

In our study, the overall response rates in the mixed invitation group and the Web-only invitation group did not differ. This result is in line with the results found in the study of Quigley et al [25]. In the study by Quigley et al, military personnel were requested to participate in a survey on information services. In one study group, a paper-based questionnaire was used with an added option of completing the questionnaire via the Internet. In the other study group, an Internet-based questionnaire was used with an added option of completing a paper version of the questionnaire by mail. Although response rates in both study groups were lower (42% and 37% respectively) than the response rates found in our study, differences in response rates between the two groups were not found, as was the case in our study. Furthermore, of the participants receiving the paper-based questionnaire with the Internet option, 77% chose to complete paper-based questionnaire. In our study, a similar proportion of participants in the mixed invitation group filled out the paper-based questionnaire, that is, 83%. However, Quigley et al found that of the participants receiving the Web-based questionnaire with the option of the paper-based version, 73% chose to complete the Web-based questionnaire, while in our study the proportion of women in the Web-only invitation group who filled out the Web-based questionnaire was much lower (35%).

Furthermore, the participation rates (ie, the proportion of women who filled out the questionnaire) measured in our study can be considered as being rather high (66% in the mixed invitation group and 59% in the Web-only invitation group). In other studies using Web-based questionnaires in combination with paper-based versions these rates are, in general, lower [17,25,30]. A possible explanation for the high response rates found in our study might be the salience of the study topic. It is known that potential participants are more likely to respond to both paper-based and Web-based surveys when the salience of the topic, defined as the degree to which the topic is of interest or is relevant for participants, is high [5,12,31]. Moreover, the questionnaire used in the current study was one of the three study components used in a nationwide study on fertility issues in female childhood cancer survivors, with the other two study components being the provision of a blood sample and a transvaginal ultrasound measurement of the reproductive organs. It is known that female survivors of childhood cancer are in need of information regarding their reproductive function [32,33]. Therefore, participation in this study might be appealing for a large group of the invited females, resulting in higher response rates compared with studies in which a questionnaire is the only measurement instrument used.

Our results show that the use of reminders improved the response rates substantially. After the first reminder (a letter sent by postal mail), the response almost doubled in both randomization groups. Other studies support our finding that both postal and telephone reminders are effective in increasing response rates for both Web-based surveys as well as traditional paper-based surveys [16,34,35].

In our study, the majority of the respondents preferred filling out the paper-based version of the questionnaire over filling out the Web-based version. Moreover, age, educational level, and employment status appeared to be important factors influencing the decision to fill out either version of the questionnaire. Our finding that women with a high education level as well as students tended to choose Web-based questionnaires over paper-based questionnaires is in line with previously published results [5,30,36]. However, results of the present study could not endorse other study results stating that Web-based questionnaires are more likely to attract younger respondents than paper-based questionnaires [6,36].

Another factor that may have played a role in the decision of the respondents to fill out either the paper-based or the Web-based version of the questionnaire is the length of the questionnaire used in the current study, which was rather long. The paper-based version consisted of 122 questions covering 32 pages. The Web-based questionnaire required several computer screens, with the number of questions on one screen depending on the type and length of the questions. For women filling out the Web-based version of the questionnaire, the median (IQR) time spent on filling out the questionnaire, which was automatically registered by the Web-based questionnaire tool, was 42.7 minutes (28.7 minutes to 67.8 minutes). Unfortunately, in the current study these data were not collected for the group of women filling out the paper-based version of the questionnaire. However, in the larger nationwide study, of which this study is part, a question was added to the paper-based questionnaire at a later point in time asking how much time was spent filling out the questionnaire. Median (IQR) time spent in this group (n = 145) was 30.0 minutes (30.0 minutes to 60.0 minutes) minutes. Thus, although this information was recorded among a different group of participants, it seems that filling out the paper-based version of the questionnaire took less time compared with the Web-based version. Various studies have shown the length of both Web-based and paper-based questionnaires to be negatively related to response rates [31,37-39]. However, no literature is available on differences in response rates to paper-based and Web-based questionnaires related to the length of a questionnaire. The results of our study seem to indicate that people tend to choose the paper-based version of a questionnaire when it concerns a long questionnaire. However, whether shorter surveys result in higher response rates when they are offered through the Web and longer surveys result in higher response rates when offered on paper needs further investigation.

In addition, the topic of our questionnaire can be considered to be rather personal. It is known that questionnaires containing questions of a sensitive nature result in lower response rates [31,38]. Although to our knowledge, studies investigating differences in response rates to paper-based and Web-based questionnaires taking into account the degree of sensitivity of the questions are lacking, one could assume that the sensitivity of our topic may have resulted in more respondents filling out the paper-based questionnaire especially since it is known that respondents filling out questionnaires through the Internet have doubts about their privacy and the confidentiality of their responses [10,40]. Despite the fact that data security and confidentiality were stressed in the letter accompanying our Web-based questionnaire, this might have led to more women filling out the paper-based questionnaire.

Limitations of Current Study

An important limitation of our study is the generalizability of the results found. Our study population mainly consisted of relatively young women, and thus the results may be less representative of older age groups or mixed groups including males. Moreover, our study population represents a rather unique clinic population, that is, long-term survivors of childhood cancer. In addition, the topic of the questionnaire used cannot be considered a conventional subject. Therefore, caution should be exercised when translating the results found in the current study to other study groups or other study topics. Furthermore, the available data on the nonresponders in the present study were limited to age, age at diagnosis, and type of diagnosis. As a consequence, potential bias introduced due to nonresponse, also influencing the generalizability, could not be investigated extensively. However, in many of the studies using both paper-based as well as Web-based questionnaires, data on nonresponders are not available at all. As nonresponse to surveys seems to be increasing in recent years [41,42], future studies investigating the degree of bias as well as its consequences for the interpretation of data collected by paper-based and Web-based questionnaires are of great importance.

Conclusions

Survivors of childhood cancer from this era represent a highly mobile group, and they may not be as available or as responsive to contact by traditional mail methods [43]. Successful recruitment of this population will require new methods of contact such as email and Web-based methods. Therefore, although our findings indicate that most survivors preferred the paper-based version over the Web-based version when offered both, we conclude that Web-based questionnaires are promising data collection tools for childhood cancer survivors. However, researchers should carefully weigh the methodological benefits and barriers of using either a paper-based or a Web-based questionnaire for this group of subjects, taking into account possible response bias.

Acknowledgments

This study was financially supported by the Dutch Cancer Society (grant number VU 2006-3622) and by the Children Cancerfree Foundation.

Conflicts of Interest

None declared

References

1. International Telecommunication Union. The world in 2010. ICT facts and figures. Geneva, Switzerland: Telecommunication Development Bureau; 2010. URL: http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf [accessed 2011-08-19] [WebCite Cache ID 613PgTZtL]
2. Internet World Stats. Bogata, Columbia: Miniwatts Marketing Group URL: http://www.internetworldstats.com/ [accessed 2010-12-13] [WebCite Cache ID 5uwVh5p9h]
3. 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]
4. Vereecken CA, Maes L. Comparison of a computer-administered and paper-and-pencil-administered questionnaire on health and lifestyle behaviors. J Adolesc Health 2006 Apr;38(4):426-432. [doi: 10.1016/j.jadohealth.2004.10.010] [Medline: 16549304]
5. Hoonakker P, Carayo P. Questionnaire Survey Nonresponse: A Comparison of Postal Mail and Internet Surveys. International Journal of Human-Computer Interaction 2009;25(5):348-373. [doi: 10.1080/10447310902864951]
6. Klovning A, Sandvik H, Hunskaar S. Web-based survey attracted age-biased sample with more severe illness than paper-based survey. J Clin Epidemiol 2009 Oct;62(10):1068-1074. [doi: 10.1016/j.jclinepi.2008.10.015] [Medline: 19246177]
7. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
8. Braithwaite D, Emery J, De Lusignan S, Sutton S. Using the Internet to conduct surveys of health professionals: a valid alternative? Fam Pract 2003 Oct;20(5):545-551 [FREE Full text] [Medline: 14507796]
9. Fleming CM, Bowden M. Web-based surveys as an alternative to traditional mail methods. J Environ Manage 2009 Jan;90(1):284-292. [doi: 10.1016/j.jenvman.2007.09.011] [Medline: 18082316]
10. Sax LJ, Gilmartin SK, Bryant AN. Assessing response rates and nonresponse bias in Web and paper surveys. Research in Higher Education 2003;44(4):409-432.
11. Greenlaw C, Brown-Welty S. A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. Eval Rev 2009 Oct;33(5):464-480. [doi: 10.1177/0193841X09340214] [Medline: 19605623]
12. Dillman DA, Smyth JD. Design effects in the transition to web-based surveys. Am J Prev Med 2007 May;32(5 Suppl):S90-S96. [doi: 10.1016/j.amepre.2007.03.008] [Medline: 17466824]
13. Kiernan NA, Kiernan M, Oyler MA, Gilles C. Is a web survey as effective as a mail survey? A field experiment among computer users. American Journal of Evaluation 2005;26(2):245-252. [doi: 10.1177/1098214005275826]
14. McCabe SE. Comparison of web and mail surveys in collecting illicit drug use data: a randomized experiment. J Drug Educ 2004;34(1):61-72. [Medline: 15468748]
15. Guise V, Chambers M, Välimäki M, Makkonen P. A mixed-mode approach to data collection: combining web and paper questionnaires to examine nurses' attitudes to mental illness. J Adv Nurs 2010 Jul;66(7):1623-1632. [doi: 10.1111/j.1365-2648.2010.05357.x] [Medline: 20497273]
16. McCabe SE, Diez A, Boyd CJ, Nelson TF, Weitzman ER. Comparing web and mail responses in a mixed mode survey in college alcohol use research. Addict Behav 2006 Sep;31(9):1619-1627. [doi: 10.1016/j.addbeh.2005.12.009] [Medline: 16460882]
17. Chizawsky LL, Estabrooks CA, Sales AE. The feasibility of Web-based surveys as a data collection tool: a process evaluation. Appl Nurs Res 2011 Feb;24(1):37-44. [doi: 10.1016/j.apnr.2009.03.006] [Medline: 20974058]
18. Hawkins MM, Robison LL. Importance of clinical and epidemiological research in defining the long-term clinical care of pediatric cancer survivors. Pediatr Blood Cancer 2006 Feb;46(2):174-178. [doi: 10.1002/pbc.20609] [Medline: 16333860]
19. Hawkins MM, Lancashire ER, Winter DL, Frobisher C, Reulen RC, Taylor AJ, et al. The British Childhood Cancer Survivor Study: Objectives, methods, population structure, response rates and initial descriptive information. Pediatr Blood Cancer 2008 May;50(5):1018-1025. [doi: 10.1002/pbc.21335] [Medline: 17849473]
20. Robison LL, Mertens AC, Boice JD, Breslow NE, Donaldson SS, Green DM, et al. Study design and cohort characteristics of the Childhood Cancer Survivor Study: a multi-institutional collaborative project. Med Pediatr Oncol 2002 Apr;38(4):229-239. [Medline: 11920786]
21. Thompson AL, Marsland AL, Marshal MP, Tersak JM. Romantic relationships of emerging adult survivors of childhood cancer. Psychooncology 2009 Jul;18(7):767-774. [doi: 10.1002/pon.1471] [Medline: 19061200]
22. Cantrell MA, Lupinacci P. Investigating the determinants of health-related quality of life among childhood cancer survivors. J Adv Nurs 2008 Oct;64(1):73-83. [doi: 10.1111/j.1365-2648.2008.04760.x] [Medline: 18808594]

23. Finnegan L, Wilkie DJ, Wilbur J, Campbell RT, Zong S, Katula S. Correlates of physical activity in young adult survivors of childhood cancers. Oncol Nurs Forum 2007 Sep;34(5):E60-E69 [FREE Full text] [doi: 10.1188/07.ONF.E60-E69] [Medline: 17878118]
24. de Boer EJ, den Tonkelaar I, te Velde ER, Burger CW, van Leeuwen FE, OMEGA-project group. Increased risk of early menopausal transition and natural menopause after poor response at first IVF treatment. Hum Reprod 2003 Jul;18(7):1544-1552 [FREE Full text] [Medline: 12832386]
25. Quigley B, Riemer RA, Cruzen D, Rosen S. Internet versus paper survey administration: Preliminary finding on response rates. 2000 Nov 9 Presented at: 42nd Annual Conference of the International Military Testing Association; November 7-9, 2000; Edinburgh, Scotland URL: http://www.imta.info/PastConferences/Year2000.aspx
26. Twisk JWR. Inleiding in de toegepaste biostatistiek [Introduction to Applied Biostatistics]. Maarssen, The Netherlands: Elsevier Gezondheidszorg; 2007.
27. Partridge AH, Gelber S, Peppercorn J, Sampson E, Knudsen K, Laufer M, et al. Web-based survey of fertility issues in young women with breast cancer. J Clin Oncol 2004 Oct 15;22(20):4174-4183. [doi: 10.1200/JCO.2004.01.159] [Medline: 15483028]
28. Partridge AH, Gelber S, Peppercorn J, Ginsburg E, Sampson E, Rosenberg R, et al. Fertility and menopausal outcomes in young breast cancer survivors. Clin Breast Cancer 2008 Feb;8(1):65-69. [Medline: 18501060]
29. Shrestha A, Nohr EA, Bech BH, Ramlau-Hansen CH, Olsen J. Parental age at childbirth and age of menarche in the offspring. Hum Reprod 2010 Mar;25(3):799-804 [FREE Full text] [doi: 10.1093/humrep/dep473] [Medline: 20089523]
30. Shih T, Fan X. Comparing response rates from web and mail surveys: a meta-analysis. Field Methods 2008;20(3):249-271. [doi: 10.1177/1525822X08317085]
31. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Computers in Human Behavior 2010;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]
32. Zebrack BJ, Casillas J, Nohr L, Adams H, Zeltzer LK. Fertility issues for young adult survivors of childhood cancer. Psychooncology 2004 Oct;13(10):689-699. [doi: 10.1002/pon.784] [Medline: 15386645]
33. Langeveld NE, Grootenhuis MA, Voûte PA, de Haan RJ, van den Bos C. Quality of life, self-esteem and worries in young adult survivors of childhood cancer. Psychooncology 2004 Dec;13(12):867-881. [doi: 10.1002/pon.800] [Medline: 15386796]
34. Schonlau M, Asch BJ, Du C. Web surveys as part of a mixed-mode strategy for populations that cannot be contacted by e-mail. Social Science Computer Review 2003;21(2):218-222. [doi: 10.1177/0894439303021002007]
35. Mavis BE, Brocato JJ. Postal surveys versus electronic mail surveys. The tortoise and the hare revisited. Eval Health Prof 1998 Sep;21(3):395-408. [Medline: 10350958]
36. Smith B, Smith TC, Gray GC, Ryan MA, Millennium Cohort Study Team. When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study. Am J Epidemiol 2007 Dec 1;166(11):1345-1354 [FREE Full text] [doi: 10.1093/aje/kwm212] [Medline: 17728269]
37. Walston JT, Lissitz RW, Rudner LM. The influence of web-based questionnaire presentation variations on survey cooperation and perceptions of survey quality. Journal of Official Statistics 2006;22(2):271-291.
38. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. BMJ 2002 May 18;324(7347):1183 [FREE Full text] [Medline: 12016181]
39. Galesic M, Bosnjak M. Effects of questionnaire length on participation and indicators of response quality in a web survey. Public Opinion Quarterly 2009;73(2):349-360. [doi: 10.1093/poq/nfp031]
40. Eaton DK, Brener ND, Kann L, Denniston MM, McManus T, Kyle TM, et al. Comparison of paper-and-pencil versus Web administration of the Youth Risk Behavior Survey (YRBS): risk behavior prevalence estimates. Eval Rev 2010 Apr;34(2):137-153. [doi: 10.1177/0193841X10362491] [Medline: 20234000]
41. De Leeuw E, De Heer W. Trends in household survey nonresponse: A longitudinal and international comparison. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, editors. Survey Nonresponse. New York, NY: John Wiley & Sons, Inc; 2002:41-54.
42. Connelly NA, Brown TL, Decker DJ. Factors affecting response rates to natural resource - focused mail surveys: Empirical evidence of declining rates over time. Society & Natural Resources 2003;16(6):541-549. [doi: 10.1080/08941920309152]
43. Leisenring WM, Mertens AC, Armstrong GT, Stovall MA, Neglia JP, Lanctot JQ, et al. Pediatric cancer survivorship research: experience of the Childhood Cancer Survivor Study. J Clin Oncol 2009 May 10;27(14):2319-2327. [doi: 10.1200/JCO.2008.21.1813] [Medline: 19364957]

Abbreviations

BCCSS: British Childhood Cancer Survivor Study
CCSS: Childhood Cancer Survivor Study
CI: confidence interval
IQR: interquartile range
OR: odds ratio
SD: standard deviation

Edited by G Eysenbach; submitted 20.12.10; peer-reviewed by Z Yan, I mohd amin; comments to author 22.04.11; revised version received 14.06.11; accepted 14.06.11; published 29.09.11.

Please cite as:
van den Berg MH, Overbeek A, van der Pal HJ, Versluys AB, Bresters D, van Leeuwen FE, Lambalk CB, Kaspers GJ, van Dulmen-den Broeder E
Using Web-Based and Paper-Based Questionnaires for Collecting Data on Fertility Issues Among Female Childhood Cancer Survivors: Differences in Response Characteristics
J Med Internet Res 2011;13(3):e76
URL: http://www.jmir.org/2011/3/e76/
doi:10.2196/jmir.1707
PMID:21955527

Original Paper

3D Visualization as a Communicative Aid in Pharmaceutical Advice-Giving over Distance

Martin Östlund[1,2], MSc, PhLic; Nils Dahlbäck[3], PhD; Göran Ingemar Petersson[1], MD, PhD

[1]eHealth Institute, Linnaeus University, Kalmar, Sweden
[2]School of Computer Science, Physics and Mathematics, Linnaeus University, Kalmar, Sweden
[3]Department of Computer and Information Science, Linköping University, Linköping, Sweden

Corresponding Author:
Martin Östlund, MSc, PhLic
eHealth Institute
Linnaeus University
Kalmar, SE-39182
Sweden
Phone: 46 480 497726
Fax: 46 480 497110
Email: martin.ostlund@lnu.se

Abstract

Background: Medication misuse results in considerable problems for both patient and society. It is a complex problem with many contributing factors, including timely access to product information.

Objective: To investigate the value of 3-dimensional (3D) visualization paired with video conferencing as a tool for pharmaceutical advice over distance in terms of accessibility and ease of use for the advice seeker.

Methods: We created a Web-based communication service called AssistancePlus that allows an advisor to demonstrate the physical handling of a complex pharmaceutical product to an advice seeker with the aid of 3D visualization and audio/video conferencing. AssistancePlus was tested in 2 separate user studies performed in a usability lab, under realistic settings and emulating a real usage situation. In the first study, 10 pharmacy students were assisted by 2 advisors from the Swedish National Co-operation of Pharmacies' call centre on the use of an asthma inhaler. The student-advisor interview sessions were filmed on video to qualitatively explore their experience of giving and receiving advice with the aid of 3D visualization. In the second study, 3 advisors from the same call centre instructed 23 participants recruited from the general public on the use of 2 products: (1) an insulin injection pen, and (2) a growth hormone injection syringe. First, participants received advice on one product in an audio-recorded telephone call and for the other product in a video-recorded AssistancePlus session (product order balanced). In conjunction with the AssistancePlus session, participants answered a questionnaire regarding accessibility, perceived expressiveness, and general usefulness of 3D visualization for advice-giving over distance compared with the telephone and were given a short interview focusing on their experience of the 3D features.

Results: In both studies, participants found the AssistancePlus service helpful in providing clear and exact instructions. In the second study, directly comparing AssistancePlus and the telephone, AssistancePlus was judged positively for ease of communication ($P = .001$), personal contact ($P = .001$), explanatory power ($P < .001$), and efficiency ($P < .001$). Participants in both studies said that they would welcome this type of service as an alternative to the telephone and to face-to-face interaction when a physical meeting is not possible or not convenient. However, although AssistancePlus was considered as easy to use as the telephone, they would choose AssistancePlus over the telephone only when the complexity of the question demanded the higher level of expressiveness it offers. For simpler questions, a simpler service was preferred.

Conclusions: 3D visualization paired with video conferencing can be useful for advice-giving over distance, specifically for issues that require a higher level of communicative expressiveness than the telephone can offer. 3D-supported advice-giving can increase the range of issues that can be handled over distance and thus improve access to product information.

(J Med Internet Res 2011;13(3):e50) doi:10.2196/jmir.1437

KEYWORDS

Pharmaceutical instruction; 3D visualization; distance communication

Introduction

People are not very good at using their medications correctly. It has been shown that only about 50% of all patients manage to follow the instructions they have been given to the full [1-3]. The effects of this failure to comply with given instructions are serious. Nearly half of all medication-related admissions to hospital are directly related to the patients not following their prescriptions [4]; and 8%–10% of the total number of hospital admissions are directly related to incorrect use of medication [5]. Many different strategies have been attempted to improve on the situation. In a comprehensive review of intervention studies, Haynes et al [6] list the following interventions as the most promising to improve compliance: providing more and better instructions, fostering a good relationship between the health care professional and the patient, scheduling regular follow-up meetings, and supplying the patient with appropriate reminders. The list shows that patient–caregiver communication is very important in fostering good use of medication.

Many types of communication tools have been tried out in the health care domain, and experience shows that the technology does not need to be advanced to be helpful. Indeed, many positive results have been achieved with simple solutions. For instance, many studies have investigated the usefulness of email-type messaging between patient and physician [7-11] and, of course, the telephone is a well-established tool for patient–caregiver consultation [12,13]. More sophisticated communication technologies such as video conferencing and groupware solutions have also been used, for example to review patient data [14,15] or to discuss medical imaging data such as from scans [16]. Results have been positive and, although initial costs can be high for the more high-tech variants, distance communication solutions of this kind have generally been shown to be cost effective [17-19].

In this paper we explore the usefulness of 3-dimensional (3D) visualization technology as a communication tool. 3D technology has been put to good use in medicine, for example in the exploration of scanning data [16], laparoscopic surgery training [20], and phobia therapy [21]. However, it has not been used to any great extent as an instrument of communication. In other domains, most prominently that of computer-aided design [22], it has been used as a communication enhancer for some time, for example to support remote collaboration between geographically dispersed teams [23-25], to review designs over distance [26-28], and to offer training and support to remotely located service personnel [29].

One exception where 3D technology has been used to support communication in health care is with the use of 3D virtual worlds such as Second Life. Second Life has been used in medical training to provide a virtual meeting place for health-related dialogue and for disseminating health care information [30,31]. This is an interesting exploration into the realm of possibilities that 3D technologies offer, but it is not a practical solution for most types of caregiver–patient communication.

In the health care domain, the preference seems to have been either for technical solutions that are very simple, such as email-based messaging systems, or for those that are very advanced, such as state-of-the-art video conferencing systems or virtual worlds. The problem is that *simple* solutions might reach many people, but the expressive power is often limited. *Advanced* solutions have more expressive power, but the number of people who can be reached may be limited because they might not have the required hardware or software, or they find it difficult to use. The challenge is to find a solution that gives room for expressiveness but also can be made available for a broad audience.

The objective of the present exploratory descriptive study was to examine the value of 3D visualization as a communicative aid in pharmaceutical advice-giving over distance. We created a Web-based application named AssistancePlus that uses interactive 3D representations to demonstrate handling instructions for complex pharmaceutical products. For the application to be useful in practical terms, it needs to be easy to access and easy to use for members of the general public. Thus, the application also presents a proof-of-concept that an advanced technology such as 3D visualization can be packaged and presented in such a way that the resulting application is accessible to this group. We carried out two evaluation studies to investigate its practical usefulness for pharmaceutical advice-giving. Comparisons were made with telephone communication and face-to-face interaction in terms of expressiveness and accessibility.

Methods

Materials

AssistancePlus is a communication tool that uses co-manipulation of 3D representations of products, audio/video communication, and co-browsing to support advice-giving on pharmaceutical products over distance. Building on Clark's theory of communication and specifically his theory of common ground [32-34], we believe that using interactive virtual 3D objects together with audio/video communication will support communicative efficiency by creating a sense of *co-presence* between the patient and advisor. Our hope is that this will translate into a natural feel to the communication, as well as providing a high level of expressiveness.

AssistancePlus has been tailored to meet the specific needs of the 2 parties involved in the advice-giving interchange: the *advice seeker* and the *advisor*. The advice seeker is typically a novice and a beginner—a novice to the content domain and a beginner at using the communication tools at hand. The advisor is an expert and a professional—an expert in the content domain and a professional at using the communication tools (see [35,36] for a similar description of the advice seeker–advisor relationship). Due to the different nature of the 2 roles, their needs and wants are different. The advice seeker wants a service that is easy to access and easy to use. The advisor wants a service that allows him or her to communicate efficiently and effectively. This is achieved easily enough by providing them with different versions of the user interface. However, although it is true that most of the information in an advice-giving setting flows from the person giving advice to the person receiving advice, it is also necessary to have a flow in the other direction.

No advice-giving session should be considered complete until the advice seeker has *internalized* the given information—that is, not only has received the information, but also has understood it and is motivated to act in accordance with what has been learned (see, for example, Clark [32,33]). It must be the responsibility of the advisor in the role of expert professional to make sure that this is so, and to do this the advisor needs feedback from the advice seeker. Therefore, it is important that advice seekers also can be expressive so that they can signal or demonstrate their understanding back to the advisor. The challenge here is to provide the right level of expressiveness for each party and, for the advice seeker, to find an appropriate balance between ease of use and expressive power. Details on how this challenge was approached in the design of AssistancePlus can be found in Multimedia Appendix 1, together with a technical description of the service.

Figure 1 shows a screenshot of AssistancePlus with the first version of the 3D player being used to demonstrate handling instructions for an asthma inhaler. The main features of AssistancePlus are (1) co-manipulation of shared 3D objects representing pharmaceutical products, (2) two-way text and audio/video communication, and (3) co-browsing of webpages with synchronized page scrolling and remote cursors

Figure 1. AssistancePlus, a tool for 3-dimensional (3D) -supported advice-giving on pharmaceutical products over distance

Evaluation Studies

The value of 3D-supported advice-giving was assessed in 2 separate studies, both set up in the same way. Study participants (advice seekers) were exposed to AssistancePlus in a realistic usage situation after which they were questioned about their experience. The first study was explorative and probed the general usefulness of 3D visualization for advice-giving purposes. The second study was focused on the details of how 3D visualization contributed to raising the level of expressive power. A telephone session was included in the second study to provide an explicit (and controlled) reference point with which to compare their experience with AssistancePlus. The type of questions that AssistancePlus is appropriate for was also examined. A large amount of data was collected in the 2 studies, of which a selection is presented here (the full range of the data is presented in [37]).

User Study I: Simple 3D Player

The objective of this substudy was to broadly explore the issue of how 3D visualization can be used as a tool for advice-giving on pharmaceutical products.

Materials

The AssistancePlus application, with version 1 of the 3D player, allowed advice seekers to communicate with an advisor over distance using 2-way audio/video communication and co-browsing of webpages (see Multimedia Appendix 1 for a description of AssistancePlus and the 3D player). We made 2 webpages available for the advice seekers and adviser to browse together: (1) a text-based page with general information about the product (side effects, storage directions, etc), and (2) a page with a 3D model of an asthma inhaler (Ingelheim; Boehringer Ingelheim GmbH, Ingelheim, Germany) (Figure 2).

Participants and Recruitment

Playing the role of advice seeker were a group of 10 pharmacy students (9 women and 1 man; age range 22–40 years). Pharmacy students, with obvious preknowledge of the subject area, were selected because their professional insight into the subject area was considered helpful at this explorative stage (the pharmacy students had no connection with the AssistancePlus project and received no reward for participating in the study). The students assumed the role of a relative of an asthma patient (the role of relative is relevant because many people are involved in their relations' medication, and it was a more realistic role for the pharmacy students to play than patient). The role of advisor was assumed by 2 licensed pharmacists from the National Cooperation of Swedish Pharmacies' call centre (both women; age range 40–50 years, each with several years' experience with telephone consultation).

Procedure

The study was performed in a usability lab using standard computer equipment and software (Web browser). Advice seekers and advisors were in separate rooms and did not meet each other before or after the trial. Advice seekers, who were assigned the task of trying to find out as much as possible about the inhaler, interacted through AssistancePlus for 15 minutes with a randomly assigned advisor. No special instructions were given on how much time to spend on any specific topic. The first author conducted individual in-depth, semistructured interviews on a separate occasion, 1 hour each for the pharmacy students and 2 hours each for the advisors.

Data Collection

The students and advisors were filmed on video to allow for analysis of the advice seeker–advisor interaction. The video on the students' side recorded keyboard and mouse use as well as the screen; and the video on the advisor side recorded only the screen. The audio from the interview was recorded, transcribed, and summarized.

A semistructured interview was administered on a separate occasion. The interview covered the following areas: general impressions; specific impressions concerning audio and video features, co-browsing features, and 3D; locus of control (self/advisor); level of activeness (self/advisor); usefulness of service; and target audience.

Figure 2. The Ingelheim asthma inhaler displayed using the first version of the 3-dimensional player

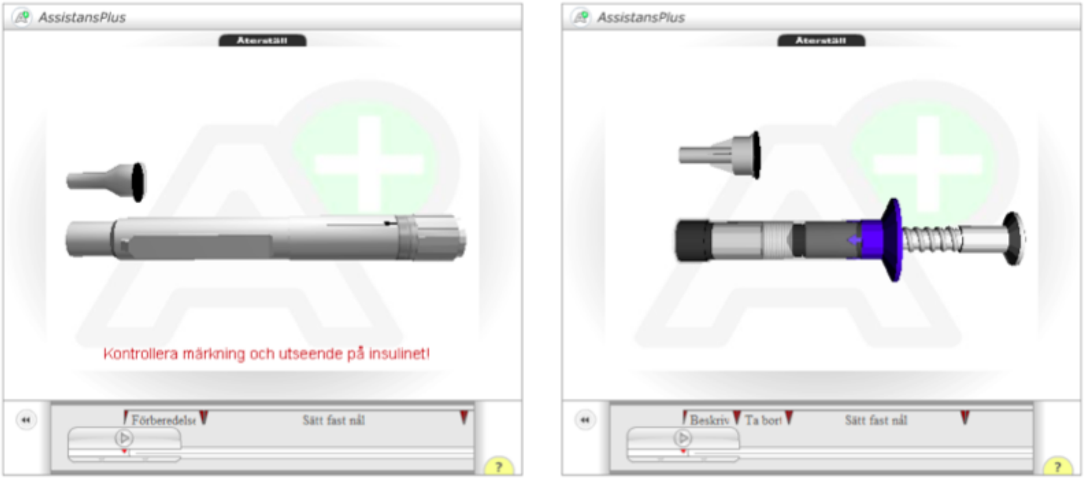
User Study II: Advanced 3D Player and Comparison With Telephone

The objective of this substudy was to compare the experience of using AssistancePlus with that of the telephone; to identify the type of questions for which use of AssistancePlus is appropriate; and to investigate the dynamics of how 3D visualization contributes to communicative expressiveness.

Materials

As in the first study, AssistancePlus, with the second version of the 3D player, was used to allow advice seekers to communicate with an advisor over distance with 2-way audio/video communication and co-browsing (Multimedia Appendix 1). As in the first study, 2 webpages were available to the participants: (1) a text-based page with general information (side effects, storage directions, etc), and (2) a page with a 3D model, this time either of a pen injector for insulin glargine (Lantus OptiSet, sanofi-aventis, Paris, France) or of a disposable syringe for human growth hormone equivalent (Genotropin MiniQuick, Pfizer, New York, USA). Figure 3 shows screenshots of their respective 3D representations as displayed with the second version of the 3D player (Multimedia Appendix 1; see also Multimedia Appendix 1 for additional screenshots of the 3D sequences used).

Figure 3. The 2 pharmaceutical products used in the second user study displayed using the second version of the 3-dimensional player: Lantus OptiSet (left) and Genotropin MiniQuick (right)



Kontrollera märkning och utseende på insulinet!

Participants and Recruitment

For the role of advice seeker, 23 participants were recruited from various night-school classes: 13 women and 10 men with varied backgrounds, interests, and ages (18–74) (participants received a nominal gift in the form of a lottery ticket with a value of less than $10). As in the first study, the participants played the role of a relative, this time of a patient with either diabetes (Lantus OptiSet) or a growth hormone deficit (Genotropin MiniQuick). The role of advisor was assumed by 3 licensed pharmacists from the National Cooperation of Swedish Pharmacies' call centre (2 women, ages 36 and 40 years; and 1 man, age 50 years), each with several years' experience with telephone consultation.

Procedure

The advice seekers were instructed to try to find out as much as possible about the 2 products, first using the telephone for one product and then, on a separate occasion, using AssistancePlus for the other product. The telephone condition was first for all participants. The telephone session was conducted from home, while the AssistancePlus session was carried out in a usability lab. Advice seekers and advisors were in separate rooms and did not meet each other before or after the trial. Each session lasted 10 minutes. The order of the products was balanced, with 12 advice seekers starting with Lantus OptiSet and 11 starting with Genotropin MiniQuick. The advisors were randomly assigned with a different advisor in the AssistancePlus session and the telephone session. After the AssistancePlus session, the advice seekers filled in a paper-based questionnaire and then participated in a 20-minute semistructured interview led by the first author.

Data Collection

The audio from the telephone sessions was recorded. The AssistancePlus sessions were filmed on video and the screen image was recorded. Questionnaire data were collected and the audio from the interviews was recorded. The questionnaire covered the following areas: background demographic variables, general attitudes toward computers and the Web, impressions of AssistancePlus and its 3D features, comparisons between the telephone and the AssistancePlus session, and which channel (medium) they would use to inform themselves about pharmaceutical products of varying complexity. In the comparison between the telephone and the AssistancePlus session, participants were asked to distribute 10 points between the telephone and AssistancePlus for each of the following 7 factors: ease of use, ease of communication, sense of personal contact, explanatory power, level of understanding, level of trust, and efficiency. For example, if AssistancePlus is given a rating of 6 on ease of use, the rating for the telephone for this factor must be 4, to make a total of 10. (The factors were gleaned from a discussion seminar held with representatives from the pharmaceutical industry, health care representatives, and advice-giving professionals [38].) Participants were also asked which channel they would choose to inform themselves about products of varying complexity level. In addition to telephone and AssistancePlus, they were able to choose from pharmacy store and the Internet, which were added to provide a more complete range of alternatives (note that the participants' acquaintance with AssistancePlus and telephone channels was controlled in the study, but that with the Internet and the pharmacy store channels were based on uncontrolled personal experience). Participants were asked to distribute 10 points among these 4 channels for each of the following 4 product types: simple nonprescription medications, simple prescription medications, complex prescription medications, and complex medications requiring handling. The interview questions focused on their experience of the 3D features and covered the following areas: perceived level of expressiveness, experience of 3D content, sense of presence, level of activeness, and locus of control.

Statistical Analysis

For each of the 7 factors on which telephone and AssistancePlus were compared, a set of preference scores were calculated by subtracting the rating given to the telephone from that given to AssistancePlus (a positive preference score thus indicating a preference for AssistancePlus and a negative score indicating a preference for the telephone). Wilcoxon signed rank test was used to assess the statistical significance of the preference scores. Friedman tests were used to assess whether the complexity level of the product affected the choice of communication channel. Descriptive statistics were calculated for other survey data.

Results

Findings From User Study I: Simple 3D Player

Being accustomed to telephone consultation, the 2 advisors could make an experience-based comparison between using the telephone and using AssistancePlus to give advice. The advisors welcomed the power of expression the 3D content provided. They felt that this extra expressiveness made it possible to handle more complex issues than could be dealt with using the telephone. In particular, the advisors valued the potential for increased precision, for example that they could rotate and zoom in on the model and then use their remote cursor to point to the exact detail they were referring to in their spoken communication.

The participants in the role of advice seekers were also pleased with their experience of AssistancePlus, both in general and with the 3D features specifically. All of the participants found the 3D features useful as a pedagogical aid and found the experience informative. The reasons they gave varied in wording, but centered on issues of clearness and distinctiveness. Several of the participants (7/10) also described the interaction as similar in feel to that of a face-to-face conversation concerning both level of expressiveness and sense of personal contact.

One somewhat surprising finding was that the activity level of the advice seekers was seemingly not very high. None of them attempted to use the 3D controls themselves and only 1 used the remote cursor purposefully. Reasons for not being more active given in the interviews were that they felt they might disrupt the advisor (6/10 participants), that the content “belonged” to the advisor (5/10), and that they felt a bit intimidated by the many buttons on the 3D player control panel (3/10). However, the main reason was that they did not feel it necessary to be more active than they were. They were active with their voice and, if they wanted to see a different view of the 3D model, they could simply ask the advisor. This said, the advice seekers wanted to keep the feature (9/10). It seems that knowing that they had the option to control the model made them feel more involved in the advice-giving process (even if they did not actually do this). Similarly, the advisors felt that, while the vocal feedback the advice seekers gave was sufficient for the advisors to judge understanding in the present situation, in other situations it might be helpful having the advice seekers control the model themselves to demonstrate what they have learned.

Findings From User Study II: Advanced 3D Player and Comparison With Telephone

Figure 4 shows the median value, interquartile range, and total range of the preference scores that were calculated by subtracting telephone ratings from AssistancePlus ratings for each of the 7 factors on which telephone and AssistancePlus were compared (see Methods section). All the preference scores were positive in AssistancePlus’ favor, and all scores except *ease of use* were shown to be statistically significant (Table 1). For *ease of communication*, *personal contact*, and *trust*, the magnitude was moderate; for *explanatory power*, *understanding*, and *efficiency* it was more pronounced. The results were corroborated in the interviews, where participants described their experience with AssistancePlus as being more clear and expressive, being more personal and present, and giving a deeper sense of understanding. Also, several participants described the feel of the interaction when using AssistancePlus as being more similar to a face-to-face meeting than that of a telephone conversation. It should be noted, though, that both the questionnaire and the interview were administered in conjunction with the AssistancePlus session, which might have introduced a positive bias in AssistancePlus’ favor, being the more recently experienced.

Figure 4. Preference scores comparing AssistancePlus and telephone. Positive scores (above the 0 baseline) indicate that AssistancePlus received a higher rating. The vertical lines show the total range, the filled blue sections show the interquartile ranges, and the horizontal red lines show the median values

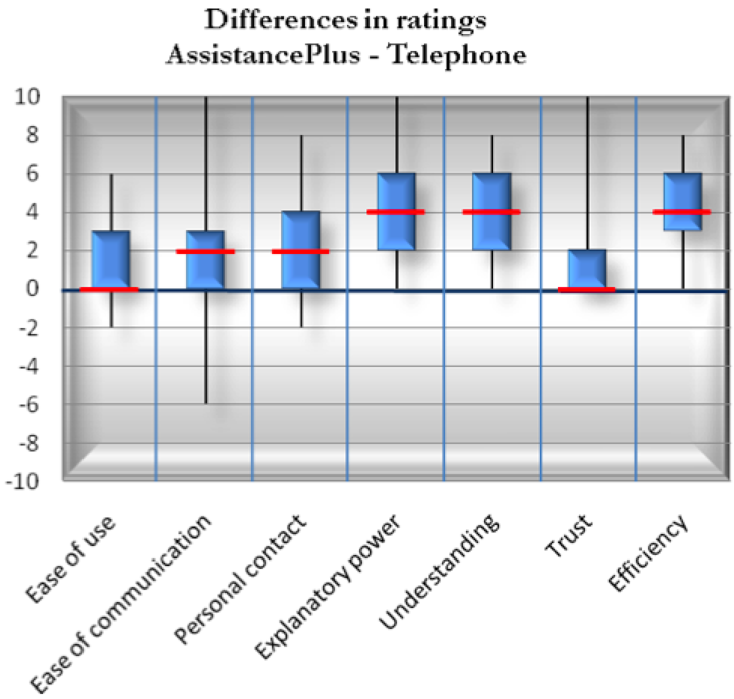


Table 1. Preference scores comparing AssistancePlus and telephone (N = 23)

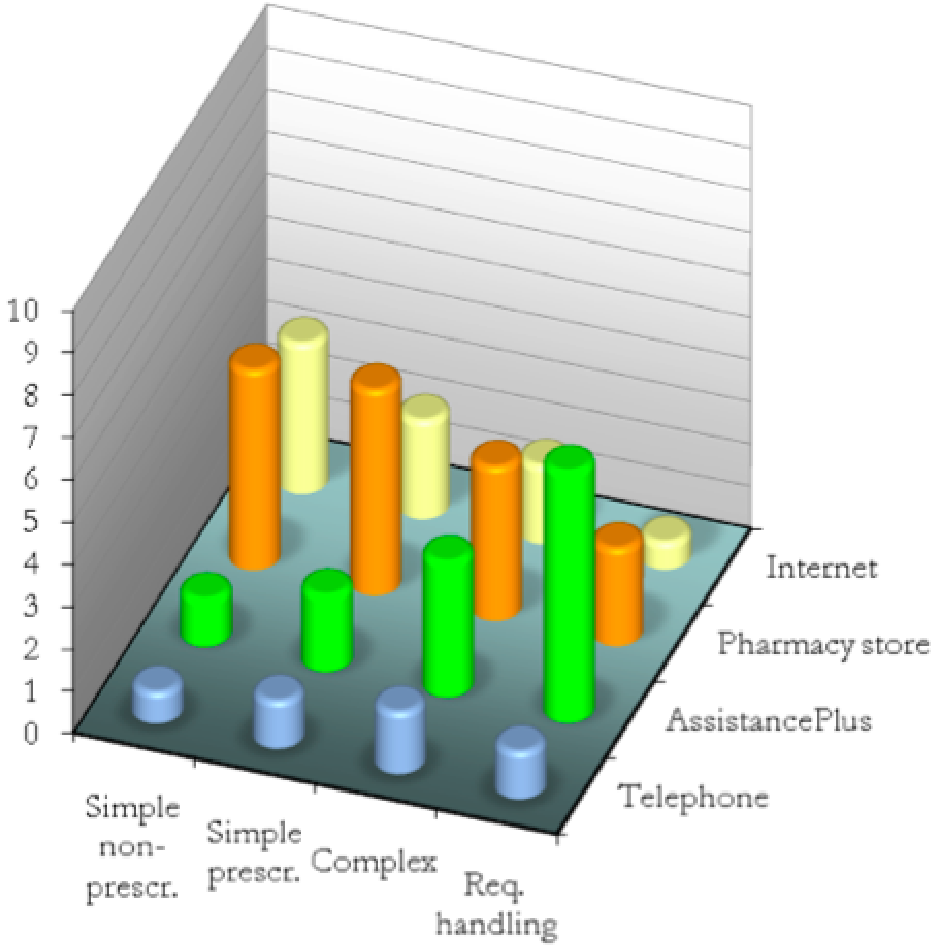
Factor	Median (quartile deviation)	Mean (SD)	Z score (Wilcoxon signed-rank test)[a]	2-tailed[b] *P* value
Ease of use	0 (3)	1.04 (2.69)	–1.596	.110
Ease of communication	2 (3)	2.26 (2.58)	–3.448	.001
Personal contact	2 (4)	2.26 (2.43)	–3.454	.001
Explanatory power	4 (4)	4. 22 (2.63)	–4.042	< .001
Understanding	4 (4)	3.57 (2.63)	–3.758	< .001
Trust	0 (2)	1.48 (2.78)	–2.401	.016
Efficiency	4 (3)	3.96 (2.06)	–4.078	< .001

[a] Based on negative ranks for [AssistancePlus score] – [telephone score].

[b] H_0: no difference between scores.

Figure 5 displays the results of the preference ratings given for the 4 types of products and 4 different channels. On type of product (front to back in Figure 5), the variation was found to be significant for all product categories except *complex medications* (Table 2). On type of channel (left to right in Figure 5), *AssistancePlus* and *Internet* differed significantly, but *telephone* and *pharmacy store* did not. The results indicate that the advice seekers prefer a simple channel for simpler questions and a more powerful channel for more complex questions—that is, that they make their choice with consideration to the complexity of the issue at hand.

Figure 5. Relative preferences among information channels for different levels of product complexity (prescr = prescription, req = requiring)



Somewhat surprisingly, though, the advice seekers did *not* rate the face-to-face setting—ostensibly the most expressive channel among the alternatives—as the preferred channel for the most complex type. In fact, the results instead showed an inverse relationship between the *pharmacy store* and product complexity. The reason for this was revealed in the interviews. The participants saw the meeting in the pharmacy store not as an expressive and personal face-to-face interaction, but as a fleeting encounter in a nonprivate, noisy, and stressful environment. The telephone received surprisingly low ratings for all categories indicating that the telephone is simply not a popular alternative when seeking medical advice.

Table 2. Friedman tests on preferences among information channels (N = 23)

Type of product		Telephone	AssistancePlus	Pharmacy store	Internet	Friedman test (c^2)[a]	*P* value
Simple nonprescription	Mean	0.70	1.09	4.70	3.52	22.35	< .001
	SD	1.06	1.20	3.42	3.25		
Simple prescription	Mean	1.09	1.78	4.83	2.30	17.15	.001
	SD	1.35	1.93	2.95	2.40		
Complex	Mean	1.43	3.17	3.61	1.83	7.608	.06
	SD	1.56	2.98	2.66	1.83		
Requiring handling	Mean	1.09	5.91	2.27	0.61	43.05	< .001
	SD	1.31	2.25	1.86	0.89		
Friedman test (c^2)[a]		3.329	43.01	5.949	24.53		
	P value	.36	.001	.11	.001		

[a] Two-tailed; df = 3 for all tests.

The level of activity on the part of the advice seekers was very low in the first study, with none of them even attempting to use the 3D controls. This behavior changed in the second study, where they were more active. Typically, the actions they performed were simple ones such as using their remote cursor to point to details on the 3D model (16/23, 70% of participants), moving the playhead to go to a specific point in the animation (2/23, 9%), and starting and stopping playback (3/23, 13%). Only 1 participant used the click-and-drag controls to rotate and zoom in on the model. However, the preference for simpler action was not because they did not know how to perform these more advanced actions. As in the first study, they simply found it easier to have the advisor do it for them. Still, the participants were very positive about having the option to be able to control the 3D content themselves (N = 23, mean [SD] 5.42 [0.90], range 1–6) and wanted to keep the feature. The advisors were also pleased that the advice seekers had the option to control the 3D content themselves.

Discussion

We have shown that 3D visualization techniques paired with video conferencing can be useful in supporting advice-giving over distance and can be used even in low-bandwidth settings. AssistancePlus offers a proof-of-concept of how 3D visualization can be used to extend the range of questions about pharmaceutical products that can be handled over distance. Distance communication cannot, and should not, be used to replace the physical meeting between patient and caregiver, but in many situations distance communication can save time and resources and give patients quicker and easier access to the information they need. Haynes et al [6] listed more and better information and fostering a good relationship with caregivers as the 2 top interventions to improve compliance. We believe that a service such as AssistancePlus can contribute to both of these by making it easier for patients or clients to communicate with their caregivers. We also found it interesting that in some cases our participants actually preferred AssistancePlus to face-to-face communication, as this usually takes place at the pharmacy. This illustrates our belief that Internet-based communication not only should be seen as a substitute for face-to-face interaction but also can in some cases offer a better alternative.

Choice of Channel Based on Question Content

Advice seekers and advisors alike found AssistancePlus useful as a pedagogical aid for advice-giving and particularly appreciated the clarity and exactness brought to instructions by the 3D features. This does not mean that AssistancePlus is the right choice for all types of advice-giving situations. The results from the second study indicate that information seekers do not just go for the most expressive channel, but differentiate their choice. Perceived level of expressiveness is one factor affecting choice, but other factors are also important. For instance, the pharmacy store was an unpopular alternative for the reason that it was seen as lacking in privacy. If we try to isolate the issue of level of expressiveness here, it seems that participants preferred the channel that offers not too much, not too little, but just the right level of expressiveness. The implication is that we should provide advice seekers not with a single catch-all channel, but with a selection so that they can choose the right channel for the question at hand. Of course, the downside of providing choice is that it places a burden on the advice seeker to choose among an ever-increasing range of channels. We believe that a major future challenge will be to provide guidance services to help advice and information seekers to find not only the right information, but also the right channel.

Empowering the Advice Seeker

In the advice-giving situation, most of the communication naturally flows from the person giving the advice to the person receiving the advice, but we have argued that it is important that the advice seeker provide feedback so that the advisor can properly judge understanding and motivation. Here, care must be taken not to overwhelm or intimidate the advice seeker, who perhaps for the first time is communicating over distance using advanced technology. Our solution is to provide self-use of, for example, the 3D controls as an option, something that one can do, but does not have to do.

An interesting—and we believe important—finding in this area was that advice seekers did feel more involved when presented with the option of using the 3D controls, but this sense of involvement was not the result of actual action. Instead it seemed to stem from the awareness that they could be active. Both those who actively used the 3D controls and those who did not felt empowered by knowing they could control the 3D content (theoretical support for this empowerment effect can be found in Clark's theory of common ground [32-34]). This finding suggests that it is positive to present advice seekers with more functionality than they are expected to use. However, we hypothesize that the empowerment effect will occur only if the functionality is perceived as something they can manage. This means that either the functionality needs to be simple enough to be self-explanatory, or it can be easily explained by way of demonstration, which is the case with the playback and 3D controls used in the studies. We believe that this finding, provided that it can be corroborated in further studies, can have important consequences for the design of online communication services of the kind presented here.

Adapting to the Specific Communication Situation

In this paper we have exemplified how 3D visualization can be used in health communication for a specific advice-giving situation. It is in the custom tailoring to the specific communication setting and to its actors that the service gains its edge, but we believe that the concept it presents—ease of access with low technical and usability thresholds and a level of expressiveness tailored to the demands of the specific communication situation—is relevant for any area of health care where there is a need for improving access to information and for bringing patients and caregivers closer to each other.

Limitation of Study and Cost of Service

It must be noted that both presented studies used convenience samples and were performed in usability laboratories. There may also be a novelty effect, and it remains to be seen how a service such as AssistancePlus would work in a real-world setting with real patients and caregivers who might react differently from the test participants in our studies. That said, the results are promising and we believe AssistancePlus can provide both an alternative to the telephone when it is not expressive enough and a complement to the face-to-face interaction when a physical meeting is not convenient.

We have not performed a detailed analysis of the cost of setting up and running a service such as AssistancePlus, but it is our assessment that it should be comparable with that of setting up a telephone-based call centre. Of course, the easiest way to get a service like AssistancePlus up and running would be to add it to the list of services offered by an existing call centre. Since all the functionality demonstrated in the AssistancePlus application can be implemented using open source free software, this means development and running costs can be kept down and, while the production of the 3D content is an extra cost for each product that is added, it is a cost of moderate size.

Conclusion

We have shown that 3D visualization techniques paired with video conferencing can be useful in supporting advice-giving over distance and can be used even in low-bandwidth settings. We also showed that 3D visualization brings clarity and exactness to given instructions and contributes to communicative efficiency. It is a technology that can be made available to the general public and, if implemented with proper consideration to the role and character of the advice seeker, will be both easy to access and easy to use and, in some cases, even preferred to face-to-face communication.

Acknowledgments

The AssistancePlus project has been sponsored through the eHealth Institute by the National Cooperation of Swedish Pharmacies–Apoteket AB, Linnaeus University (previously University of Kalmar), the Regional Council in Kalmar County, the Kalmar County Council, and the Municipality of Kalmar. The eHealth Institute is part of Linnaeus University. Thanks to Daniel Fredriksson, who helped with the practical construction of the 3D models, and Lina Hellström, who advised on the content of the 3D sequences.

Conflicts of Interest

The first author has founded a company – Instant Connect Sweden AB – for the purpose of exploring various commercial applications of the described communication technology.

Multimedia Appendix 1

Technical description of AssistancePlus and screenshots of the second version of the 3D player displaying animation sequences for Lantus OptiSet and Genotropin.

[PDF file (Adobe PDF File), 487 KB - jmir_v13i3e50_app1.pdf]

Multimedia Appendix 2

Survey from user study II in English translation (original is in Swedish).

[PDF file (Adobe PDF File), 1025 KB - jmir_v13i3e50_app2.pdf]

References

1. Arbetsgruppen för bättre läkemedelsanvändning (ABLA II). Mindre sjukdom och bättre hälsa genom ökad följsamhet till läkemedelsordinationerna. Stockholm, Sweden: ABLA II; 2001. URL: http://www.nepi.net/res/dokument/ablaiimindresjukdomochbattrehalsa.pdf [accessed 2011-06-27] [WebCite Cache ID 5zlE8A1E9]
2. Läkemedelsindustriförbundet LIF. FAKTA 2006 Läkemedelsmarknaden och hälso- och sjukvården. Stockholm, Sweden: LIF; 2006.
3. Sackett DL, Snow JC. The magnitude of adherence and nonadherence. In: Haynes RB, Taylor DW, Sackett DL, editors. Compliance in Health Care. Baltimore, MD: Johns Hopkins University Press; 1979:11-22.
4. Bergman U, Wiholm BE. Drug-related problems causing admission to a medical clinic. Eur J Clin Pharmacol 1981;20(3):193-200. [Medline: 7286037]
5. Manasse HR. Toward defining and applying a higher standard of quality for medication use in the United States. Am J Health Syst Pharm 1995 Feb 15;52(4):374-379. [Medline: 7757862]
6. Haynes RB, Yao X, Degani A, Kripalani S, Garg A, McDonald HP. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2005(4):CD000011. [Medline: 16235271]
7. Couchman GR, Forjuoh SN, Rascoe TG, Reis MD, Koehler B, Walsum KL. E-mail communications in primary care: what are patients' expectations for specific test results? Int J Med Inform 2005 Jan;74(1):21-30. [doi: 10.1016/j.ijmedinf.2004.08.005] [Medline: 15626633]
8. 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]
9. Liederman EM, Lee JC, Baquero VH, Seites PG. The impact of patient-physician Web messaging on provider productivity. J Healthc Inf Manag 2005;19(2):81-86. [Medline: 15869217]
10. Liederman EM, Morefield CS. Web messaging: a new tool for patient-physician communication. J Am Med Inform Assoc 2003;10(3):260-270. [doi: 10.1197/jamia.M1259] [Medline: 12626378]
11. Umefjord G, Sandström H, Malker H, Petersson G. Medical text-based consultations on the Internet: a 4-year study. Int J Med Inform 2008 Feb;77(2):114-121. [doi: 10.1016/j.ijmedinf.2007.01.009] [Medline: 17317292]
12. Bunn F, Byrne G, Kendall S. The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review. Br J Gen Pract 2005 Dec;55(521):956-961. [Medline: 16378566]
13. Clark RA, Inglis SC, McAlister FA, Cleland JGF, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007;334(7600):942. [doi: 10.1136/bmj.39156.536968.55] [Medline: 17426062]
14. LeRouge C, Garfield MJ, Hevner AR. Quality attributes in telemedicine video conferencing. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences. Hawaii: January 7-10. Washington, DC: IEEE; 2002.
15. Wootton R, Craig J, Patterson V. In: Wootton R, Craig J, Patterson V, editors. Introduction to Telemedicine. King of Prussia, PA: Rittenhouse Book Distributors; 2006.
16. Hartvigsen G, Johansen MA, Hasvold P, Bellika JG, Arsand E, Arild E, et al. Challenges in telemedicine and eHealth: lessons learned from 20 years with telemedicine in Tromsø. Stud Health Technol Inform 2007;129(Pt 1):82-86. [Medline: 17911683]
17. Chang T, Lee J, Wu S. The telemedicine and teleconsultation system application in clinical medicine. Conf Proc IEEE Eng Med Biol Soc 2004;5:3392-3395. [doi: 10.1109/IEMBS.2004.1403953] [Medline: 17271012]
18. Clemensen J, Larsen SB. Cooperation versus coordination: using real-time telemedicine for treatment at home of diabetic foot ulcers. J Telemed Telecare 2007;13(3):32-35. [doi: 10.1258/135763307783247428]
19. Koch S. Home telehealth: current state and future trends. Int J Med Inform 2006 Aug;75(8):565-576. [doi: 10.1016/j.ijmedinf.2005.09.002] [Medline: 16298545]
20. Sutherland LM, Middleton PF, Anthony A, Hamdorf J, Cregan P, Scott D, et al. Surgical simulation: a systematic review. Ann Surg 2006 Mar;243(3):291-300. [doi: 10.1097/01.sla.0000200839.93965.26] [Medline: 16495690]
21. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J Behav Ther Exp Psychiatry 2008 Sep;39(3):250-261. [doi: 10.1016/j.jbtep.2007.07.007] [Medline: 17720136]
22. Li WD, Ong SK, Nee AYC. Integrated And Collaborative Product Development Environment: Technologies And Implementations (Series on Manufacturing Systems and Technology). Hackensack, NJ: World Scientific Publishing Company; 2006.
23. Fuh JYH, Li WD. Advances in Collaborative CAD: The state-of-the-art. Computer-Aided Design 2005;37:571-581. [doi: 10.1016/j.cad.2004.08.005]
24. Rosenman MA, Smith G, Ding L, Marchant D, Maher ML. Multidisciplinary design in virtual sorlds. In: Computer Aided Architectural Design Futures 2005, Vienna, Austria. New York, NY: Springer; 2005.
25. Wang L, Shen W, Xie H, Neelamkavil J, Pardasani A. Collaborative conceptual design: state of the art and future trends. Computer-Aided Design 2002;34:981-996. [doi: 10.1016/S0010-4485(01)00157-9]
26. Min T, Shang-Ching C, Jin-Xiang D. Collaborative virtual environment for feature based modeling. In: Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in industry. New York, NY: ACM; 2004.

27. Pekkola S. Critical approach to 3D virtual realities for group work. In: NordiCHI 2002. 2002 Presented at: Second Nordic Conference on Human-Computer Interaction; October 19-23; Aarhus, Denmark.
28. Zhang S, Shen W, Ghenniwa H. A review of Internet-based product information sharing and visualization. Computers in Industry 2004;54:1-15. [doi: 10.1016/j.compind.2003.09.002]
29. Adobe Creative Team. Adobe Acrobat 3D Version 8 Classroom in a Book. San Jose, CA: Adobe Press; 2007.
30. Beard L, Wilson K, Morra D, Keelan J. A survey of health-related activities on second life. J Med Internet Res 2009;11(2):e17 [FREE Full text] [doi: 10.2196/jmir.1192] [Medline: 19632971]
31. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
32. Clark HH. Arenas of Language Use. Chicago, IL: University of Chicago Press; 1992.
33. Clark HH. Using Language. Cambridge, UK: Cambridge University Press; 1996.
34. Clark HH, Brennan SA. Grounding in communication. In: Resnick LB, Levine JM, Teasley SD, editors. Perspectives on Socially Shared Cognition. Washington, DC: APA Books; 1991.
35. Bromme R, Jucks R, Runde A. Barriers and Biases in Computer-Mediated Knowledge Communication: and How They May Be Overcome. New York, NY: Springer; 2005:89-118.
36. Van Servellen GM. Communication Skills for the Health Care Professional: Concepts and Techniques. Gaithersburg, MD: Aspen Publishers; 1997.
37. Östlund M. Linköping University. 2008. AssistancePlus: 3D-Mediated Advice-Giving on Pharmaceutical Products [licentiate thesis] URL: http://liu.diva-portal.org/smash/record.jsf?pid=diva2:18453 [WebCite Cache ID 5ki7xWaiq]
38. Östlund M, Révay P. The virtual pharmacy: the use of applied virtual reality technology in presenting pharmaceutical information. In: Proceedings. 2003 Presented at: 4th International Conference of PhD Students; August 11-17, 2003; Miskolc, Hungary.

Abbreviations

3D: 3-dimensional

Edited by G Eysenbach; submitted 21.12.09; peer-reviewed by F Amenta, J Sankaranarayanan; comments to author 18.06.10; revised version received 25.05.11; accepted 26.05.11; published 18.07.11.

Please cite as:
Östlund M, Dahlbäck N, Petersson GI
3D Visualization as a Communicative Aid in Pharmaceutical Advice-Giving over Distance
J Med Internet Res 2011;13(3):e50
URL: http://www.jmir.org/2011/3/e50/
doi:10.2196/jmir.1437
PMID:21771714

Original Paper

Cool but Counterproductive: Interactive, Web-Based Risk Communications Can Backfire

Brian J Zikmund-Fisher[1,2,3,4], PhD; Mark Dickson[3], MA; Holly O Witteman[2,3], PhD

[1]Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, United States
[2]Division of General Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States
[3]Center for Bioethics and Social Sciences in Medicine, Medical School, University of Michigan, Ann Arbor, MI, United States
[4]Risk Science Center, School of Public Health, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Brian J Zikmund-Fisher, PhD
Department of Health Behavior and Health Education
School of Public Health
University of Michigan
1415 Washington Heights
Ann Arbor, MI, 48109-2029
United States
Phone: 1 734 936 9179
Fax: 1 734 763 7379
Email: bzikmund@umich.edu

Abstract

Background: Paper-based patient decision aids generally present risk information using numbers and/or static images. However, limited psychological research has suggested that when people interactively graph risk information, they process the statistics more actively, making the information more available for decision making. Such interactive tools could potentially be incorporated in a new generation of Web-based decision aids.

Objective: The objective of our study was to investigate whether interactive graphics detailing the risk of side effects of two treatments improve knowledge and decision making over standard risk graphics.

Methods: A total of 3371 members of a demographically diverse Internet panel viewed a hypothetical scenario about two hypothetical treatments for thyroid cancer. Each treatment had a chance of causing 1 of 2 side effects, but we randomly varied whether one treatment was better on both dimensions (strong dominance condition), slightly better on only one dimension (mild dominance condition), or better on one dimension but worse on the other (trade-off condition) than the other treatment. We also varied whether respondents passively viewed the risk information in static pictograph (icon array) images or actively manipulated the information by using interactive Flash-based animations of "fill-in-the-blank" pictographs. Our primary hypothesis was that active manipulation would increase respondents' ability to recognize dominance (when available) and choose the better treatment.

Results: The interactive risk graphic conditions had significantly worse survey completion rates (1110/1695, 65.5% vs 1316/1659, 79.3%, *P* < .001) than the static image conditions. In addition, respondents using interactive graphs were less likely to recognize and select the dominant treatment option (234/380, 61.6% vs 343/465, 73.8%, *P* < .001 in the strong dominance condition).

Conclusions: Interactivity, however visually appealing, can both add to respondent burden and distract people from understanding relevant statistical information. Decision-aid developers need to be aware that interactive risk presentations may create worse outcomes than presentations of static risk graphic formats.

(J Med Internet Res 2011;13(3):e60) doi:10.2196/jmir.1665

KEYWORDS

Patient-provider communication

Introduction

Ample evidence exists that even highly educated adults can have poor numeracy skills [1-4]. As a result, patient decision aids and other patient communications that incorporate risk statistics will only be effective in improving patient decision making if they use design features that make these risk communications easier to understand. In this vein, researchers have evaluated the benefits of giving people more intuitive representations of risks by, for example, using frequencies instead of percentages [5-8] and testing a wide variety of visual displays, such as bar graphs, pie charts, and pictographs or icon arrays [9-14].

Improving risk-communication methods is also important because patients are likely to translate risk statistics into intuitive “gist” representations [15,16] that may influence anxiety or worry, powerful emotions that have significant impacts on people’s responses to health risks and disease [17-22]. In fact, emotional responses may mediate cognitive risk perceptions or shape behavior independently, or both [23-25].

Recently, Natter and Berry suggested that communications that force the audience to actively process risk information may be more effective than more passive displays [26]. In their study, participants were better calibrated in their perceptions of medication side effects when they created a bar graph of the risk instead of just viewing one. Similarly, Ancker and colleagues found that a Web-based, game-like, interactive risk graphic in which participants clicked in a matrix until they uncovered a risk event had the effect of reducing disparities in risk perceptions between high- and low-numeracy participants [27]. While such interactive graphing tasks were formerly difficult to implement, advances in the interactive capabilities of the Internet provide an opportunity to integrate interactive graphing tasks into patient decision-support materials. Such exercises could be seen as one method of increasing patients’ active processing of risk information. Indeed, Ancker and colleagues also showed that their interactive risk graphics may elicit emotional responses that reflect the potential for increased understanding of actual risks and better ability to compare and contrast risks [28].

We therefore hypothesized that a task designed for active processing might help patients not only to comprehend the statistics presented to them but also to integrate these facts into their decision making and thereby make better decisions.

To test the hypothesis that interactive graphing tasks could improve knowledge of risk information and decision making relevant to the risks, we conducted an Internet-administered experiment in which participants read a hypothetical treatment decision-making scenario that included information about the risks of 2 possible side effects. While some study participants passively viewed the risk information in state-of-the-art static risk images, others had to actively complete an interactive graphing task that asked them to fill in an icon array to match risk statistics that were provided numerically. We assessed whether the interactive task affected survey completion rates, treatment choices, and gist knowledge about the treatment options.

Methods

Recruitment

A stratified random sample of US adults age 21 years and older was selected from a panel of Internet users administered by Survey Sampling International (Shelton, CT, USA). To ensure at least moderate demographic diversity (but not representativeness) and offset large expected variations in response rates, we drew distinct subsamples by both age and race, and dynamically adjusted the number of email invitations in each demographic subsample until all quotas were achieved. Selected panel members received email invitations to complete the online survey. Upon completing the survey, participants were entered into both an instant-win contest and a monthly drawing administered by Survey Sampling International for modest prizes.

Design of the Study

Respondents read a short vignette in which they imagined being diagnosed with thyroid cancer and discussing treatment options with their doctor. The vignette discussed 2 types of radiation treatment, external beam therapy and seed therapy, which were described as being equally effective in treating the patient’s type of thyroid cancer. Both therapies were described as having a chance of causing 2 side effects: (1) fatigue, and (2) mouth and throat problems. We chose thyroid cancer as the disease context because its comparative unfamiliarity (versus, for example, breast or prostate cancer) meant that few study participants would be likely to have preconceived beliefs about treatment options or their associated risks.

Our primary research question was to determine whether an interactive graphing task would increase respondents’ ability to recognize a better treatment option (ie, one with lower risks of side effects) when such was available. To do so, we experimentally varied the content of the 2 survey pages that presented the risk of each side effect with the two treatment options. On those pages, we varied 2 factors in a 3 (risk levels) × 2 (graphic type) between-subjects design.

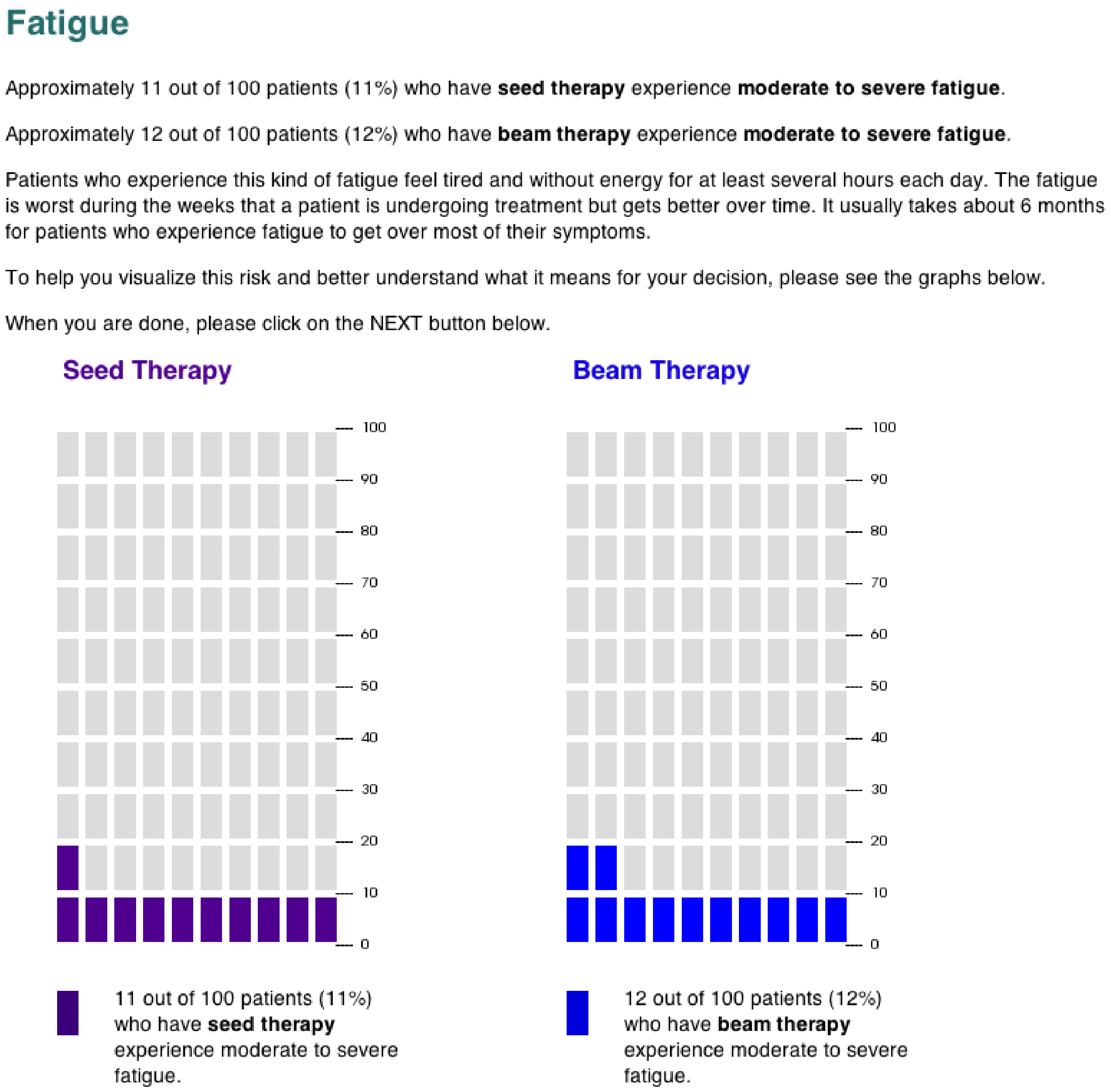
Table 1. Side-effect risks presented in the hypothetical vignette

	Risk of fatigue		Risk of mouth or throat problems	
	Beam therapy	Seed therapy	Beam therapy	Seed therapy
Trade-off condition	12%	11%	13%	15%
Mild dominance condition	12%	11%	15%	15%
Strong dominance condition	12%	11%	21%	15%

The risk-level manipulation is summarized in Table 1. By varying the likelihood of experiencing mouth or throat problems with beam therapy, we created 3 distinct decision contexts: (1) a trade-off condition, in which each treatment had a higher likelihood of one side effect and a lower likelihood of the other, (2) a mild dominance condition, in which seed therapy had a very slightly lower rate of fatigue and the same rate of mouth and throat problems, and (3) a strong dominance condition in which seed therapy had a lower rate of both side effects.

All risk information for the side effects was presented both numerically and in side-by-side graphics on sequential screens (first fatigue and then mouth and throat problems). Half of study participants viewed the risks displayed in static 100-unit icon arrays (also called pictographs: Figure 1), a format demonstrated to improve risk communication in a variety of medical contexts [9,13,29-33]. The remaining participants received an interactive version of the same icon array format (Figure 2), programmed in Flash (Adobe Systems Inc, San Jose, CA, USA). In the interactive exercise, the risk information was provided numerically but the graphics were initially blank (ie, all gray blocks). Participants were then instructed to use their mouse to click and/or drag in each graph to set it to the appropriate risk level. The graph continually adjusted to provide both visual and numerical feedback. (See Multimedia Appendix 1 for a movie demonstrating the interaction.) Participants were required to interact with and set the first of the 2 graphs on the page before they were allowed to go on in the study. The specific instructions given to participants are shown in Figure 1. Note, however, that participants were not prevented from advancing if they did not set the second graph or were inaccurate in their graphing of the risks; minimal use of the interactive graphics was sufficient to allow them to continue in the study. This design received Institutional Review Board exempt status approval as anonymous survey research.

Figure 1. Screen image of the static pictographs



Fatigue

Approximately 11 out of 100 patients (11%) who have **seed therapy** experience **moderate to severe fatigue**.

Approximately 12 out of 100 patients (12%) who have **beam therapy** experience **moderate to severe fatigue**.

Patients who experience this kind of fatigue feel tired and without energy for at least several hours each day. The fatigue is worst during the weeks that a patient is undergoing treatment but gets better over time. It usually takes about 6 months for patients who experience fatigue to get over most of their symptoms.

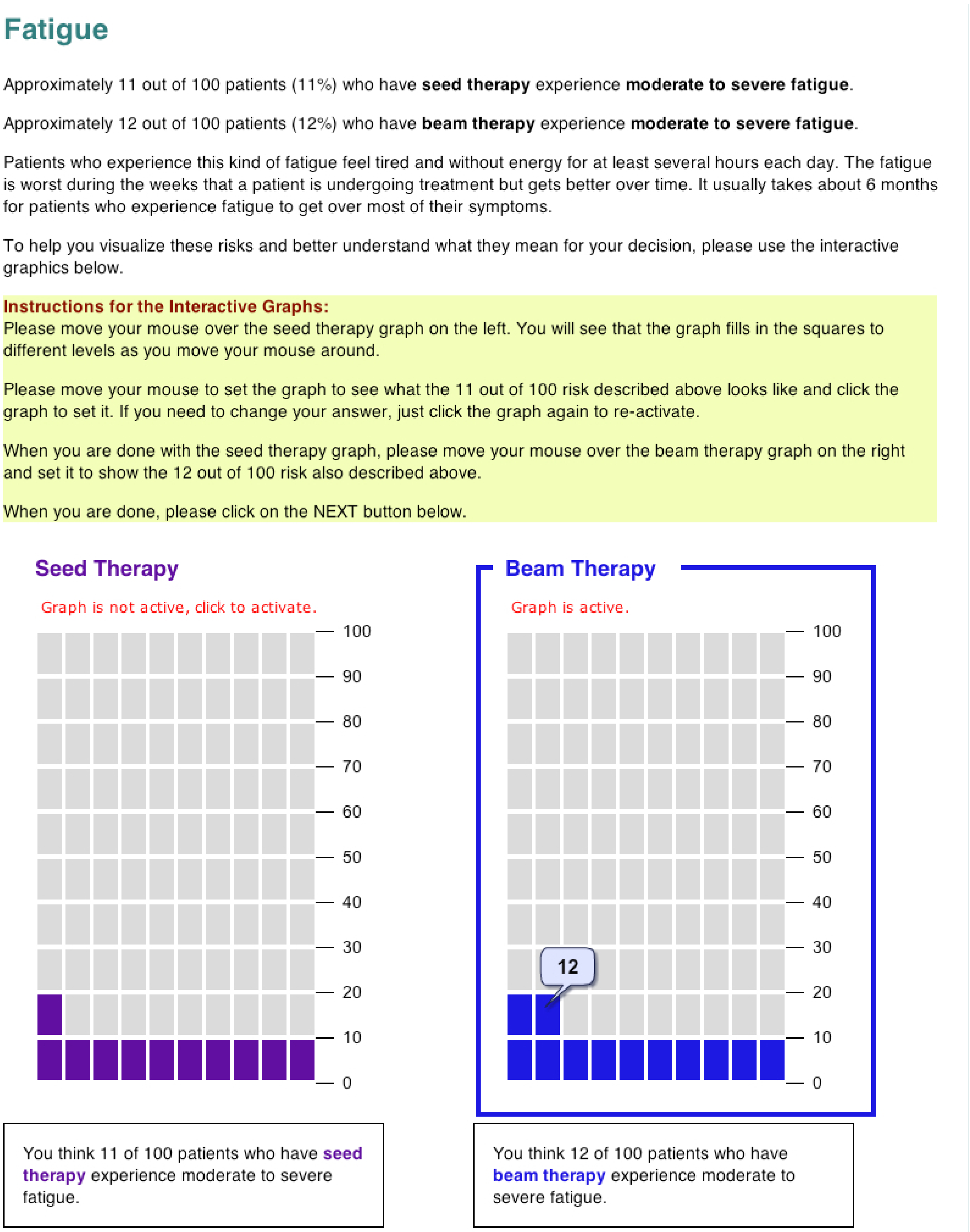
To help you visualize this risk and better understand what it means for your decision, please see the graphs below.

When you are done, please click on the NEXT button below.

■ 11 out of 100 patients (11%) who have **seed therapy** experience moderate to severe fatigue.

■ 12 out of 100 patients (12%) who have **beam therapy** experience moderate to severe fatigue.

Figure 2. Screen image of the interactive graphing task



Fatigue

Approximately 11 out of 100 patients (11%) who have **seed therapy** experience **moderate to severe fatigue**.

Approximately 12 out of 100 patients (12%) who have **beam therapy** experience **moderate to severe fatigue**.

Patients who experience this kind of fatigue feel tired and without energy for at least several hours each day. The fatigue is worst during the weeks that a patient is undergoing treatment but gets better over time. It usually takes about 6 months for patients who experience fatigue to get over most of their symptoms.

To help you visualize these risks and better understand what they mean for your decision, please use the interactive graphics below.

Instructions for the Interactive Graphs:
Please move your mouse over the seed therapy graph on the left. You will see that the graph fills in the squares to different levels as you move your mouse around.

Please move your mouse to set the graph to see what the 11 out of 100 risk described above looks like and click the graph to set it. If you need to change your answer, just click the graph again to re-activate.

When you are done with the seed therapy graph, please move your mouse over the beam therapy graph on the right and set it to show the 12 out of 100 risk also described above.

When you are done, please click on the NEXT button below.

Seed Therapy

Graph is not active, click to activate.

Beam Therapy

Graph is active.

You think 11 of 100 patients who have **seed therapy** experience moderate to severe fatigue.

You think 12 of 100 patients who have **beam therapy** experience moderate to severe fatigue.

Measures and Covariates

Our primary outcome measure was the preferred treatment choice (beam or seed). We also asked respondents 2 gist knowledge questions in which they were to indicate which therapy had a higher risk of each of the side effects. In addition, we electronically tracked 2 measures of respondent burden: (1) time spent reading the static graphs pages or interacting with the interactive graphs pages, and (2) survey break-offs (to assess whether the interactive task annoyed people sufficiently to make them stop taking the survey).

Because individuals vary in their numeracy (ie, their facility and comfort with quantitative health information such as risk statistics), all study participants also completed the Subjective Numeracy Scale (SNS) [34]. The SNS is a validated measure of quantitative ability and of preferences for receiving information in numerical form that has previously been shown to correlate with the ability to recall and comprehend both textual and graphical risk communications [32,35,36]. A participant's SNS score is calculated as his or her mean rating across the 8 SNS questions and ranges from 1 (least numerate) to 6 (most numerate). We also assessed participants' level of education, which we model for analysis purposes as a 3-level variable: high school or less, some post-high school education but no bachelor's degree, and bachelor's degree or more, as well as standard demographics measures.

Statistical Analyses

We used chi-square tests of proportions to test whether graph type affected treatment choices, knowledge recall, and survey discontinuation rates across the 3 risk-level conditions and logistic regression models that included both design factors, an interaction term, SNS score, and education level to assess the impact of these covariates on treatment choices. Because time spent on a given survey page has a highly skewed distribution, we then used Wilcoxon rank-sum tests to compare the distributions of time spent reading the static risk graphics versus completing the interactive graphing task. All analyses were performed using Stata (release 11; StataCorp LP, College Station, TX, USA), and all tests of significance were 2-sided and used alpha = .05.

Results

Recruitment

In total, 3371 people age 21 years and older reached the survey website and viewed the first content page. Of these, 17 reported having been actually diagnosed with thyroid cancer and were excluded as having pre-existing knowledge of the relevant treatment options, leaving 3354 possible participants.

Overall, 2426 (72%) of participants completed the entire survey, including questions on demographics that came toward the end of the survey instrument. Characteristics for those participants who answered each demographic question are reported in Table 2. We observed a wide range of educational achievement, with 889 participants (36.7%) having a bachelor's or higher college degree but also 441 (18.2%) having completed only high school or less education. The SNS numeracy measure showed high reliability (Cronbach alpha = .84). Mean SNS score was 4.63 (SD 1.0), with substantial variation (range 1.75–6.0). Because questions about participant demographics came at the end of the survey, we cannot know whether the demographics of those who dropped out differ from those who completed the survey.

Statistical Analyses

Dropout Rates and Time Spent on Risk Graphic Pages

As shown in Table 3, participants randomly assigned to complete the interactive graphics task were significantly less likely to complete the survey than those randomly assigned to view static graphs (65.5% vs 79.3%). More detailed examination confirmed that this difference was specifically due to the interactive graphics task. Over 23% of participants in the interactive graphics condition dropped out of the survey during that section of the survey, while less than 4% of participants viewing static graphics dropped out on those pages. Even among those who did progress beyond that point in the survey, it took significantly longer to complete the interactive graphing task than to view the static graphs.

Table 2. Participant[a] demographic characteristics

Characteristic	Distribution	Mean (SD)
Age (years)		49.1 (16.3)
21–29	351 (14.5%)	
30–39	448 (18.5%)	
40–49	435 (18.0%)	
50–59	396 (16.3%)	
60–69	528 (21.8%)	
≥70	266 (11.0%)	
Gender		
Male	1212 (50.0%)	
Female	1211 (50.0%)	
Ethnicity		
Hispanic (any race)	269 (11.2%)	
Race [b]		
White	2008 (82.7%)	
African American	236 (9.7%)	
All other	213 (8.8%)	
Education		
≤ High school	44 (1.8%)	
High school only	397 (16.4%)	
Some college/trade	1093 (45.1%)	
Bachelor’s degree	593 (24.5%)	
Master’s/doctorate	296 (12.2%)	
Subjective numeracy scale (mean rating)		4.63 (1.00)
1.00–1.99	45 (1.8%)	
2.00–2.99	174 (6.8%)	
3.00–3.99	500 (19.6%)	
4.00–4.99	894 (35.1%)	
5.00–5.99	862 (33.8%)	
6.00	72 (2.8%)	

[a] Reports results only for those respondents who completed each question or measure.

[b] Respondents could mark more than 1 race.

Table 3. Survey completion times rates, by graphic type

	Static graphs	Interactive graphs	Significance
Discontinued survey at the risk graphic pages	58/1659 (3%)	391/1695 (23.1%)	$\chi^2_1 = 277$; $P < .001$
Median time spent on the 2 pages with risk graphics (seconds)[a]	52	155	$z = 33.63$; $P < .001$
Completed entire survey	1316/1659 (79.3%)	1110/1695 (65.5%)	$\chi^2_1 = 80.2$; $P < .001$

[a] Among respondents who did not discontinue at the risk graphic pages.

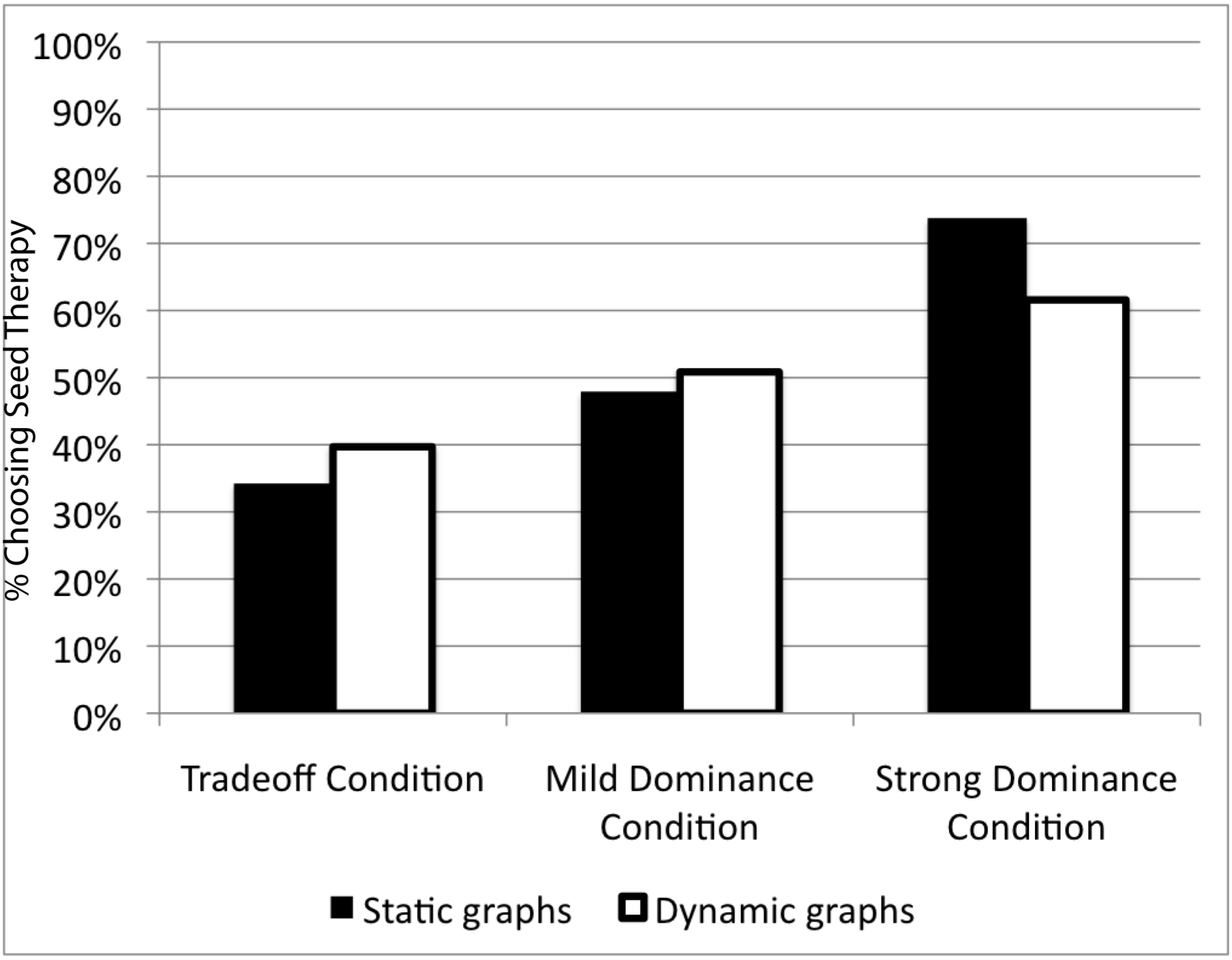
Treatment Choices

However problematic, the longer completion times and lower completion rates noted above for participants completing the interactive graphing task might be acceptable if it resulted in improved treatment choices among those who did complete the task. Unfortunately, such was not the case. As shown in Figure 3, selection of the best (dominant) treatment option of seed therapy in the strong dominance condition was significantly higher among those in the static graphics group than in the interactive graphics group (343/465, 73.8% vs 234/380, 61.6%, $\chi^2_1 = 14.3$, $P < .001$, among only those participants who answered the choice question; 343/536, 64.0% vs 234/558, 41.9%, $\chi^2_1 = 53.4$, $P < .001$, in an intent-to-treat analysis that included all participants, including those who dropped out of the survey). There was no significant difference in treatment choices in the mild dominance condition, and a mild trend toward less selection of seed therapy by participants viewing static graphs versus interactive graphs in the trade-off condition.

In fact, across all 3 risk-level conditions, the interactive graphics condition made those participants less sensitive to variations in the risk of mouth and throat problems than were participants in the static graphs condition. A logistic regression analysis that included both experimental factors, as well as participant numeracy and education, confirmed a significant graph type × condition interaction ($\chi^2_2 = 18.4$, $P < .001$). In addition, participants with higher SNS scores were significantly more likely to choose the best treatment (odds ratio = 1.18 per step on the 1–6 scale, $z = 3.78$, $P < .001$), but participant education had no effect. More complicated models found no significant interaction between participants' numeracy levels and the effect of graph type.

We also examined whether those participants in the interactive graphics condition who accurately graphed the risk statistics made different treatment choices from those who had difficulty completing this task. In the strong dominance condition (the only one that showed a significant difference overall between participants who viewed interactive and static graphics), study participants who accurately graphed all 4 side-effect risks (243, 64.0%, of the 380 who answered the treatment choice question) were significantly more likely to choose the dominant option of seed therapy (164/243, 67.5% vs 70/137, 51%, $\chi^2_1 = 10.0$, $P = .002$). We note, however, that optimal decision making remained at somewhat lower levels than observed in the static graphics condition (164/243, 67.5% vs 343/465, 73.8%, $\chi^2_1 = 3.1$, $P = .08$), thereby providing no evidence that even completely accurate use of an interactive graphing task would improve people's treatment selections.

Figure 3. Treatment choices, by risk level and graph type conditions. Note: graph reports choices among those participants who completed the survey. Selection of seed therapy is a dominant choice in the mild and strong dominance conditions but not in the trade-off condition



Gist Knowledge

As shown in Table 4, study participants who had to complete interactive graphing tasks were also less able to identify which treatment had the higher risk of mouth or throat problems in the trade-off condition (43.8% vs 51.6%). Logistic regression analysis found this effect only marginally significant (odds ratio = 0.77, z = –1.77, P = .08) when controlling for the highly significant effect of numeracy (odds ratio = 1.52 per step on the 1–6 scale, z = 5.17, P < .001). There were no significant differences in gist knowledge about mouth and throat problems in the other 2 conditions. Nor were there any significant differences in gist knowledge about which treatment had a higher rate of fatigue (a side effect that did not vary across the 3 risk-level conditions.)

Table 4. Percentage of participants correctly identifying which treatment had a higher rate of mouth and throat problems, by condition and graphic type

	Static graphs	Interactive graphs	Significance
Trade-off condition	225/436 (51.6%)	165/377 (43.8%)	χ^2_1 = 5.0; P = .03
Mild dominance condition	404/502 (80.5%)	338/414 (81.6%)	χ^2_1 = 0.2; P = .66
Strong dominance condition	288/452 (63.7%)	226/374 (60.4%)	χ^2_1 = 0.9; P = .33

Discussion

Principal Results

The promise of interactivity is to potentially encourage people not just to read relevant information but also to actively process it. Such active processing might help people to remember the information and use it to make better treatment decisions.

Unfortunately, the promise of interactive risk-communication exercises remained unfulfilled in our study. Our interactive task, which asked participants to graph pairs of risk statistics that were clearly presented in numerical format on the same page, increased survey dropouts and resulted in lower knowledge and poorer treatment choices than among those who viewed the information already displayed in equivalent static graphs.

Limitations

Our study has several key limitations. First and foremost, we tested a single type of interactive task with a particular user interface, one that did not undergo external usability testing. It is certainly possible that other tasks or more intuitive interfaces could have improved the use and user experience of the interactive task and potentially thereby improved outcomes. Second, as noted above, we asked otherwise healthy adults to answer a hypothetical scenario in an Internet survey. As a result, it is likely that these participants were considerably less motivated to work out how to do the interactive task and to think about the risk information we provided than real patients would be when reading a detailed decision aid designed to inform them and help them make their own medical decisions. Yet, we observed no advantage of the interactive task even for those participants who completed it perfectly. It is hard to imagine that even high levels of motivation would be enough to not just offset the knowledge and decision-making deficits we observed but actually accrue significant benefits from completing the interactive graphing exercise.

Comparison with Prior Work

As Natter and Berry note, although other research exists that uses interactive graphics in risk communication [37], such research has not generally evaluated the interactive graphics against their passive counterparts. Accordingly, their study was the first to demonstrate the potential of interactive graphics over standard graphics by showing that people who draw a bar graph rather than simply viewed one had better recall of medication side effects [26]. More recent research has suggested that interactive risk graphics in which participants explore a risk matrix until they find a risk event could also reduce numeracy disparities in risk perceptions [27].

By contrast, our study showed no such positive effect. In fact, participants who were asked to interactively graph the risk information in our study did significantly worse on multiple outcome measures. This reversal may be partly accounted for by the fact that we used pictographs (icon arrays), which are less familiar to people than bar graphs, and also partly by the fact that our study was conducted online, whereas Natter and Berry's used paper booklets. We also speculate post hoc that the novelty of the interactive task may have led participants to devote a majority of their attention and cognitive resources to figuring out what they were supposed to do. This effect may therefore have prevented our participants from engaging in the type of deeper, meaning-finding processing that we were hoping to stimulate. As noted above, a more in-depth, user-centered design process might have resulted in a better interface and hence reduced this problem.

Our different result may also derive from the fact that our task implemented a form of "teach-back" instructional methodology that encouraged participants to restate information they already had, whereas Ancker and colleagues' research used an exploratory task that encouraged discovery of the risk itself. In qualitative analyses of focus group transcripts, Ancker and colleagues concluded that their interactive risk graphics elicited more emotional responses than static graphics did, with more participants expressing concern about large risks and/or relief about small ones [28]. Their subsequent experimental study found no overall effect of graphic type on risk estimates or risk feelings [27]. While our study did not ask participants about how the graphics made them feel about their risk, we did find that people were less likely to choose the clearly optimal option after having used an interactive risk graphic than they were with a standard risk graphic. This suggests that the increased emotional response observed by Ancker et al may be an artifact of their game-like task and that more straightforward types of interactivity may not lead to improved choices.

Research on other forms of interactivity in health education such as video games [38] and immersive 3-dimensional environments [39] has suggested caution in considering how these applications might or might not improve outcomes in health education. Our results complement these prior findings by showing that interactive approaches to risk communication should come with similar caveats.

Conclusions

Interactive risk graphics are intriguing, but their use may be counterproductive to the purpose of effectively communicating to patients about health risks. While the movement of patient decision aids from booklet and DVD formats to online resources has created opportunities for greater technological sophistication, our research should serve as a cautionary tale to developers and educators seeking to ensure that their materials are “cutting edge.” Our teach-back interactive graphing task was intended to reinforce deeper cognitive processing of decision-relevant risk information, exactly the type of thinking most decision aids seek to promote in patients. Yet, in this case, our intervention backfired and created worse decision outcomes. More research is clearly needed to evaluate different types of interactive risk communications and to identify the design features of interactive exercises that lead to better results versus the features of those that do not. In the meantime, decision-aid designers should proceed with caution when considering the use of flashy risk graphics.

Acknowledgments

This research was supported by a grant from the Foundation for Informed Medical Decision Making (IIG-0126-1) to Dr. Zikmund-Fisher. Dr. Zikmund-Fisher is supported by a career development award from the American Cancer Society (MRSG-06-130-01-CPPB). The funding agreements ensured the authors’ independence in designing the study, interpreting the data, and publishing the report. The authors would like to thank Angela Fagerlin for helpful comments on an earlier draft of this manuscript, Bob Burbach and Mark Swanson for development of the Flash-based interactive graphics, and Nicole Exe for her exceptional research assistance.

Conflicts of Interest

None declared

Multimedia Appendix 1

MP4 movie of the interactive graphing task.

[MP4 File, 301KB - jmir_v13i3e60_app1.mp4]

References

1. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001;21(1):37-44. [Medline: 11206945]
2. Nelson W, Reyna VF, Fagerlin A, Lipkus I, Peters E. Clinical implications of numeracy: theory and practice. Ann Behav Med 2008 Jun;35(3):261-274. [doi: 10.1007/s12160-008-9037-8] [Medline: 18677452]
3. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Aff (Millwood) 2007;26(3):741-748. [doi: 10.1377/hlthaff.26.3.741] [Medline: 17485752]
4. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997 Dec 1;127(11):966-972 [FREE Full text] [Medline: 9412301]
5. Waters EA, Weinstein ND, Colditz GA, Emmons K. Formats for improving risk communication in medical tradeoff decisions. J Health Commun 2006 Mar;11(2):167-182. [doi: 10.1080/10810730500526695] [Medline: 16537286]
6. Gigerenzer G. The psychology of good judgment: frequency formats and simple algorithms. Med Decis Making 1996;16(3):273-280. [Medline: 8818126]
7. Schapira MM, Nattinger AB, McHorney CA. Frequency or probability? A qualitative study of risk communication formats used in health care. Med Decis Making 2001;21(6):459-467. [Medline: 11760103]
8. Miron-Shatz T, Hanoch Y, Graef D, Sagi M. Presentation format affects comprehension and risk assessment: the case of prenatal screening. J Health Commun 2009;14(5):439-450. [doi: 10.1080/10810730903032986] [Medline: 19657924]
9. Waters EA, Weinstein ND, Colditz GA, Emmons KM. Reducing aversion to side effects in preventive medical treatment decisions. J Exp Psychol Appl 2007 Mar;13(1):11-21. [doi: 10.1037/1076-898X.13.1.11] [Medline: 17385998]
10. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. Patient Educ Couns 2008 Dec;73(3):448-455. [doi: 10.1016/j.pec.2008.07.023] [Medline: 18755566]
11. Lipkus IM, Hollands JG. The visual communication of risk. J Natl Cancer Inst Monogr 1999(25):149-163. [Medline: 10854471]

12. Stone ER, Yates JF, Parker AM. Effects of numerical and graphical displays on professed risk-taking behavior. J Exp Psychol Appl 1997;3(4):243-256. [doi: 10.1037/1076-898X.3.4.243]
13. Feldman-Stewart D, Brundage MD, Zotov V. Further insight into the perception of quantitative information: judgments of gist in treatment decisions. Med Decis Making 2007;27(1):34-43. [doi: 10.1177/0272989X06297101] [Medline: 17237451]
14. Garcia-Retamero R, Galesic M. Who profits from visual aids: overcoming challenges in people's understanding of risks [corrected]. Soc Sci Med 2010 Apr;70(7):1019-1025. [doi: 10.1016/j.socscimed.2009.11.031] [Medline: 20116159]
15. Reyna VF, Ellis SC. Fuzzy-trace theory and framing effects in children's risky decision making. Psychol Sci 1994;5(5):275-279. [doi: 10.1111/j.1467-9280.1994.tb00625.x]
16. Reyna VF, Hamilton AJ. The importance of memory in informed consent for surgical risk. Med Decis Making 2001;21(2):152-155. [Medline: 11310949]
17. McCaul KD, Tulloch HE. Cancer screening decisions. J Natl Cancer Inst Monogr 1999(25):52-58. [Medline: 10854458]
18. Rothman AJ, Kiviniemi MT. Treating people with information: an analysis and review of approaches to communicating health risk information. J Natl Cancer Inst Monogr 1999(25):44-51. [Medline: 10854457]
19. Cameron LD, Leventhal H, Love RR. Trait anxiety, symptom perceptions, and illness-related responses among women with breast cancer in remission during a tamoxifen clinical trial. Health Psychol 1998 Sep;17(5):459-469. [Medline: 9776005]
20. Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. Ann Intern Med 1991 Apr 15;114(8):657-661. [Medline: 2003712]
21. Lofters A, Juffs HG, Pond GR, Tannock IF. "PSA-itis": knowledge of serum prostate specific antigen and other causes of anxiety in men with metaststic prostate cancer. J Urol 2002 Dec;168(6):2516-2520. [doi: 10.1097/01.ju.0000032824.52830.55] [Medline: 12441952]
22. Trask PC, Paterson AG, Wang C, Hayasaka S, Milliron KJ, Blumberg LR, et al. Cancer-specific worry interference in women attending a breast and ovarian cancer risk evaluation program: impact on emotional distress and health functioning. Psychooncology 2001;10(5):349-360. [Medline: 11536413]
23. Slovic P, Peters E, Finucane ML, Macgregor DG. Affect, risk, and decision making. Health Psychol 2005 Jul;24(4 Suppl):S35-S40. [doi: 10.1037/0278-6133.24.4.S35] [Medline: 16045417]
24. Loewenstein GF, Weber EU, Hsee CK, Welch N. Risk as feelings. Psychol Bull 2001 Mar;127(2):267-286. [Medline: 11316014]
25. Finucane ML, Alhakami A, Slovic P, Johnson SM. The affect of heuristic judgments of risks and benefits. J Behav Decis Making 2000;13:1-17. [doi: 10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S]
26. Natter HM, Berry DC. Effects of active information processing on the understanding of risk information. Appl Cogn Psychol 2005;19:123-135. [doi: 10.1002/acp.1068]
27. Ancker JS, Weber EU, Kukafka R. Effects of game-like interactive graphics on risk perceptions and decisions. Med Decis Making 2011;31(1):130-142. [doi: 10.1177/0272989X10364847] [Medline: 20393103]
28. Ancker JS, Chan C, Kukafka R. Interactive graphics for expressing health risks: development and qualitative evaluation. J Health Commun 2009;14(5):461-475. [doi: 10.1080/10810730903032960] [Medline: 19657926]
29. Fagerlin A, Wang C, Ubel PA. Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? Med Decis Making 2005;25(4):398-405. [doi: 10.1177/0272989X05278931] [Medline: 16061891]
30. Zikmund-Fisher BJ, Ubel PA, Smith DM, Derry HA, McClure JB, Stark A, et al. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. Patient Educ Couns 2008 Nov;73(2):209-214. [doi: 10.1016/j.pec.2008.05.010] [Medline: 18602242]
31. Price M, Cameron R, Butow P. Communicating risk information: the influence of graphical display format on quantitative information perception-Accuracy, comprehension and preferences. Patient Educ Couns 2007 Dec;69(1-3):121-128. [doi: 10.1016/j.pec.2007.08.006] [Medline: 17905553]
32. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Improving understanding of adjuvant therapy options by using simpler risk graphics. Cancer 2008 Dec 15;113(12):3382-3390 [FREE Full text] [doi: 10.1002/cncr.23959] [Medline: 19012353]
33. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol 2009 Mar;28(2):210-216. [doi: 10.1037/a0014474] [Medline: 19290713]
34. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
35. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making 2007;27(5):663-671. [doi: 10.1177/0272989X07303824] [Medline: 17652180]
36. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. A demonstration of "less can be more" in risk graphics. Med Decis Making 2010;30(6):661-671. [doi: 10.1177/0272989X10364244] [Medline: 20375419]
37. Woloshin S, Schwartz LM, Byram S, Fischhoff B, Welch HG. A new scale for assessing perceptions of chance: a validation study. Med Decis Making 2000;20(3):298-307. [Medline: 10929852]

38. Annetta LA, Minogue J, Holmes SY. Cheng M-T. Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education 2009;53(1):74-85. [doi: 10.1016/j.compedu.2008.12.020]
39. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]

Abbreviations

SNS: Subjective Numeracy Scale

Edited by G Eysenbach; submitted 05.10.10; peer-reviewed by J Ancker, T Miron-Shatz; comments to author 26.01.11; revised version received 24.02.11; accepted 05.05.11; published 25.08.11.

Please cite as:
Zikmund-Fisher BJ, Dickson M, Witteman HO
Cool but Counterproductive: Interactive, Web-Based Risk Communications Can Backfire
J Med Internet Res 2011;13(3):e60
URL: http://www.jmir.org/2011/3/e60/
doi:10.2196/jmir.1665
PMID:21868349

Original Paper

Managing the Personal Side of Health: How Patient Expertise Differs from the Expertise of Clinicians

Andrea Hartzler[1], PhD; Wanda Pratt[1,2], PhD

[1]Division of Biomedical and Health Informatics, School of Medicine, University of Washington, Seattle, WA, United States

[2]The Information School, University of Washington, Seattle, WA, United States

Corresponding Author:
Andrea Hartzler, PhD
Division of Biomedical and Health Informatics
School of Medicine
University of Washington
Box 357240
1959 NE Pacific Street, HSB I-264
Seattle, WA, 98195
United States
Phone: 1 206 616 0369
Fax: 1 206 221 2671
Email: andreah@uw.edu

Abstract

Background: When patients need health information to manage their personal health, they turn to both health professionals and other patients. Yet, we know little about how the information exchanged among patients (ie, patient expertise) contrasts with the information offered by health professionals (ie, clinician expertise). Understanding how patients' experiential expertise contrasts with the medical expertise of health professionals is necessary to inform the design of peer-support tools that meet patients' needs, particularly with the growing prevalence of largely unguided advice sharing through Internet-based social software.

Objective: The objective of our study was to enhance our understanding of patient expertise and to inform the design of peer-support tools. We compared the characteristics of patient expertise with that of clinician expertise for breast cancer.

Methods: Through a comparative content analysis of topics discussed and recommendations offered in Internet message boards and books, we contrasted the topic, form, and style of expertise shared in sources of patient expertise with sources of clinician expertise.

Results: Patient expertise focused on strategies for coping with day-to-day personal health issues gained through trial and error of the lived experience; thus, it was predominately personal in topic. It offered a wealth of actionable advice that was frequently expressed through the narrative style of personal stories about managing responsibilities and activities associated with family, friends, work, and the home during illness. In contrast, clinician expertise was carried through a prescriptive style and focused on explicit facts and opinions that tied closely to the health care delivery system, biomedical research, and health professionals' work. These differences were significant between sources of patient expertise and sources of clinician expertise in topic ($P < .001$), form ($P < .001$), and style ($P < .001$).

Conclusion: Patients offer other patients substantial expertise that differs significantly from the expertise offered by health professionals. Our findings suggest that experienced patients do not necessarily serve as "amateur doctors" who offer more accessible but less comprehensive or detailed medical information. Rather, they offer valuable personal information that clinicians cannot necessarily provide. The characteristics of patient expertise and the resulting design implications that we identified will help informaticians enhance the design of peer-support tools that will help meet patients' diverse information needs.

(J Med Internet Res 2011;13(3):e62) doi:10.2196/jmir.1728

KEYWORDS

Health knowledge; attitudes; practice; social support; community networks; peer group; consumer health informatics; online communities; patient expertise; personalized health

Introduction

In addition to the indispensable information received from health professionals, patients use information and advice offered by other patients to help them actively participate in their own health care and make informed personal health decisions. Although patients are best known for providing emotional support, they also offer other patients personal health guidance based on the expertise they have gained from managing similar health situations. We define *patient expertise* as experiential knowledge gained from personally managing the day-to-day experience of illness. For example, patients develop expertise in the self-management of chronic conditions through their everyday experience with self-care over time [1-3]. Some experienced patients can even come to know as much as their doctors about aspects of their health [4]. In the context of breast cancer, patient expertise reflects practical know-how and coping strategies exchanged among patients and their peers, including other patients, cancer survivors, and their caregivers, family, and friends (ie, other patients). In contrast, we define *clinician expertise* as knowledge gained from professional training and practice. Clinician expertise is shared with patients by health professionals (eg, physicians, nurses, therapists, and support staff).

In contrast to other forms of social support, including *emotional* support (ie, communication of caring and concern) and *instrumental support* (ie, provision of material goods), patients commonly share their expertise through *informational support*, which involves the provision of information and advice used to guide one's personal health management in new ways [5]. A patient's need for informational support is thought to be strongest during periods of transition, once emotional support needs associated with a crisis have been met [6]. The need for guidance, which is carried through informational support, is commonly met by people with expertise [7]. In the context of patient-expertise sharing, experienced patients serve as experts by providing informational support for personal health management through experiential guidance. Peer-support programs for cancer patients that provide opportunities to exchange such guidance are associated with benefits for patients, including improved coping skills, understanding of the cancer experience, and psychosocial outcomes [8,9].

Patient expertise has been valued in varied and growing contexts. For example, personal knowledge, such as lifestyle, priorities, and experiences, is an important contribution patients make to shared decision making with health professionals [10]. Providing patients with decision-support tools to share their personal knowledge and experiences with health professionals can improve nursing care and patient outcomes [11]. Other research demonstrates the value of involving patients as teachers who share their illness experiences through medical education [12].

In this work, we focus on patients sharing their expertise with one another. Breast cancer patients, for example, have expressed a strong need for experiential health information provided by peers [13]. When those needs for patient expertise are met, patients might be better able to receive and appropriately use health information than when patients interact with an "ask-the-expert" service that offers clinically oriented resources [14]. However, we lack a deep understanding of the characteristics of patient expertise, which hinders clear guidance on how to design peer-support tools that facilitate patient-expertise sharing.

Historically, patients who share similar health situations have helped one another to cope with illness by sharing their expertise through participation in patient-led support groups [15], as mentors in pioneering programs such as "Reach to Recovery" [16], and as instructors for chronic disease self-management programs [17]. Although the Internet has facilitated expertise sharing among patients in online health communities, many content analyses of interactions among correspondents of online health communities for conditions such as breast cancer [18-22] focus on emotional support [23]. Yet, growing evidence highlights the high prevalence of patient expertise exchanged through informational support in online health communities [24-26].

Patient expertise has continued to gain visibility as Web-based social software (eg, forums, social networking tools, blogs, and wikis) helps patients to readily exchange information and advice with others who are facing similar health situations [27-31]. For example, patient-led support groups provide a longstanding online resource for patient-generated guidance and advice on treatments, personal histories, shared experiences, and medication side effects for epilepsy [24]. More recently, personal profiles and commenting features of PatientsLikeMe.com allow users to ask one another about specific health experiences and to offer advice, suggestions, and tips that stem from those experiences [32]. In a recent survey, 1 in 5 Internet users reported going online to find others with similar health concerns, particularly for chronic or rare conditions [31]. Indeed, many patients now use social software more often to obtain expertise from other patients than to obtain their emotional support [28]. However, social software still needs enhancements to facilitate this important peer interaction around health [33], particularly in making the expertise of patients more prominent, explicit, and accessible [34]. For example, tools could do more to help users relate to the health experiences of other users [35] or to help users gain awareness of the expertise available from other users without the time required for multiple interactions when building relationships [9]. Research to improve our understanding of the characteristics of patient expertise could deeply inform design enhancements that facilitate expertise sharing.

Despite the growing prevalence of expertise sharing among patients on the Internet, disparate views about the characteristics of that expertise remain. For example, Meier et al [25], through a content analysis of cancer-related Internet mailing lists, found that patient exchanges of information and advice clustered squarely around medically oriented topics (eg, treatments and communication with health care providers). Our preliminary work on breast cancer message boards shows additional clusters around personal topics related to the cancer experience, including the family, home, work, advocacy, and interactions with social networks [26]. Similarly, other work notes that nearly half of interactions between cancer patients and survivors through a telephone-based peer-support system revolved around psychosocial and day-to-day issues, such as the impact of cancer on family and friends, compared with interactions revolving around more medically focused treatments and side effects [36]. A mixture of treatment-related and personal topics has also been noted in discussions among other patient groups, such as questions posed to peers on online forums for epilepsy [24], amyotrophic lateral sclerosis [37], and anterior cruciate ligament injuries [23]. These contrasting views of patient expertise (ie, whether it is more medial or more personal in nature) warrant detailed investigation that explores how the expertise exchanged among patients contrasts with the medical expertise offered by health professionals.

Facilitating patient-expertise sharing with innovative technology will depend on a solid understanding of the fundamental characteristics of the expertise that patients share. For example, could we meet patients’ needs for information solely by enhancing communication between patients and health professionals? Alternatively, do patients also need help finding other patients who have had similar health experiences because clinicians have neither the time nor the expertise to meet all their needs? An important first step to answering these questions is to understand the role of both patient expertise and clinician expertise in meeting patients’ needs.

Our aim in this study was to enhance our understanding of patient expertise by assessing how it differs from clinician expertise. In the context of breast cancer, we conducted an in-depth and comparative content analysis [38] to investigate how patient expertise and clinician expertise compare with respect to topics discussed and the types of recommendations offered. Through a multiphased approach, we enhanced our previous work [26] by (1) extending our characterization of patient expertise through the analysis of content from 2 patient-authored books, (2) characterizing clinician expertise through the analysis of content from a leading breast cancer resource book written by a health professional and an “ask-the-doctor” message board, and (3) contrasting the characteristics of patient expertise and clinician expertise. We conclude with suggestions for how our results can be used to enhance the design of collaborative peer-support tools for patients that will help them cope with their health issues and make informed health care decisions.

Methods

Using an evolving coding scheme that was grounded in the data [39], we conducted our content analysis of patient and clinician expertise in 4 phases. In phase 1, we analyzed content from sources of patient expertise to identify emergent *topics* discussed (ie, medical or personal) and the types of *recommendations* offered (ie, advice in the form of suggested action strategies, knowledge, perspectives, or information resources). We also noted the style through which recommendations were carried (ie, narrative or prescriptive). Next, in phase 2, we expanded our coding scheme by repeating this procedure using content from sources of clinician expertise. We then solidified our coding scheme into a codebook and tested the reliability of our coding procedure in phase 3. Finally, in phase 4, we contrasted the topics discussed and recommendations offered in sources of patient expertise and sources of clinician expertise. We describe each phase in detail in the Procedure section. This multiphased analysis answered 2 main research questions: (1) How do topics discussed in sources of patient expertise compare with topics discussed in sources of clinician expertise? (2) How do the recommendations offered by sources of patient expertise compare in form and style with recommendations offered by sources of clinician expertise?

Content Sources

We analyzed sources of patient expertise and sources of clinician expertise from both online message boards and books. Message boards are common online resources for patients to seek advice from peers through online communities or to seek advice from health professionals through ask-the-expert forums. Books are traditional resources that patients also commonly turn to for advice, both authored by health professionals (ie, clinician expertise) and authored by peer cancer survivors (ie, patient expertise). Books are particularly important because they are one of the few written forms available to patients who have no Internet access. Although books offer a limited source of perspectives because of the short list of authors, they attempt to provide in-depth expertise from that perspective. In contrast, message boards bring insights into the kinds of expertise actively sought by patients from a breadth of perspectives. Although patients have available to them a spectrum of valuable resources, our combined analysis of message boards and books enabled us to capture a variety of expertise that is available to and sought by patients at both ends of that spectrum, both online and offline.

Sources of patient expertise in our analysis included 3 online message boards that support correspondence among breast cancer patients, and 2 books written by cancer survivors. We selected the 3 patient message boards (message boards A, B, and C) based on public accessibility, high volume of use, longevity, and variation in formality (ie, varied levels of moderation and affiliation with health-related organizations). We selected the 2 patient books because they are autobiographical yet differ in style. The first patient book (book 1) is highly narrative in its compilation of experiences contributed by 10 breast cancer survivors. The second patient book (book 2) is an interactive guide written by a 6-year survivor of metastatic cancer who provides extensive strategies for staying organized and informed during the cancer experience.

Sources of clinician expertise included an ask-the-doctor message board that supports correspondence between breast cancer patients and health professionals. We selected this message board over clinical advice summaries or health professionals’ blogs to enable analysis of questions from patients and answers from health professionals. As an additional source of clinician expertise, we chose *Dr. Susan Love’s Breast Book* [40], which is a popular book written for breast cancer patients. We selected this book because correspondents in the message boards we analyzed often recommended this popular resource to one another. For example, this book was referred to as the “bible” of breast cancer and is a source of clinician expertise that many patients turn to.

Table 1 shows the content sources, including the number of text pages we analyzed and the number of content units (see Procedure) that each source contributed to the analysis. Table 2 shows characteristics of the 4 message boards.

Table 1. Content sources

	Source	Text pages	Content units
Patient expertise	Message board A	174	50
	Message board B	316	50
	Message board C	276	50
	Book 1: McCarthy and Loren, 1997 [41]	230	79
	Book 2: Willis, 2001 [42]	220	131
	Total	1216	360
Clinician expertise	Ask-the-doctor message board	277	150
	Book: Love and Lindsey, 2000 [40]	552	225
	Total	829	375

Table 2. Characteristics of message boards

	Message board			
	A	B	C	Ask the doctor
Year of inception	1998	1994	1998	2000
Moderation	Yes	No	No	Yes
Affiliation with a health-related organization	Yes	Yes	No	Yes
Total messages	379	152	316	300
Mean messages/thread (range)	8 (1–31)	3 (1–10)	8 (1–27)	2 (1–2)
Days’ worth of threads	5	24	55	85

Procedure

In phase 1, we analyzed content from the sources of patient expertise. Our unit of analysis for message boards was the *thread* (ie, 1 or more related messages) and for books was the *subsection* (ie, a titled section within a chapter). Our inclusion criteria for the analysis were *content units* (ie, message board thread or book subsection) that solicited or offered informational support (ie, information used to guide or advise) related to the diagnosis, treatment, or long-term management of breast cancer.

We collected archived threads from the patient message boards with posting dates starting in February 2006 until we obtained 50 content units from each board that met our inclusion criteria. Obtaining an equal number of content units from each patient message board required the collection and filtering of more threads from message board B (130 threads) than from message board A (66 threads) or message board C (81 threads). Common kinds of threads that we excluded from the analysis reflected exchanges of pure emotional support, technical support issues, threads labeled by correspondents as *off topic*, and spam-like advertisements. For our corresponding examination of patient expertise in books, we divided both patient books into subsections. All subsections from both patient books met our inclusion criteria. Sources of patient expertise contributed 360 content units in total. The patient message boards contributed 150 content units and patient books contributed 210 content units (see Table 1).

Based on themes that emerged from our preliminary analysis of informational support exchanged in the patient message boards [26], we coded content units from the sources of patient expertise while expanding our coding scheme. Our coding procedure was shaped by the challenge of identifying explicitly formulated questions within long, evolving discussion threads on patient message boards. In our preliminary work, we noticed that questions were often formulated as threads evolved. Other times, discussion was initiated by comments, rather than questions. More generally, we observed that threads typically discussed an overarching problem, or “topic,” whether initiated by an explicit question, an evolving question, or a comment. In response, other users would post potential solutions, or “recommendations,” for that problem. Rather than limiting our analysis to only those threads initiated by explicitly formulated questions, we framed our coding procedure more generally in terms of topics and recommendations that content units reflect. Thus, for each content unit, we used our evolving coding scheme to capture emergent topics discussed and recommendations offered:

- **Topics** reflect a spectrum of personal health issues discussed, such as specific situations or problems (eg, choosing a doctor). Topics represent the predominate theme of a content unit, which was typically most clearly identified within the introductory paragraph of a book subsection or within the initiating message and subject line of a message board thread. We assigned 1 topic to each content unit.

- **Recommendations** reflect a range of advice offered for dealing with the personal health issue (ie, topic) discussed within a content unit, such as a suggested solution to a problem. Unlike the breadth captured by topics, recommendations are fine-grained statements of advice serving as potential problem solutions in the form of action strategies, knowledge, perspectives, and information resources that were often sprinkled throughout a content unit. Recommendations reflect either a prescriptive style (ie, “you should...”) or a narrative style when carried through a personal story (ie, “when I was in that situation, this is what I did...”). We assigned 1 or more recommendation to each content unit.

In **phase 2**, we expanded the coding scheme by repeating our coding procedure on content from sources of clinician expertise. We collected threads from the ask-the-doctor message board until we obtained 150 content units that met our inclusion criteria. Unlike the threads from the patient message boards, the threads from the ask-the-doctor message board were generally short, consisting of a question posted by a patient and a response posted by a health professional, and each met our inclusion criteria. We divided the clinician book into subsections and excluded subsections that did not meet our inclusion criteria. We excluded subsections from chapters 1 through 9 of the clinician book because content from those chapters describes the development of healthy breasts and common breast problems rather than relating directly to the diagnosis, treatment, or long-term management of breast cancer. Sources of clinician expertise contributed 375 content units in total. The ask-the-doctor message board contributed 150 content units and the clinician book contributed 225 content units (see Table 1).

In **phase 3**, we used card sorting [43] and discussion to solidify our evolving coding scheme into a codebook made up of 2 main overlapping topics (medical and personal) and 4 types of recommendations (action strategies, knowledge, perspectives, and information resources). Our codebook, which reflects the end point of our evolving coding scheme, includes clusters of subtopics discussed, as well as different styles (ie, prescriptive or narrative) through which recommendations were expressed across all content units. Table 3 shows our codebook and details the distribution of topics (part a) and recommendations (part b) in patient and clinician sources. We describe the codes making up our codebook in greater detail in the Results section.

Table 3. Codebook of (a) topics and (b) recommendations (percentages reflect proportions of content units from each type of content source)

	Patient message boards	Patient books	Ask the doctor	Clinician book
a. Topics				
Medical				
Deciding on care teams, treatments, and procedures	16 (11%)	9 (4%)	19 (13%)	14 (6%)
Understanding biomedical concepts and processes	49 (33%)	6 (3%)	102 (68%)	145 (65%)
Managing interactions with health professionals	2 (1%)	15 (7%)	17 (11%)	1 (0.5%)
Managing information to collaborate with clinicians or understand biomedical issues	2 (1%)	3 (1%)	6 (4%)	5 (2%)
Personal				
Managing life at home	8 (5%)	16 (8%)	0 (0%)	0 (0%)
Managing work life	3 (2%)	14 (7%)	1 (0.7%)	1 (0.5%)
Managing one's emotional response to illness	12 (8%)	11 (5%)	0 (0%)	5 (2%)
Managing interactions with social networks	8 (5%)	18 (9%)	1 (0.7%)	3 (1%)
Managing personal tasks and projects	16 (11%)	86 (41%)	1 (0.7%)	19 (9%)
Managing advocacy and volunteer work	6 (4%)	2 (1%)	0 (0%)	0 (0%)
Both medical and personal	28 (19%)	30 (14%)	3 (2%)	32 (14%)
Total content units	150	210	150	225
b. Recommendations				
Action strategies				
Prescriptive	248 (23%)	303 (14%)	122 (35%)	474 (13%)
Narrative	192 (18%)	223 (10%)	0 (0%)	27 (1%)
Knowledge				
Prescriptive	159 (15%)	419 (19%)	225 (65%)	1,620 (45%)
Narrative	204 (19%)	264 (12%)	0 (0%)	133 (4%)
Perspectives				
Prescriptive	96 (9%)	70 (3%)	0 (0%)	76 (2%)
Narrative	48 (4%)	97 (4%)	0 (0%)	3 (<1%)
Information resources				
Books	13 (1%)	11 (1%)	0 (0%)	195 (6%)
Contact information	17 (2%)	23 (1%)	0 (0%)	314 (9%)
Magazines and magazine articles	2 (<1%)	7 (<1%)	0 (0%)	15 (<1%)
Multimedia	0 (0%)	4 (<1%)	0 (0%)	140 (4%)
News articles	19 (2%)	11 (1%)	0 (0%)	2 (<1%)
Poems and quotes	5 (<1%)	24 (1%)	0 (0%)	0 (0%)
Research articles and academic journals	11 (1%)	64 (3%)	0 (0%)	350 (10%)
Templates	0 (0%)	115 (5%)	0 (0%)	4 (<1%)
Webpages	70 (6%)	482 (22%)	1 (<1%)	118 (3%)
Miscellaneous publications	0 (0%)	86 (4%)	1 (<1%)	102 (3%)
Total recommendations	1084	2203	349	3573

We used the codebook to test the reliability of our coding procedure using a 10% reliability sample of content units. Based on the number of contributing units, we randomly selected a set of content units from each content source for the reliability sample. An independent coder (CL) applied the codebook to the reliability sample. We calculated kappa scores to determine the level of intercoder agreement between codes applied to the reliability sample by AH (1 of the authors) and by CL. We applied linear weighting to our kappa calculations [44] for recommendations to account for the level of agreement between coders for both types and numbers of recommendations (ie, coders could assign multiple types and numbers of recommendations to each content unit). Reliability test results show good intercoder agreement for topics (κ = .71), action strategies (κ = .69), knowledge (κ = .72), and perspectives (κ = .54), as well as excellent intercoder agreement for information resources (κ = .94).

In phase 4, we compared the kinds of topics discussed and the types of recommendations offered in sources of patient expertise versus sources of clinician expertise. We compared the distribution of topics and recommendations across patient sources and across clinician sources. Then, we explored differences in the proportions of subtopics as well as the types and styles of recommendations among content sources. Finally, we used Pearson’s chi square statistic to assess differences in the frequencies of topics and recommendations between sources of patient expertise and sources of clinician expertise.

Ethical Considerations

We thought deeply about ethical considerations and evolving guidelines for conducting Internet-based research [45-49] as we obtained, analyzed, and reported our findings from online message board content. Thus, we selected public message boards for which membership was not required to access content, collected archived threads, removed identifiers from collected threads, and took care in reporting our results to balance the anonymity of correspondents with research trustworthiness. Our approach aligns closely with the approach taken in other content analyses of online health message boards, such as Finn [50]. We obtained institutional review board approval through the University of Washington before collecting data from the message boards.

Results

We analyzed 735 content units across all sources. Patient sources contributed 360 content units and clinician sources contributed 375 content units. Each content unit was associated with 1 topic, but usually with many recommendations. Content units contained 7209 recommendations in total. Content units from patient sources contained 3287 recommendations and content units from clinician sources contained 3922 recommendations. Table 4 shows the distribution of topics (part a) and recommendations (part b) across individual content sources.

Table 4. Distribution of (a) topics and (b) recommendations across content sources (percentages reflect proportions from individual sources)

	Patient message boards			Patient books		Ask the doctor	Clinician book
	A	B	C	1	2		
a. Topics							
Medical	25 (50%)	12 (24%)	32 (64%)	12 (15%)	21 (16%)	144 (96%)	165 (74%)
Personal	18 (36%)	22 (44%)	13 (26%)	58 (73%)	89 (68%)	3 (2%)	28 (12%)
Both	7 (14%)	16 (32%)	5 (10%)	9 (12%)	21 (16%)	3 (2%)	32 (14%)
Total content units	50	50	50	79	131	150	225
b. Recommendations							
Action strategies	215 (39%)	119 (47%)	106 (38%)	300 (36%)	226 (17%)	122 (35%)	501 (14%)
Knowledge	200 (36%)	52 (21%)	111 (40%)	368 (44%)	315 (23%)	225 (64%)	1753 (49%)
Perspectives	86 (15%)	33 (13%)	25 (9%)	121 (14%)	46 (3%)	0 (0%)	79 (2%)
Information resources	54 (10%)	47 (19%)	36 (13%)	49 (6%)	778 (57%)	2 (1%)	1240 (35%)
Total recommendations	555	251	278	838	1365	349	3573

Next, we detail the kinds of topics and recommendations that emerged from our analysis across content units from all sources. The descriptive detail we provide on topics and recommendations corresponds to the codes making up our codebook (see Table 3). We then describe how sources of patient expertise and sources of clinician expertise compare with respect to those topics and recommendations.

Topics Discussed Across Content Sources

Most content units fell into 2 broad topic categories: discussion that was mostly *medical* in nature (411/735 content units, 56%) or discussion that was mostly *personal* in nature (231/735 content units, 31%). A smaller proportion of content units contained discussion that shared aspects that were *both medical and personal* in nature (93/735 content units, 13%). Next, we provide representative quotes to describe the kinds of personal health issues that emerged as subtopics in each topic category.

Medical Topics

Topics that were medical in nature involved problems or concerns about constructs or processes that are tied closely to the health care delivery system, biomedical research, and health professionals’ work. Medical topics often reflected discussion that could stimulate an improved understanding of the medical domain or strategies to better fit one’s life to the health care delivery system. Common clusters of subtopics that fell in the medical category with representative examples include the following:

(a) Deciding on health care teams, treatments and procedures, and research trial enrollment

- “Tackling the selection of our health care team”
- Being in a “dilemma about reconstruction”
- Dealing with competing recommendations from different doctors
- Deciding between “radiation and tamoxifen or chemo and radiation”
- Deciding whether to have a biopsy
- Determining eligibility for participation in “genetic research.”

(b) Understanding biomedical concepts and processes, clinical treatments, procedures and tests, side effects, and biomedical research

- Understanding “cancer staging” and other medical terminology
- Determining whether a “bone scan” is a typical part of cancer care
- Discussing a “pathology report question”
- Uncovering how the diagnostic process typically flows
- Understanding effects of Arimidex on cholesterol.

(c) Managing interactions with health care professionals

- “Good care is also about communication”
- Determining when to seek a second opinion
- “What can I expect” for my upcoming appointment
- Understanding considerations doctors make when recommending treatments.

(d) Managing information to collaborate with clinicians or to understand biomedical issues

- Tracking medications, pain, or side effects to share with your health care provider
- “Questions to ask your oncologist”
- Preparing information for appointments
- “I was supposed to take the reports to a general surgeon, but I wonder if this is necessary, since nothing was found?”
- Discussing a research article on the effectiveness of Herceptin.

Personal Topics

Topics that were personal in nature involved problems or concerns around constructs or processes that are tied closely to one’s personal life, including ongoing responsibilities and day-to-day activities associated with family, friends, the home, work, and health-related activities that occur outside of the health care delivery system. Personal topics often reflected discussion that could stimulate the development of practical strategies to fit health management into one’s ongoing life. Common clusters of subtopics that fell in the personal category with representative examples include the following:

(a) Managing life at home

- Recovering from medical treatments and procedures: “What to expect following surgery”
- Keeping up with family and household responsibilities: Sharing my experiences with hospice planning
- Maintaining oversight of legal, financial, and insurance issues: how to “keep track of your medical expenses.”

(b) Managing life at work

- Shifting your work load during treatment: “Worry about health, not your job performance”
- Considering the impact of cancer on work prospects and insurance: “We are uncertain about what would happen if we were to change employers or careers”
- Interacting with coworkers, colleagues, or clients during treatment: “Maintaining a work persona”
- Deciding whether to work during treatment: “Have any of you gone back to work during part of your chemo?”

(c) Managing the emotional response to cancer

- Coping with anxiety, anger, depression, and fear
- “Finding ways to cope with the emotional roller coaster”
- Managing the “fear of recurrence”
- “Humor is a necessary healing component.”

(d) Managing interactions with one’s social network

- “What to tell your children”
- The “Fears of our loved ones”
- Getting help to find others with a similar diagnosis
- “Letting our partners know what we expected and needed.”

(e) Managing personal tasks and projects

- Managing lifestyle and self-care, including diet, exercise, and meditation: the “Dixie cup method” to organize medications and supplements; dealing with scalp pain while losing one’s hair; seeking a good “self-massage video”
- Focusing on spirituality and hobbies
- Managing personal health information: “Identifying and utilizing information resources.”

(f) Managing advocacy and volunteer work

- Sharing notices about upcoming cancer-related fundraisers
- Reaching out to others: “Breast cancer has helped us discover our mission and taught us that we can make a difference.”

Both Medical and Personal

We placed content units that shared medical and personal topics fairly equally into the overlapping category *both medical and personal*. Some examples that fell in this category include the following:

(a) Understanding biological concepts and processes AND Managing interactions with one’s social network

> *The risk of developing breast cancer is higher for women who have family history of cancer...Telling our mothers about our diagnosis and anticipating their responses were a source of major concern and anxiety for all of us.*

(b) Managing interactions with health care professionals AND Managing personal tasks and projects

> *After all of your treatments are completed...write down how you feel in general terms about once a month. Not only will it assist you in communicating with your doctor but it will also give you a barometer by which to measure your progress.*

(c) Deciding on treatments and procedures AND Managing work life

> *Schedule your chemotherapy right before the weekend so that it interferes with work as little as possible.*

Recommendations Offered Across Content Sources

Recommendations offered across content units fell into 4 major types: action strategies (1589/7209 recommendations, 22%), knowledge (3024/7209 recommendations, 42%), perspectives (390/7209 recommendations, 5%), and information resources (2206/7209 recommendations, 31%). Whereas recommendations in the form of action strategies offered procedural knowledge through suggested tasks (ie, "things to do"), recommendations in the form of knowledge offered declarative knowledge through facts and opinions (ie, "things to know"). Perspectives recommended attitudes or belief systems (ie, ways of believing or approaching situations), and information resources recommended tangible artifacts (ie, "things to obtain and use"). We describe each type of recommendation in greater detail below.

During our analysis we also recognized style differences between the recommendations; some action strategies, knowledge, and perspectives were direct, or *prescriptive*, in nature (ie, "you should..."), while others were carried implicitly through the *narrative* style of personal stories (ie, "when I was in your situation, I..."). We also recognized occasional overlap between action strategies, recommended knowledge, and perspectives. For example, taking action (ie, action strategy) can rely on acquiring knowledge. We can also learn (ie, knowledge) through our actions. Although chunks of text in a content unit could contain combinations of related recommendations in these different forms, we broke the text down (eg, sentence level) to code each recommendation with the type it best fit rather than allowing overlap between these types of recommendations.

Action Strategies: Things to Do

Action strategies are recommended tasks to deal with a personal health issue. This procedural knowledge about specific and actionable tasks can contribute toward solving a problem—for example, "It may be helpful for her to meet with a radiation oncologist before the surgery to discuss the pros and cons [of mastectomy versus lumpectomy]." Prescriptive action strategies reflected direct instructions for prescribed actions—for example, "One piece of advice is to use a pillow or some sort of padding for your over the shoulder seatbelt [following mastectomy]." In contrast, narrative action strategies reflected personal stories—for example, "What helped me [when deciding between mastectomy and lumpectomy] was searching the Internet for photos and having various women who had been through it send me their [postsurgery] photos."

Knowledge: Things to Know

Recommended knowledge refers to informative facts and opinions that one can learn to deal with a personal health issue. Unlike action strategies that represent tasks, recommended knowledge reflects declarative descriptions of concepts or ideas a person comes to understand—for example, you should know that "both lumpectomy and mastectomy require anesthesia." Prescriptive knowledge included subjective perceptions, opinions, or prescribed facts—for example, "Your pathology report [describes the] tumor and...nearby lymph nodes"). We also considered descriptive explanations and term definitions as prescriptive—for example, "Staging breast cancer is the process of...". Knowledge that was narrative in style included recommendations—for example, "The surgery for the tissue expander was painful for me."

Perspectives: Ways of Believing or Approaching Situations

Perspectives are recommended belief systems, attitudes, or philosophies that drive an overarching approach for dealing with a personal health issue, such as putting one's faith in God or acting as a strong advocate for oneself. In contrast to action strategies and recommended knowledge, perspectives reflect high-level and generalized beliefs, values, or attitudes toward an overarching experience—for example, "I made this [treatment] decision to be comfortable with my body." We differentiated between perspectives that were prescriptive—for example, "I know it's hard but I think you are actually mourning your old life...you need to give yourself time to do that"—and perspectives that were narrative in style—for example, "One of my big moments came when I really understood that everything will always be different from the 'before' and that I must adjust to that."

Information Resources: Things to Obtain and Use

Information resources are recommendations for obtaining and using specific tools and tangible items to deal with a personal health issue. A diverse range of information resources were recommended (see Table 3), including books, contact information (eg, for health professionals, health organizations, and local services), magazines and magazine articles, multimedia (eg, videos, graphs, figures, audiotapes, calculators), news articles, poems and quotes, academic journals and research articles, templates, webpages, and miscellaneous types of publications (eg, conference papers, pamphlets, brochures, white papers, and recipes). We were struck by the diversity of recommended webpages and multimedia. Webpages ranged from cancer-related organizations to personal websites and blogs. Multimedia ranged from static figures and graphs to audio, video, and interactive programs.

Templates, which included structured lists, tables, and worksheets for correspondents to personalize by filling them in, were an unexpected type of information resource. Templates reflect an embodiment of expertise that offer scaffolding to organize thoughts or actions surrounding personal health issues, such as tracking one's health status, side effects, medical expenditures, or day-to-day events, recording research evidence on treatments, preparing for medical procedures, or assessing one's personal level of resilience, pain, or nutrition. In contrast to clinician-oriented templates (eg, drain logs for patients to record and communicate postsurgery recovery to clinicians) that draw upon professional expertise, templates created by patients draw upon patients' personal health experiences. For example, patient book 2 and the clinician book both offered templates that outline considerations for choosing a clinician or care team. Both templates suggested assessing clinicians' communication style, their involvement with clinical studies, and the ability to tape-record visits. However, the patient book also reflected the patient's experience by recommending consideration of clinicians' personal character, professional reputation, availability, and payment options. In contrast, the clinician book reflected the clinician's experience by recommending consideration of clinicians' explanations for clinical tests, their interactions around complimentary and alternative treatments, and whether they are threatened when patients bring in information from the media to discuss.

Differences in Topics Discussed Between Sources of Expertise

Although sources of patient expertise and sources of clinician expertise contained content units that spanned both medical and personal topics, the proportions of content units falling under each topic (ie, medical, personal, or both medical and personal) differed significantly between those sources (χ^2_2[N = 735] = 233.4, $P < .001$). This significant difference held when we compared topics discussed in message boards alone (χ^2_2[N = 300] = 91.2, $P < .001$) and in books alone (χ^2_2[N = 435] = 168.7, $P < .001$). On average, patient sources discussed more personal topics and clinician sources discussed more medical topics. For example, the maximum proportion of content units from patient sources that discussed medical topics was 64% (32/50) on message board C, whereas the minimum for clinician sources was 74% (165/225) in the clinician book. In contrast, the minimum proportion of content units from patient sources that discussed personal topics was 26% (13/50) on message board C, whereas the maximum proportion for clinician sources was only 12% (28/225) in the clinician book (see Table 4).

Although sources of patient expertise showed a high proportion of personal topics relative to clinician sources, the degree to which personal topics were discussed varied across individual books and message boards. For example, 58 out of 79 content units (73%) from the patient book 1 contained personal topics, whereas only 13 out of 50 content units (26%) from patient message board C did so (see Table 4). In contrast, the sources of clinician expertise were both predominantly medical in focus. Only 28 out of 225 content units (12%) from the clinician book discussed personal topics and only 3 out of 150 content units (2%) from the ask-the-doctor message board did so (see Table 4). Although patient sources were more variable, the minimum proportion of personal topics discussed in the patient sources (13/50 content units, 26%, from patient message board C) was more than twice the maximum proportion of personal topics discussed in clinician sources (28/225 content units, 12%, from the clinician book).

The most common medical topic discussed across all sources was "understanding biomedical concepts and processes," making up 49 out of 150 content units (33%) discussed in patient message boards and 6 out of 210 content units (3%) in patient books, as well as 102 out of 150 content units (68%) discussed in the ask-the-doctor message board and 145 out of 225 content units (65%) in the clinician book (see Table 3). Within this medical subtopic, correspondents on patient message boards discussed clinical procedures and side effects, test results, and research or news articles about breast cancer. For example, after describing their medical history one correspondent asked "what is DIEP reconstruction?" Another correspondent asked whether "routine bone scans" are a standard part of long-term follow-up for breast cancer metastasis. Another correspondent looked for other patients in a similar situation to double-check what her "next step should be" in treatment. Many correspondents, particularly from message boards B and C, discussed research or news articles about various breast cancer treatments. In contrast, discussion on the ask-the-doctor message board revolved mostly around understanding the diagnosis and prognosis of breast cancer. For example, after describing aspects of their medical histories, correspondents commonly asked whether symptoms they were experiencing, such as "burning," "aching," or "swollen" breasts, could signal breast cancer.

The most common personal topic discussed across all sources was "managing personal tasks and projects," making up 16 out of 150 content units (11%) among patient message boards and 86 out of 210 content units (41%) in patient books, as well as 1 out of 150 content units (<1%) in the ask-the-doctor message board and 19 out of 225 content units (9%) in the clinician book (see Table 3). All but 1 content unit from clinician sources came from the clinician book and covered issues ranging from exercise, meditation, and diet to prostheses. The 1 content unit from the ask-the-doctor message board that discussed this personal topic pertained to dietary advice to "boost" blood counts. Content units from the patient message boards commonly pertained to self-care, such as managing hair loss and seeking a "good wig shop." In contrast, the patient books focused more on gathering and organizing information to play an active role in health care and self-care activities, such as nutrition and poetry writing.

Differences in Recommendations Offered by Sources of Expertise

Although content units from both patient and clinician sources offered recommendations falling under all 4 types (action strategies, knowledge, perspectives, and information resources), the proportions of those types differed significantly between patient and clinician sources (χ^2_3[N = 7209] = 482.1, $P < .001$). On average, patient sources offered a greater proportion of action strategies and perspectives but a smaller proportion of knowledge than clinician sources. However, both types of sources offered similar proportions of information resources on average. For instance, the proportion of perspectives from patient sources ranged from 46 out of 1365 recommendations (3%) in patient book 2 to 86 out of 555 of recommendations (15%) in message board A. In contrast, a maximum of only 79 out of 3573 recommendations (2%) from the clinician book offered perspectives (see Table 4). Furthermore, the maximum proportion of knowledge from patient sources was 368 out of 838 recommendations (44%) from patient book 1, whereas the minimum proportion of knowledge from clinician sources was 1753 out of 3573 recommendations (49%) from the clinician book (see Table 4).

When we compared message boards alone and books alone, we also found significant differences in the types of recommendations offered between patient message boards and the ask-the-doctor message board (χ^2_3[N = 1435] = 153.5, $P < .001$), as well as between the patient and clinician books (χ^2_3[N = 5774] = 274.9, $P < .001$). Table 3indicates that the patient message boards offered larger proportions of action strategies, perspectives, and information resources than the ask-the-doctor message board. Similarly, the patient books offered larger proportions of action strategies, perspectives, and information resources than the clinician book.

When we delved further into the styles used to express recommendations, we found that action strategies, knowledge, and perspectives were frequently expressed implicitly through personal narratives in sources of patient expertise compared with the prescriptive style that was common to sources of clinician expertise. This difference in style was significant for action strategies (χ^2_1[N = 1589] = 281.4, $P < .001$), knowledge (χ^2_1[N = 3024] = 621.0, $P < .001$), and perspectives (χ^2_1[N = 390] = 49.1, $P < .001$). Nearly half of all action strategies from patient sources were narrative in style, whereas almost all action strategies from clinician sources were prescriptive in style (see Table 3). The same pattern holds between sources for recommended knowledge and recommended perspectives.

Although sources of patient expertise and sources of clinician expertise offered similar proportions of information resources on average, the types of information resources that were most commonly exchanged differed between those sources. For example, sources of patient expertise offered more webpages, poems, quotes, and news articles, but fewer books, contact information, and academic journals or research articles than sources of clinician expertise (see Table 3).

Discussion

Results from this analysis show that patient expertise differs significantly from clinician expertise in topic (medical, personal, or both), type of recommendation (action strategies, knowledge, perspectives, and information resources), and style of recommendation (narrative vs prescriptive). Sources of clinician expertise were predominately medical in topic, knowledge-oriented in type, and prescriptive in style, whereas sources of patient expertise contained more personal topics that were carried through narrative-style action strategies and perspectives. These findings suggest that patients, by sharing their expertise about personal health, meet an important information need unmet by clinician sources. Our findings extend prior analyses of patient interaction with supportive evidence that differentiates patients' experiential knowledge about personal health from the medical expertise of professionals. This contribution enhances our understanding about the fundamental nature of patient expertise and guides the design of peer-support tools that facilitate patient-expertise sharing.

The Unique Nature of Patient Expertise

Differences in the characteristics of patient expertise and clinician expertise support the notion that patients and health professionals possess different domains of health expertise [51]. Rather than filling the role of an amateur doctor (ie, claiming professional-like medical knowledge about the treatment of disease without having professional training), the experiential knowledge offered by patients appears to focus on coping with highly personal issues drawn from the context of daily life. This characterization of expertise over managing the personal side of health supports the claim that such knowledge is gained not through professional training, but rather through the trial and error of managing the lived experience of illness [1,3]. Reports on the expertise of patient groups who manage conditions other than cancer, such as epilepsy [24], share a similar illustration of the strength of expertise on the personal side of health. Although some patient expertise appears condition specific (eg, tips for managing hair loss from chemotherapy), other expertise appears transferable between patient groups (eg, what to look for in a clinician).

In addition to clinician expertise obtained from health professionals, patients are finding new ways to reach out to other patients to exchange complementary personal health advice based on their own experiences through collaborative tools on the Internet [27-31]. Our findings suggest that patients are filling an important and valuable function that is not fully served by traditional health care and medical information delivery models that lack focus on the personal side of health. Thus, patients could benefit from informatics tools designed to help them share their expertise with one another. Our findings provide a strong foundation for designing new patient-centered tools that meet patients' needs for sharing expertise with peers.

Design Implications for Patient Expertise-Sharing Tools

Patient expertise-sharing tools are technologies that bring patients together to interact and exchange their personal health knowledge. Enhancements to tools that patients already use to exchange personal health information, such as health-related social software [27-31], are sensible targets for facilitating patient-expertise sharing. Blogs, wikis, forums, social networking tools, and other collaborative tools are being increasingly used by patients to exchange personal health information [28,30,31]. For example, patients contribute and rate recommended websites on health-related wikis [52], exchange health-related information on Facebook [53-56], track and share their condition-specific symptoms and treatments through profiles with health-specific social networking tools [32], and search member directories to find patients who share the same diagnosis [57,58]. Given the high prevalence of seeking health advice from peers on the Internet [30,31], designers should explore enhancements that will make the expertise of patients more prominent, explicit, and accessible.

Design efforts to facilitate patient-expertise sharing can offer patients opportunities to interact with these collaborative technologies in ways that extend beyond the traditional, text-based message boards of the past. For example, participatory design work illustrates patients' strong desire for online collaboration and networking tools, such as Facebook [56] or Myspace [59], to connect and share common illness experiences and valuable health resources [60,61]. Our findings on the fundamental nature of patient expertise provide a strong foundation for such innovative design efforts. In particular, as we detail in the following subsections, our findings have bearing on enhancing social software by including support for (1) collaboratively managing shared resources, (2) locating patient expertise, and (3) safeguarding against misinformation (Table 5).

Table 5. Design opportunities to facilitate patient-expertise sharing

Type of support	Design feature
Collaboratively managing shared resources	Common space to share and interact with varied resources
	User-generated tags and folksonomies that are meaningful to patients
	Methods for rating and recommending tailored resources
	Narrative and template formats for sharing experiences and expertise
Locating patient expertise	Detailed user profiles that illustrate areas of experience and expertise
	Methods for people finding and matchmaking
	Analytic tools for identifying topics of interest from user contributions
Safeguarding against misinformation	Features that preserve natural safeguarding strategies in a public context
	Change log to provide audit trail of corrections to content
	Vetting features to note affirmation, rebuttal, or reference sources

Collaboratively Managing Shared Resources

Designers should focus on developing common spaces for patients to manage the multitude of information resources they share together. The wide range of information resources (eg, webpages, books, articles, and multimedia) that patients exchange suggests the need for tools that enable patients to work together to create, annotate, store, share, and reuse content across a diverse range of formats and topics. Patients need help managing this full range of resources they recommend to and garner from one another. Collaborative features of social software, such as user-generated tags to organize content shared through a wiki, have the potential to support this need for collaboratively managing shared resources. For example, Weiss and Lorenzi [62] synthesized community wisdom about local cancer programs and services using collaborative Web-based tools for sharing community-based cancer resources. Others are developing recommendation systems for patients [63], such as a tag-based recommendation system that leverages community ratings of health content to rank tailored suggestions it provides to users [64].

Collaborative recommendation systems like these help users share their expertise by rating resources and benefit from each other's views, opinions, and experiences through collaborative filtering [65]. We envision this collaborative space much like an updated version of an "expert patient knowledge base" [24]. Our findings on the breadth of information resources that patients exchange (eg, books, contacts, news and academic articles, poems and quotes, and recipes) suggest that incorporating collaborative recommendation methods into health-related social software could help users work together to manage this range of content and recommend useful resources to one another.

Given the range of medical and personal topics discussed among patients in our analysis, medically oriented resources (eg, medical dictionaries and patient information summaries) would certainly make up a valuable component of collaboratively managed collections of patient resources. However, the prominence of personal topics (eg, tracking medical expenses, working during treatment, what to tell your children, and selecting wigs) suggests that a fundamental component of such collections must include nonclinical resources as well. These resources should provide advice on personal topics related to work, family, the home, and social relationships in the context of illness. For example, one of the threads we analyzed consisted of dozens of suggestions from patients on considerations to make when writing an "end of life memoir" for family members (eg, your favorite books, family heirlooms, hobbies you enjoy, and world travels). Other examples include discussions about favorite "juicer recipes" and "experiences with sick-leave policies." The breadth of these personal topics could link to a full range of relevant information resources from multiple domains (eg, medicine, law, social work, art, cooking, community resources, and finance). Users could tag and annotate these resources collaboratively in ways that capture important contextual ties to their specific experiences and facilitate later reuse by other users [66].

A medical library model [67] might provide only a partial fit for organizing the breadth of information needs met by the collections of resources shared among patients [68]. Our findings expand on work that shows a poor mapping between many concepts that patients use and controlled medical terminologies, such as the Unified Medical Language System (UMLS) metathesaurus [69]. For example, UMLS was not designed to capture many of the nonmedical concepts, such as the family, work, and social matters, for which patients turn to one another for help. Alternative organizational structures could allow users to compile shared information resources in personally meaningful, yet diverse, ways. For example, tools could encourage users to create *consumer health folksonomies* [70] to organize documents around their own conceptualization of health-related issues. Such folksonomies might resemble collaboratively constructed and tag-based systems that have emerged in other contexts, such as Web bookmarking [71].

Our findings provide additional insights for supporting collaboratively managed collections of resources. The common style of personal stories used to express patient expertise (see also [21,72]) suggests the potential value of narrative-based formats, such as "war stories," that have been a highly valued format for sharing expertise in some professional work settings [73]. Repositories of personal health stories that are surfacing through social software, such as personal blogs [29] and CarePages' "Stories of Inspiration" [74], might facilitate a natural expression of patient expertise and provide contextual detail upon which to create experience maps that guide patients' problem solving surrounding specific health situations. Vetting features (ie, ratings, awards, and crowdsourcing techniques), whereby users associate comments or affirmations with personal stories, could help users to assess the fit of implicit advice those personal stories provide to their own health situations. Furthermore, patient expertise in the key form of action strategies could be exchanged through "how to" pages [75] that communicate strategies for dealing with personal health issues, or through templates that provide guidance by scaffolding action plans around personal health activities (eg, a preparing-for-surgery checklist). Patients could later recommend useful personal health practices through such templates to other patients.

Locating Patient Expertise

Designers should focus on developing tools that help patients find and connect with other patients who have specific kinds of experiences or expertise. As health-related use of social software grows [28,30,31], patients will need help locating other patients with the specific expertise they seek. During our analysis, for example, we observed patients posting requests to find other users with specific experiences or wisdom (eg, has anyone on this forum dealt with this particular rare side effect?). A common complaint about message boards is the challenge of determining who knows what, because the expertise of users gets lost within the volumes of threads in the forum. Design enhancements that make users' requests and their expertise more explicit [8,34] will help to overcome these limitations.

Features of social networking tools, such as user profiles, can help bring users' expertise to the surface, enabling a targeted search for patients with specific health and personal characteristics [34]. For example, a closer match in lifestyle and belief system leads to peer support that is perceived as more helpful [76]. Yet, most user profiles of social software are limited to a single health condition and a small set of demographics, and provide little if any indication of the kinds of expertise users can offer. Whether a user finds other patients for advice by posting personal health data to their personal profile [32] or by posting forum questions, this common broadcast strategy works only if people with the requisite expertise notice and respond. Once the user garners that expertise, they must determine the suitability of peers who provided the expertise for meeting their specific needs. Although traditional user profiles and broadcast mechanisms help, findings from our analysis point to expertise cues that could facilitate locating patient expertise through people-finding or matchmaking features. For example, user profiles could display the topics a user commonly discusses or information resources (eg, webpages, articles, and books) that users recommend one another but otherwise get buried within threads [34]. Consider a message board thread we analyzed in which a patient provided an extensive critique on a recent article about access to breast cancer treatments. Given a user profile, this patient could post the recommended article and her critique, thus making them visible and easily accessible to other users later.

We could also leverage the solid foundation of expertise-sharing research conducted in other settings to make progress in this important design direction of supporting the locating of patient expertise. For example, when confronted with an unfamiliar problem, people in professional work settings locate needed expertise by identifying potential sources (eg, other people and artifacts) and selecting specific sources to approach for help [77]. This practice of locating expertise has informed the design of tools that help professionals find colleagues with the desired, and often specialized, knowledge within a professional organization through features, such as user profiles and social networks [78-80]. With guidance from related work on expertise locators and our understanding of the characteristics of patient expertise, we can enhance social software and make it easier for patients to identify other users who have the specific patient expertise they need. To this end, we envision matchmaking tools that could support the following scenario:

Sally seeks advice about whether to work during chemotherapy. She wants to locate a patient who has already dealt with this situation (eg, "I want to find another mother of school-aged children who worked throughout chemotherapy"). She enters age, gender, and condition into a directory search service offered by a social networking tool for cancer patients. Unfortunately, she is overwhelmed by the large number of user profiles the system returns, which she must now manually review to find another user with the specific characteristics she is looking for. In particular, Sally needs awareness of not only the health condition and demographics of other users, but details about their specific knowledge and health-related experiences to answer questions, such as "Does this person have the experience I am interested in? If so, how recently? What is their experience level?"

Enhanced features that make specific and detailed health experiences of users more prominent could make Sally's work much easier and tailored to her needs. For example, it was common for correspondents in the message boards we analyzed to preface their thread postings with detailed descriptions of their health experiences (eg, "I'm a 4-year survivor..."). Such details could be combined with a larger range of medical and nonmedical characteristics [32] to extend user profiles. Our findings also reveal a range of personal topics that patient expertise reflects (eg, managing health-related issues that connect to work or the home). We could also facilitate patient-expertise sharing by incorporating users' topical expertise, based on the topics they discuss, into their user profiles. Access to both topical expertise and specific health experiences of other users provides important contextual cues that patients need to locate expertise that meets their specific needs [76,81].

Safeguarding Against Misinformation

Designers should also focus on features that preserve and encourage self-correction, self-monitoring, and other natural safeguarding strategies already used by patients online. Some might fear that enhancing informatics support for patient-expertise sharing could lead to the spread of mistaken, misinterpreted, outdated, incomplete, or otherwise poor-quality information. Although the potential for medical misinformation certainly exists, studies have examined patient interactions in online health communities and found minimal levels of medical misinformation [24,82-85].

We did not assess the accuracy of information exchanged in the patient message boards that we analyzed, but our observations of message board correspondents were consistent with previous research on strategies used to actively safeguard against the potential for misinformation, such as self-correction [83,86] and warnings from watchful members [23]. We also observed correspondents using additional safeguarding strategies, including source referencing (eg, “my oncologist told me that...”), advice prefacing (eg, “everyone has a different experience, [but this is what happened to me]”), rebuttal (eg, “our support group has many women’s experiences that prove otherwise”), and affirmative vetting of advice offered by other correspondents (eg, “I agree with all the ladies so far”).

Our observations point to the importance of preserving functionality that encourages patients’ natural misinformation-safeguarding strategies, such as vetting features within a public context, as health-related social software evolves to support patient-expertise sharing. In particular, our observations suggest support for audit trails that make content changes explicit and easy to log, and simple vetting features for noting affirmation or rebuttal (eg, thumps up/down) and for referencing source material.

Study Limitations and Directions for Future Research

The characteristics of patient expertise we present are derived from a deep exploration of content from selected message boards and books in the context of cancer. The codebook resulting from our analysis is necessarily shaped by diverse interests and viewpoints of book authors and message board correspondents from the sources we analyzed. For example, expertise captured from a book written by a single author might not be as diverse or transferable as the expertise of multiple voices from a book coauthored by several people. Thus, our findings might not capture the breadth of expertise across the wide array of resources available to patients. For example, message boards reflect patients’ information needs through discussion that is initiated by patients’ own questions or offers of support. In comparison, books written by cancer survivors could provide a less direct reflection of authors’ and publishers’ perceptions about patients’ needs. Furthermore, individual content sources vary in their predominance of personal topics. Despite these differences, we found a strikingly similar distinction between the patient expertise in both books and message boards and the expertise in clinician sources. While our findings illustrate the volume of patient discussion beyond the medical realm of personal health, additional research is needed to tease out issues of misinformation and deeper distinctions within medical topics discussed.

Although our in-depth effort was scoped to small diverse samples, the work yielded rich descriptions that provide a solid basis for understanding patient expertise as a critical facet within the breadth of patients’ information needs. Details of our codebook point to a range of information needs related to the personal side of health and contribute to a holistic view of the patient. Given the experiential nature of patient expertise, it is plausible that the characteristics we ascribe to this specialized form of knowledge are also reflected by the experiential knowledge that people develop from personally managing other health situations, such as diabetes, heart disease, or pregnancy. Although aspects of patient expertise we identified are specific to cancer, other aspects could be widespread and shared by patients with other conditions [24]. Future research could explore which characteristics of patient expertise reflected by our codebook are transferrable to these other health contexts.

Future work could also explore how our design implications play out within patients’ expanding space of social participation on the Internet [87]. Although our content analysis captured only a sampling of content sources, our enhanced understanding of patient expertise points to innovative directions for the design of peer-support tools that facilitate patient-expertise sharing. Our findings also provide insights for the design of tools that encourage information sharing between patients and health professionals, such as integrating a broad range of personal factors with health care planning in the context of shared decision making [10,11], as well as for the design of tools that bridge the expertise of patients and professionals [52].

Conclusion

Our findings demonstrate that patient expertise differs significantly from the expertise of clinicians in topic, type, and style. Neither increasing the amount of time that patients spend with health care providers nor training patients with medical knowledge to become amateur doctors appears sufficient to meet the needs for patient expertise. Instead, we offer alternatives in the form of design directions for facilitating patient-expertise sharing with health-related social software. Patients provide other patients with a unique and valued information resource that complements expertise provided by health professionals. Patients deserve informatics support that can fill the breadth of their health information needs by facilitating this patient-expertise sharing.

Acknowledgments

This work was supported by the National Institutes of Health (NIH) grant R01 LM009143. We thank Ching-Ping Lin for her assistance with inter-coder reliability testing, John Gennari and Richard Boyce for their assistance in preparing this manuscript, and David McDonald, Huong Nguyen, William Jones, and the iMed research group at the University of Washington for their invaluable discussions and feedback throughout this project. We thank MedHelp International (http://www.medhelp.org) for their support.

Conflicts of Interest

None declared

References

1. Paterson B, Thorne S. Developmental evolution of expertise in diabetes self-management. Clin Nurs Res 2000 Nov;9(4):402-419. [Medline: 11881697]
2. Paterson BL, Thorne S, Dewis M. Adapting to and managing diabetes. Image J Nurs Sch 1998;30(1):57-62. [Medline: 9549943]
3. Benner P, Wrubel J. The Primacy of Caring: Stress and Coping in Health and Illness. Reading, MA: Addison Wesley; 1988.
4. Petersen A. The best experts: the narratives of those who have a genetic condition. Soc Sci Med 2006 Jul;63(1):32-42. [doi: 10.1016/j.socscimed.2005.11.068] [Medline: 16431006]
5. Helgeson VS, Cohen S. Social support and adjustment to cancer: reconciling descriptive, correlational, and intervention research. Health Psychol 1996 Mar;15(2):135-148. [Medline: 8681922]
6. Jacobson DE. Types and timing of social support. J Health Soc Behav 1986 Sep;27(3):250-264. [Medline: 3772062]
7. Weiss RS. The provisions of social relationships. In: Rubin Z, editor. Doing Unto Others: Joining, Molding, Conforming, Helping, Loving. Englewood Cliffs, NJ: Prentice-Hall; 1974:17-26.
8. Campbell HS, Phaneuf MR, Deane K. Cancer peer support programs-do they work? Patient Educ Couns 2004 Oct;55(1):3-15. [doi: 10.1016/j.pec.2003.10.001] [Medline: 15476984]
9. Hoey LM, Ieropoli SC, White VM, Jefford M. Systematic review of peer-support programs for people with cancer. Patient Educ Couns 2008 Mar;70(3):315-337. [doi: 10.1016/j.pec.2007.11.016] [Medline: 18191527]
10. Berry DL, Ellis WJ, Woods NF, Schwien C, Mullen KH, Yang C. Treatment decision-making by men with localized prostate cancer: the influence of personal factors. Urol Oncol 2003;21(2):93-100. [Medline: 12856636]
11. 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. [Medline: 10428003]
12. Wykurz G, Kelly D. Developing the role of patients as teachers: literature review. BMJ 2002 Oct 12;325(7368):818-821 [FREE Full text] [Medline: 12376445]
13. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]
14. Shaw BR, Han JY, Baker T, Witherly J, Hawkins RP, McTavish F, et al. How women with breast cancer learn using interactive cancer communication systems. Health Educ Res 2007 Feb;22(1):108-119 [FREE Full text] [doi: 10.1093/her/cyl051] [Medline: 16829544]
15. Diamond RJ, Little ML. Utilization of patient expertise in medication groups. Psychiatr Q 1984;56(1):13-19. [Medline: 6522524]
16. Burdick D. Rehabilitation of the breast cancer patient. Cancer 1975 Aug;36(2):645-648. [Medline: 808265]
17. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999 Jan;37(1):5-14. [Medline: 10413387]
18. Weinberg N, Schmale J, Uken J, Wessel K. Online help: cancer patients participate in a computer-mediated support group. Health Soc Work 1996 Feb;21(1):24-29. [Medline: 8626154]
19. Sharf BF. Communicating breast cancer on-line: support and empowerment on the Internet. Women Health 1997;26(1):65-84. [doi: 10.1300/J013v26n01_05] [Medline: 9311100]
20. Klemm P, Hurst M, Dearholt SL, Trone SR. Gender differences on Internet cancer support groups. Comput Nurs 1999;17(2):65-72. [Medline: 10194883]
21. Høybye MT, Johansen C, Tjørnhøj-Thomsen T. Online interaction. Effects of storytelling in an internet breast cancer support group. Psychooncology 2005 Mar;14(3):211-220. [doi: 10.1002/pon.837] [Medline: 15386774]
22. Winefield HR. Support provision and emotional work in an Internet support group for cancer patients. Patient Educ Couns 2006 Aug;62(2):193-197. [doi: 10.1016/j.pec.2005.07.002] [Medline: 16139466]
23. Preece J. Empathetic communities: Reaching out across the web. interactions 1998;5(2):32-43. [doi: 10.1145/274430.274435]
24. Hoch DB, Norris D, Lester JE, Marcus AD. Information exchange in an epilepsy forum on the World Wide Web. Seizure 1999 Feb;8(1):30-34. [doi: 10.1053/seiz.1998.0217] [Medline: 10091845]
25. Meier A, Lyons EJ, Frydman G, Forlenza M, Rimer BK. How cancer survivors provide support on cancer-related Internet mailing lists. J Med Internet Res 2007;9(2):e12 [FREE Full text] [doi: 10.2196/jmir.9.2.e12] [Medline: 17513283]
26. Civan A, Pratt W. Threading together patient expertise. AMIA Annu Symp Proc 2007 Oct 11;2007:140-144. [Medline: 18693814]

27. Elkin N. icrossing. 2008. How America Searches: Health and Wellness URL: http://www.icrossing.com/research/how-america-searches-health-and-wellness.php [accessed 2010-12-27] [WebCite Cache ID 5vIvESPSv]
28. Sarasohn-Kahn J. California HealthCare Foundation. 2008. The Wisdom of Patients: Health Care Meets Online Social Media URL: http://www.chcf.org/publications/2008/04/the-wisdom-of-patients-health-care-meets-online-social-media [accessed 2010-12-26] [WebCite Cache ID 5vIvZHzFe]
29. Adams SA. Blog-based applications and health information: two case studies that illustrate important questions for Consumer Health Informatics (CHI) research. Int J Med Inform 2010 Jun;79(6):e89-e96. [doi: 10.1016/j.ijmedinf.2008.06.009] [Medline: 18701344]
30. Fox S, Jones S. Pew Internet. Washington, DC: Pew Internet & American Life Project; 2009 Jun. The social life of health information URL: http://www.pewinternet.org/~/media//Files/Reports/2011/Pew_P2PHealthcare_2011.pdf [accessed 2011-07-04] [WebCite Cache ID 5zxHcG9iD]
31. Fox S. Peer-to-peer healthcare. Washington, DC: Pew Internet & American Life Project; 2011 Jun. URL: http://www.pewinternet.org/Reports/2011/P2PHealthcare.aspx [accessed 2011-03-22] [WebCite Cache ID 5xNN2pQQP]
32. Frost JH, Massagli MP. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. J Med Internet Res 2008;10(3):e15 [FREE Full text] [doi: 10.2196/jmir.1053] [Medline: 18504244]
33. Lau AY, Kwok TM. Social features in online communities for healthcare consumers - A review. In: Ozok AA, Zaphiris P, editors. Online Communities and Social Computing: Third International Conference, OCSC 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, ... Applications, incl. Internet/Web, and HCI). Berlin, Germany: Springer; 2009:682-689.
34. Civan-Hartzler A, McDonald DW, Powell C, Skeels MM, Mukai M, Pratt W. Bringing the field into focus: user-centered design of a patient expertise locator. CHI Conf Proc 2010;2010:1675-1684. [doi: 10.1145/1753326.1753577] [Medline: 21243114]
35. Rimer BK, Lyons EJ, Ribisl KM, Bowling JM, Golin CE, Forlenza MJ, et al. How new subscribers use cancer-related online mailing lists. J Med Internet Res 2005 Jul 1;7(3):e32 [FREE Full text] [doi: 10.2196/jmir.7.3.e32] [Medline: 15998623]
36. Dasch K, Kendall J. What do newly diagnosed cancer patients discuss with survivors? Del Med J 2007 Dec;79(12):489-492. [Medline: 18240655]
37. Frost J, Massagli M. PatientsLikeMe the case for a data-centered patient community and how ALS patients use the community to inform treatment decisions and manage pulmonary health. Chron Respir Dis 2009;6(4):225-229. [Medline: 19858352]
38. Krippendorff K. Content Analysis: An Introduction to its Methodology. Thousand Oaks, CA: Sage Publications; 2004.
39. Crabtree BF, Miller WL. Doing Qualitative Research. Thousand Oaks, CA: Sage Publications; 1999.
40. Love SM, Lindsey K. Dr. Susan Love's Breast Book. Cambridge, MA: Perseus Publishing; 2000.
41. McCarthy P, Loren JA. Breast Cancer? Let Me Check My Schedule. Boulder, CO: Westview Press; 1997.
42. Willis J. The Cancer Patient's Workbook: Everything You Need to Stay Organized and Informed. New York, NY: DK ADULT; 2001.
43. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications; 1985.
44. Lowry R. Vassar College. 2010. Concepts and Applications of Inferential Statistics URL: http://faculty.vassar.edu/lowry/webtext.html [accessed 2010-12-27] [WebCite Cache ID 5vIvn6K8B]
45. Brownlow C, O'Dell L. Ethical issues for qualitative research in on-line communities. Disability & Society 2002;17(6):685-694. [doi: 10.1080/0968759022000010452]
46. Kraut R, Olson J, Banaji M, Bruckman A, Cohen J, Couper M. Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. Am Psychol 2004;59(2):105-117. [doi: 10.1037/0003-066X.59.2.105] [Medline: 14992637]
47. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: 11701577]
48. Moreno MA, Fost NC, Christakis DA. Research ethics in the MySpace era. Pediatrics 2008 Jan;121(1):157-161 [FREE Full text] [doi: 10.1542/peds.2007-3015] [Medline: 18166570]
49. Bruckman A. Teaching students to study online communities ethically. J Information Ethics 2006;15(2):82-98. [doi: 10.3172/JIE.15.2.82]
50. Finn J. An exploration of helping processes in an online self-help group focusing on issues of disability. Health Soc Work 1999 Aug;24(3):220-231. [Medline: 10505283]
51. Coulter A. After Bristol: putting patients at the centre. BMJ 2002 Mar 16;324(7338):648-651 [FREE Full text] [Medline: 11895824]
52. WEGO Health. WEGO Health.com. 2011 URL: http://www.wegohealth.com/ [accessed 2011-01-05] [WebCite Cache ID 5vW7EsPpz]
53. Ahmed OH, Sullivan SJ, Schneiders AG, McCrory P. iSupport: do social networking sites have a role to play in concussion awareness? Disabil Rehabil 2010;32(22):1877-1883. [doi: 10.3109/09638281003734409] [Medline: 20367328]
54. Bender JL, Jimenez-Marroquin MC, Jadad AR. Seeking support on facebook: a content analysis of breast cancer groups. J Med Internet Res 2011;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]

55. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME. It's not that I don't have problems, I'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In: Proceedings. 2011 Presented at: ACM 2011 conference on Computer supported cooperative work; 2011; Hangzhou, China. [doi: 10.1145/1958824.1958876]
56. Facebook. Facebook.com. 2011 URL: http://www.facebook.com [accessed 2011-07-05] [WebCite Cache ID 5zxKhEVqj]
57. Healia Health Communities and Support Groups. healia.com. 2008 URL: http://communities.healia.com/ [accessed 2011-01-04] [WebCite Cache ID 5vW7TfRVr]
58. PatientsLikeMe. patientslikeme.com. 2011 URL: http://www.patientslikeme.com/ [accessed 2011-01-04] [WebCite Cache ID 5vW7a3aYr]
59. Myspace. myspace.com. 2011 URL: http://www.myspace.com/ [accessed 2011-01-05] [WebCite Cache ID 5vW7hDO6y]
60. Weiss JB, Lorenzi NM. Building a shared vision for an online cancer survivorship community. AMIA Annu Symp Proc 2009 Nov 14;2009:690-694. [Medline: 20351942]
61. Skeels MM, Unruh KT, Powell C, Pratt W. Catalyzing social support for breast cancer patients. CHI Conf Proc 2010:173-182. [doi: 10.1145/1753326.1753353] [Medline: 21654894]
62. Weiss JB, Lorenzi NM. Synthesizing community wisdom: a model for sharing cancer-related resources through social networking and collaborative partnerships. AMIA Annu Symp Proc 2008 Nov 06:793-797. [Medline: 18999298]
63. Witteman H, Chignell M, Krahn M. A recommender system for prostate cancer websites. AMIA Annu Symp Proc 2008 Nov 06;2008:1177. [Medline: 18999034]
64. Khan SA, Cohall A, Kukafka R. A tag based recommendation engine to suggest information resources in an online community for health promotion. AMIA Annu Symp Proc 2008 Nov 06;2008:1002. [Medline: 18999253]
65. Terveen LG, Hill W. Beyond recommender systems: helping people help each other. In: Carroll J, editor. Human-Domputer Interaction in the New Millennium. New York, NY: ACM Press; 2001:487-509.
66. Diigo. Diigo Inc. 2011 URL: http://www.diigo.com/ [accessed 2011-01-05] [WebCite Cache ID 5vW7ltisz]
67. Cosgrove TL. Planetree health information services: public access to the health information people want. Bull Med Libr Assoc 1994 Jan;82(1):57-63. [Medline: 8136762]
68. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. J Am Med Inform Assoc 2006;13(1):24-29. [doi: 10.1197/jamia.M1761] [Medline: 16221948]
69. Keselman A, Smith CA, Divita G, Kim H, Browne AC, Leroy G, et al. Consumer health concepts that do not map to the UMLS: where do they fit? J Am Med Inform Assoc 2008;15(4):496-505. [doi: 10.1197/jamia.M2599] [Medline: 18436906]
70. Smith CA, Wicks PJ. PatientsLikeMe: Consumer health vocabulary as a folksonomy. AMIA Annu Symp Proc 2008 Nov 06;2008:682-686. [Medline: 18999004]
71. Delicious. delicious.com. 2011 URL: http://www.delicious.com/ [accessed 2011-07-05] [WebCite Cache ID 5zxLe9MKf]
72. Swift TL, Dieppe PA. Using expert patients' narratives as an educational resource. Patient Educ Couns 2005 Apr;57(1):115-121. [doi: 10.1016/j.pec.2004.05.004] [Medline: 15797160]
73. Orr JE. Talking About Machines: An Ethnography of a Modern Job. Ithaca, NY: ILR Press; 1996.
74. CarePages. CarePages, Inc. 2011 URL: http://www.carepages.com/ [accessed 2011-04-04] [WebCite Cache ID 5xhIHasDv]
75. Torrey C, McDonald DW. How-to web pages. Computer 2007;40(8):96-97. [doi: 10.1109/MC.2007.277]
76. Dunn J, Steginga SK, Occhiping S, Wilson K. Evaluation of a peer support program for women with breast cancer-lessons for practitioners. J Community Appl Soc Psychol 1999;9(1):13-22. [doi: 10.1002/(SICI)1099-1298(199901/02)9:1<13::AID-CASP488>3.0.CO;2-F]
77. McDonald DW, Ackerman MS. JustTalk to me: a field study of expertise location. In: CSCW '98: Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work. New York, NY: ACM; 1998:315-324. [doi: 10.1145/289444.289506]
78. McDonald DW, Ackerman MS. Expertise recommender: a flexible recommendation architecture. In: CSCW '00: Proceedings of the ACM Conference on Computer-Supported Cooperative Work. New York, NY: ACM; 2000:231-240. [doi: 10.1145/1316624.1316642]
79. Ehrlich K. Locating expertise: design issues for an expertise locator system. In: Ackerman MS, Pipek V, Wulf V, editors. Sharing Expertise: Beyond Knowledge Management. Cambridge, MA: MIT Press; 2003:137-158.
80. Ehrlich K, Lin C, Griffiths-Fisher V. Searching for experts in the enterprise: combining text and social network analysis. In: GROUP '07: Proceedings of the 2007 International ACM Conference on Supporting Group Work. New York, NY: ACM; 2007:117-126.
81. Civan A, McDonald DW, Unruh KT, Pratt W. Locating patient expertise in everyday life. GROUP ACM SIGCHI Int Conf Support Group Work 2009;2009:291-300. [Medline: 20953244]
82. van Uden-Kraan CF, Drossaert CHC, Taal E, Lebrun CEI, Drossaers-Bakker KW, Smit WM, et al. van de Laar MAFJ. Coping with somatic illnesses in online support groups: Do the feared disadvantages really occur? Computers in Human Behavior 2008;24:309-324. [doi: 10.1016/j.chb.2007.01.014]
83. 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] [doi: 10.1136/bmj.38753.524201.7C] [Medline: 16513686]

84. 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] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
85. Kelly W, Connolly V, Bilous R, Stewart A, Nag S, Bowes D, et al. Reader beware: diabetes advice on the web. Pract Diab Int 2002;19(4):108-110. [doi: 10.1002/pdi.340]
86. Winzelberg A. The analysis of an electronic support group for individuals with eating disorders. Computers in Human Behavior 1997;13(3):393-407. [doi: 10.1016/S0747-5632(97)00016-2]
87. Hesse BW, Hansen D, Finholt T, Munson S, Kellogg W, Thomas JC. Social Participation in Health 2.0. Computer (Long Beach Calif) 2010 Nov 11;43(11):45-52. [doi: 10.1109/MC.2010.326] [Medline: 21379365]

Abbreviations

UMLS: Unified Medical Language System

Edited by G Eysenbach; submitted 05.01.11; peer-reviewed by S Adams, J Frost; comments to author 14.03.11; revised version received 30.04.11; accepted 01.05.11; published 16.08.11.

Please cite as:
Hartzler A, Pratt W
Managing the Personal Side of Health: How Patient Expertise Differs from the Expertise of Clinicians
J Med Internet Res 2011;13(3):e62
URL: http://www.jmir.org/2011/3/e62/
doi:10.2196/jmir.1728
PMID:21846635

Original Paper

Risk Estimates From an Online Risk Calculator Are More Believable and Recalled Better When Expressed as Integers

Holly O Witteman[1,2], PhD; Brian J Zikmund-Fisher[1,2,3,4], PhD; Erika A Waters[5], MPH, PhD; Teresa Gavaruzzi[6,7], PhD; Angela Fagerlin[1,2,8,9], PhD

[1]Program in Health Communication and Decision Making, Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI, United States

[2]Division of General Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States

[3]Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[4]Risk Science Center, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[5]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, Saint Louis, MO, United States

[6]Leeds Institute of Health Sciences, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom

[7]Department of Developmental Psychology and Socialization, University of Padova, Padova, Italy

[8]VA Ann Arbor Center for Clinical Management Research, Ann Arbor, MI, United States

[9]Department of Psychology, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Holly O Witteman, PhD
Program in Health Communication and Decision Making
Center for Bioethics and Social Sciences in Medicine
University of Michigan
300 North Ingalls Building 7C27
Ann Arbor, MI, 48109-5429
United States
Phone: 1 734 615 2296
Fax: 1 734 936 8944
Email: hollywit@med.umich.edu

Abstract

Background: Online risk calculators offer different levels of precision in their risk estimates. People interpret numbers in varying ways depending on how they are presented, and we do not know how the number of decimal places displayed might influence perceptions of risk estimates.

Objective: The objective of our study was to determine whether precision (ie, number of decimals) in risk estimates offered by an online risk calculator influences users' ratings of (1) how believable the estimate is, (2) risk magnitude (ie, how large or small the risk feels to them), and (3) how well they can recall the risk estimate after a brief delay.

Methods: We developed two mock risk calculator websites that offered hypothetical percentage estimates of participants' lifetime risk of kidney cancer. Participants were randomly assigned to a condition where the risk estimate value rose with increasing precision (2, 2.1, 2.13, 2.133) or the risk estimate value fell with increasing precision (2, 1.9, 1.87, 1.867). Within each group, participants were randomly assigned one of the four numbers as their first risk estimate, and later received one of the remaining three as a comparison.

Results: Participants who completed the experiment (N = 3422) were a demographically diverse online sample, approximately representative of the US adult population on age, gender, and race. Participants whose risk estimates had no decimal places gave the highest ratings of believability ($F_{3,3384} = 2.94$, $P = .03$) and the lowest ratings of risk magnitude ($F_{3,3384} = 4.70$, $P = .003$). Compared to estimates with decimal places, integer estimates were judged as highly believable by 7%–10% more participants ($\chi^2_3 = 17.8$, $P < .001$). When comparing two risk estimates with different levels of precision, large majorities of participants reported that the numbers seemed equivalent across all measures. Both exact and approximate recall were highest for estimates with zero decimals. Odds ratios (OR) for correct approximate recall (defined as being within 50% of the original estimate) were, for one decimal place, OR = 0.65 (95% CI 0.49–0.86), for two decimal places, OR = 0.70 (95% CI 0.53–0.94), and for three decimal places, 0.61 (95% CI 0.45–0.81). Exact recall showed a similar pattern, with larger effects.

Conclusions: There are subtle but measurable differences in how people interpret risk estimates of varying precision. Adding decimal places in risk calculators offers little to no benefit and some cost. Rounding to the nearest integer is likely preferable for communicating risk estimates via risk calculators so that they might be remembered correctly and judged as believable.

(J Med Internet Res 2011;13(3):e54) doi:10.2196/jmir.1656

KEYWORDS

Risk; risk assessment; communication; risk communication; perception; risk perception; calculators, programmable; risk calculator; Internet; online

Introduction

Risk calculators abound online. Anyone with Internet access, a Web browser, some interest in their future health, and five minutes to spare can enter a few pieces of information about themselves and receive an assessment of their risk of human immunodeficiency virus infection [1], breast cancer [2], heart attack [3,4], diabetes [5,6], prostate cancer recurrence [7,8], or one of multiple options from a broad array of common conditions [9]. Online risk calculators have been heralded as a means to disseminate individualized risk prediction scores with an aim toward increased understanding of personal risk, improved health behavior, higher-quality decision making, and, ultimately, better population health [10].

However, questions remain about how to best design risk calculators to achieve the goals of risk communication. Different calculators vary significantly in terms of their adherence to best practices for risk communication [11], the amount of detail and clinical specificity in their questionnaires, and the level of precision that they offer in their risk assessments. For example, a 55-year-old woman might use two different Web-based calculators to check her risk of breast cancer in the next 10 years and find that, according to one, her risk is 2.1% [12] while, according to another, it is 5.05399% [13]. (See Multimedia Appendix 1 for details.)

Robust underlying models may enable calculators to give precise risk estimates. However, it is not known whether this additional precision is helpful or harmful for people using risk calculators. In other words, we *can* be more precise, but *should* we?

The importance of this question becomes apparent when one considers the complex range of challenges inherent in risk communication. Across levels of education and expertise, many people, particularly those with poor numeracy, have trouble interpreting numbers in health-risk communications [14,15] and demonstrate biased interpretations of proportions [16]. For example, people have been shown to rate a cancer as riskier when it "kills 1286 out of 10,000 people" (about 13%) than when it "kills 24.14 out of 100 people" (about 24%) [17], and to prefer a lottery that offers 5, 6, 7, 8, or 9 winning draws in 100 (5%–9%) over a lottery that offers 1 winning draw in 10 (10%) [18]. Furthermore, people respond differently to proportions and numbers with decimals depending on the presentation format—for example, percentages (25%), natural frequencies (25 in 100), and "1 in n" formats (1 in 4) [19,20].

The precision of a number, in particular, can affect how people perceive and act on numerical information. For example, home buyers offer bids closer to the asking price for houses with more precise list prices [21] and are more likely to choose a battery-powered product with a battery life of 120–180 minutes versus one with a battery life of 2–3 hours [22]. Consumers [23,24] and investors [25] exhibit purchasing behavior suggesting that many people may interpret prices with decimals by simply truncating digits beyond the decimal point.

The effects of estimate precision in health have thus far been studied by examining responses to point estimates (eg, 9%) versus ranges of estimates (eg, 5%–13%). Previous qualitative research suggested that ranges of risk estimates may be perceived as more credible than point estimates [26]. Subsequent experimental research showed that perceived risk was larger for a range than a point estimate when the estimate was expressed in text, but there was no main effect of precision on credibility [27]. The sparse, mixed results in previous research suggest that the effect of precision on believability of a risk estimate remains an open question. Further, we propose that, although ranges and number of significant figures are both used mathematically to convey precision or lack thereof, they present different questions when it comes to layperson responses. Ranges make the imprecision explicit, whereas the number of decimal places is an implicit signal. Many risk calculators give a single point estimate [11], and, prior to this experiment, the specific effects of the number of decimal places in a risk estimate had not been studied.

In this study, we aimed to isolate the effects of precision—that is, number of decimal places—on people's interpretations of risk estimates offered by online risk calculators. We selected believability ("How believable is this number?") and risk magnitude ("How large or small does this number feel to you?") as primary outcomes. Perceptions of believability and risk magnitude are critical to changing health attitudes and behavior [28], common goals of risk calculators, and we suggest that these are a reasonable first line of consideration when examining responses to a risk estimate. People are unlikely to ponder whether a risk is modifiable and what actions they might take if they lack confidence in the estimate in the first place, and a key measure for risk communication is the resulting subjective feeling of how large or small the risk is.

Methods

Design of Experiment

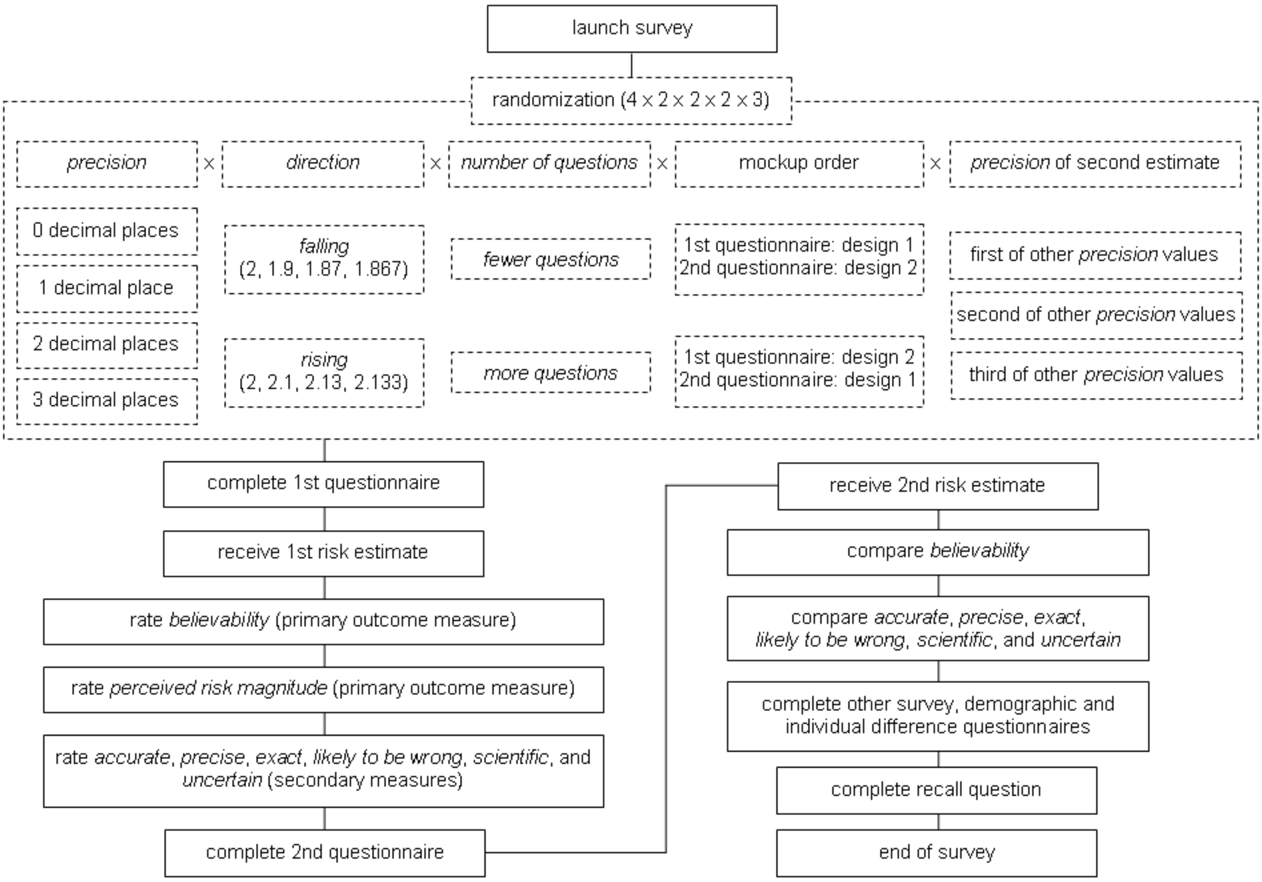
Participants were asked to imagine they were visiting a kidney cancer risk calculator. (See Multimedia Appendix 2 for full methodological details, including exact wording used, screenshots of the two mock risk calculators participants were sent to, and rationale for design choices described here.) The calculator questions used actual risk factors for kidney cancer, but randomly assigned each participant a risk estimate around 2%, slightly above the average lifetime risk statistic for US adults of 1.49% [29]. Risk estimates were expressed as integers with zero decimal places (2%), to one decimal place (1.9% or 2.1%), two decimal places (1.87% or 2.13%), or three decimal places (1.867% or 2.133%). As shown by the numbers used, numerical values either fell or rose slightly below or above 2% with increasing decimals; participants were randomly assigned to either the “falling” or “rising” group. Participants were also randomly assigned to either a shorter version (fewer questions) or longer version (more questions) of the mock risk calculator to test whether a longer questionnaire might reasonably be seen as providing a more credible estimate.

After completing the questions in the risk calculator, participants were shown the “result” that they had been randomly assigned. They were then asked to indicate the believability of the risk, how large or small it felt to them, and a series of secondary assessments about how well or poorly the following adjectives described the estimate they were given: accurate, precise, exact, likely to be wrong, scientific, and uncertain. These secondary assessments were taken from previous work done by our research group comparing point estimates and ranges [30], and were intended to collect exploratory data that might help unpack any differences found in primary outcomes.

To mimic a plausible response to receiving a risk estimate—namely, seeking a second estimate to confirm or contradict the first—participants were then directed to a second mock calculator that presented a second risk estimate. The second estimate was randomly assigned from the other three numbers in their rising or falling group of numbers. For example, participants assigned to the falling group might receive a first risk estimate of 1.9%, and their second risk estimate would be randomly assigned as either 2%, 1.87%, or 1.867%. Participants were then asked to compare the two numbers in terms of believability, as well as the secondary outcomes of accurate, precise, exact, likely to be wrong, scientific, and uncertain. To remove the possibility that recall differences might contaminate the comparisons, all comparison questions were presented with the estimates as labels, with the first estimate as the label for the first column, and the second estimate for the second column. We did not ask participants to compare the estimates in terms of risk magnitude because we predicted that the difference in expressed values (for example, 2 > 1.9) would dominate any effects of the level of precision and we therefore saw little benefit in increasing respondent burden by adding another comparison task.

Participants completed another survey about hypothetical treatment choices for colon cancer, in which participants were cross-randomized to avoid any systematic interaction between the two surveys. They then completed a brief set of demographic and individual difference measures. Finally, on the last page of the combined survey, we asked participants to recall to the best of their ability both risk estimates they had been given. (The experimental procedure is detailed in Figure 1 and Multimedia Appendix 3.)

Figure 1. Flow Diagram of Experiment.



Recruitment

Email invitations were sent to a random sample of US adults aged 30 to 70 years, selected from a panel of Internet users administered by Survey Sampling International (Shelton, CT, USA) and stratified by gender, age, and race to ensure demographic diversity. The survey did not collect identifying information. Survey Sampling International uses a complex digital fingerprinting technique to ensure respondent uniqueness [31]. The design was approved by the University of Michigan Medical Institutional Review Board.

Measures

Independent Variables

Precision was operationalized by the number of decimals shown in the risk estimate (ie, 0, 1, 2, or 3).

The *direction* of the risk estimate refers to whether the numeric value of the estimates rose or fell with increasing precision. The *rising* condition used the numbers 2, 2.1, 2.13, and 2.133. The *falling* condition used 2, 1.9, 1.87, and 1.867.

The *number of questions* in the risk calculator was operationalized as either 4 pages (screens) of questions in the mocked-up questionnaire or the same 4 plus an additional 3 pages. The former, which contained 11 questions in all, represents the condition *fewer questions* and the latter, which contained 17 questions in all, represents the condition *more questions*.

Primary Outcome Variables

Believability of the risk estimate was assessed by asking participants to respond to the question, “Think about the risk number you were given. In your opinion, how **believable** is this number?” (emphasis original) on a 6-point Likert scale anchored by “not at all” on one end and “extremely” on the other. Responses were not labeled with their numeric value, meaning that participants did not see any numbers, only a horizontal visual array of equally spaced radio buttons. Believability was collected immediately after viewing the risk estimate, on its own survey page. For examining and reporting distribution of responses, we defined the bottom two points of the scale as representing *low believability*, the middle two points as representing *moderate believability*, and the top two points as representing *high believability*.

Perceived risk magnitude was a measure of how big the risk estimate felt to participants. This outcome was assessed by the question, “How **large or small** does the risk **feel** to you?” (emphasis original). Responses were collected using a horizontal slider labeled by “extremely small” on the left and “extremely large” on the right. The slider marker was visible throughout the interaction, with its default position in the center. We anticipated that, while this might anchor results around the center of the scale, such anchoring would not interfere with our analysis, which aimed only to compare slider positions between conditions. Therefore, the increased usability of making the slider marker visible was worth the potential global bias. As with the believability scale, participants saw only the visual position of the slider, not the numeric values representing their response. Values were stored to 2 decimal places on the interval [0, 1]. Risk magnitude was collected on its own survey page.

Secondary Outcome Variables

Secondary outcomes *accurate*, *precise*, *exact*, *likely to be wrong*, *scientific*, and *uncertain* were assessed via the questions, “We would also like to know, in your opinion: How **accurate** is this number? How **precise** is this number? How **exact** is this number? How **scientific** is this number? How **likely** is this number **to be wrong**? How **uncertain** is this number?” (emphases original). All were collected using 6-point Likert scales anchored by “not at all” on one end and “extremely” on the other, with no numeric feedback. These variables were collected together on one survey page, in a randomly ordered list.

To elicit *comparisons* between the two given risk estimates, participants were first asked to compare the two numbers in terms of believability, indicating whether the first estimate they were given or the second was more believable, or if they were equal in this respect. This comparison question was completed on its own survey page, on which the response options were labeled with the first risk estimate, then the second, then the label “both equal.” On the next page, participants were given a similar comparison task, this time applied to the same secondary outcomes used earlier: accurate, precise, exact, likely to be wrong, scientific, and uncertain, once again randomly ordered in a list.

Measures of *recall* were elicited by asking participants, “To the best of your ability, can you tell us the lifetime risk of kidney cancer from the website calculators earlier in this survey? Please enter the most detailed numbers you can remember.” We defined correct *exact recall* as a recalled number that matched the given risk estimate perfectly. We defined correct *approximate recall* as a recalled estimate that was within 50% error of the given number (Figure 2).

Figure 2. Equation for calculating participants’ approximate recall (within 50% error) of their estimated risk: *recall* represents the recalled number and *orig* represents the original given estimate.

$$\textit{approximate recall} = \begin{cases} 1\ (\text{correct}) & \text{if } \left|\dfrac{recall - orig}{orig}\right| \leq 0.5 \\ 0\ (\text{incorrect}) & \text{if } \left|\dfrac{recall - orig}{orig}\right| > 0.5 \end{cases}$$

This definition enables a wide margin of error, which we deemed appropriate for such a small risk estimate. Thus, recall estimates between approximately 1% and 3% were defined as correct approximate recall, whereas those outside the defined range were defined as incorrect.

Individual Difference Measures

Individual difference measures used in this study were as follows: *subjective numeracy* [32], *cancer fear* (using an adapted breast cancer fear scale [33] that has been adapted as general cancer fear in other research [34]), and three Big-5 personality dimensions, *neuroticism*, *openness*, and *agreeableness*, which were collected using a brief index of all five dimensions [35]. We also posed an ad hoc self-assessed measure of *susceptibility to marketing*: “How much do you think your choices about the things you buy are influenced by marketing?" assessed on a 6-point Likert scale anchored by "not at all influenced" and "extremely influenced." All of these measures were selected for analysis due to their possible moderating effect on individuals' responses to different levels of precision.

Statistical Analyses

We analyzed ratings of believability and risk magnitude for the first risk estimates via multivariate analysis of variance (MANOVA). We used three independent variables: precision (number of decimals), direction of values (rising versus falling), and number of questions (4 screens versus 7). We included all main effects and all 2-way interactions in the model and conducted post hoc tests on precision (the only independent variable with more than two levels) via the Tukey least significant difference test. We performed a second MANOVA with the same model to examine the effects of the independent variables on the secondary outcomes accurate, precise, exact, likely to be wrong, scientific, and uncertain.

To explore participants' assessments of the comparisons between the two risk estimates, observed differences in proportions for each measure were tested via 2-tailed binomial tests. These tests were conducted only on data from participants who judged the two numbers as different on that measure.

Finally, we analyzed exact and approximate recall via repeated measures logistic regression, regressing recall on the number of decimals. We present both exact and approximate recall results, though we focus on approximate recall as the fair and practically relevant comparison. Recalling a value expressed to more decimal places exactly requires additional memory capacity, and it is unlikely that people would need to recall estimates to a high level of precision for any practical purpose.

Data were entered and analyzed in SPSS version 16.0 (IBM Corporation, Somers, NY, USA).

Results

Recruitment

Out of 4242 people who clicked the link to launch the survey, 4117 (97%) continued beyond the informed consent page, and 3422 (81%) completed the survey. All completed surveys were analyzed. Completion rates were consistent across experimental conditions. Characteristics of study participants are shown in Table 1. Participants' ages ranged from 30 to 70 years old with mean age 50 (SD 11) years, 1723/3305 (52%) were female, participants were racially and ethnically diverse, and 1483/3402 participants (44%) had an associate's degree or higher.

Table 1. Study participant characteristics (N = 3422)

Characteristic	
Age (years), mean (SD)	50 (11)
Gender, n (%)	
Female	1723 (52%)
Male	1582 (48%)
Ethnicity, n (%)	
Hispanic	486 (14%)
Middle Eastern	46 (1%)
Race, n (%)	
White or Caucasian	2518 (74%)
Black or African American	529 (16%)
American Indian or Alaska Native	55 (2%)
Asian or Asian American	150 (4%)
Pacific Islander or Native Hawaiian	17 (0.5%)
Other	167 (5%)
Highest education level reached, n (%)	
None	2 (0.1%)
Elementary school	4 (0.1%)
Some high school, but no diploma	72 (2%)
High school (diploma or GED[a])	665 (19%)
Trade school	186 (6%)
Some college, but no degree	990 (29%)
Associate’s degree (AA, AS, etc)	357 (11%)
Bachelor’s degree (BS, BA, etc)	759 (22%)
Master’s degree (MA, MPH, etc)	306 (9%)
Doctoral/professional degree (PhD, MD, etc)	61 (2%)

[a] General equivalency diploma.

Statistical Analyses

The precision of the risk estimate was related to believability and perceived risk magnitude. In particular, risk estimates with zero decimals yielded the highest believability scores, with scores decreasing slightly with increasing number of decimal places ($F_{3,3384}$ = 2.94, *P* = .03, partial eta squared = .003). Believability was not significantly related to the number of questions in the risk calculator. Participants whose estimates had more decimals felt that the risk magnitude was larger ($F_{3,3384}$ = 4.70, *P* = .003, partial eta squared = .004), as did participants who were assigned to a risk calculator with fewer questions ($F_{3,3384}$ = 5.85, *P* = .02, partial eta squared = .002). For both believability and risk magnitude, direction of values did not have a significant effect, and there were no significant interactions. See Table 2 for further details.

The distribution of believability ratings is shown in Table 3. Risk estimates with zero decimals yielded 7%–10% more participants who rated the estimate as highly believable (χ^2_3 = 17.8, *P* < .001).

Table 2. Primary outcomes

	Believability: 1 = not at all, 6 = extremely; mean (SD)[a]	Risk magnitude: 0 = extremely small, 1 = extremely large; mean (SD)[a]
Precision (number of decimal places)		
0	4.35 (1.24) (reference)	.21 (.24) (reference)
1	4.24 (1.23) (*P* = .07)	.24 (.24) (*P* = .03)
2	4.21 (1.26) (*P* = .02)	.23 (.24) (*P* = .20)
3	4.19 (1.22) (*P* = .006)	.26 (.25) (*P* < .001)
Overall significance	*P* = .03	*P* = .003
Direction of values (rising: 2, 2.1, 2.13, 2.133 ; falling: 2, 1.9, 1.87, 1.867)		
Rising	4.24 (1.24)	.24 (.24)
Falling	4.26 (1.24)	.23 (.24)
Overall significance	*P* = .59	*P* = .18
Number of questions (fewer: 4 screens of questions ; more: 7 screens of questions)		
Fewer	4.21 (1.22)	.25 (.25)
More	4.28 (1.26)	.22 (.24)
Overall significance	*P* = .10	*P* = .02

[a] *P* values reported next to means for precision refer to Tukey least significant difference referenced against zero decimals.

Table 3. Distribution (n, %) of believability responses by precision (also see Table 4.1 in Multimedia Appendix 4)

	Low believability	Moderate believability	High believability
Precision (number of decimal places)			
0	63 (7%)	353 (41%)	450 (52%)
1	60 (8%)	378 (47%)	362 (45%)
2	80 (9%)	389 (46%)	383 (45%)
3	69 (8%)	440 (50%)	373 (42%)

Estimates with one decimal point were rated as the least uncertain compared to estimates with zero ($P = .02$), two ($P = .001$), or three ($P = .001$) decimal places ($F_{3,3314} = 4.76$, $P = .003$, partial eta squared = .004.) However, none of the related terms accurate, precise, exact, likely to be wrong, or scientific showed a significant difference between different numbers of decimals, which may suggest possible confusion about the meaning(s) of uncertainty, which has multiple meanings, each of which ought to line up conceptually with at least one of the other terms. If participants interpreted uncertainty as an assessment of the truth of the estimate, we would also expect differences in ratings of accuracy and likelihood of being wrong. If, on the other hand, uncertainty were to be interpreted as an assessment of precision, we would expect ratings of precision and exactitude to show differences. None of these differences were in evidence and, thus, the secondary measures did not suggest potential mechanisms to explain this finding.

Ratings of accuracy were higher in the condition with more questions ($F_{1,3314} = 4.16$, $P = .04$, partial eta squared = .001), but number of questions did not have a significant effect on any other terms. Direction of values did not have a significant effect, and there were no significant interactions.

Overall, none of the secondary measures suggested potential mechanisms to explain the primary findings.

Individual difference measures demonstrated strong main effects in the expected directions. (See Multimedia Appendix 4 for details.) However, none of the individual difference measures showed any statistically significant interactions with precision, suggesting that the effects of precision are consistent across types of users.

When comparing the first and second risk estimates they were given, large majorities of participants indicated equality across all measures (see Table 4 for a summary; see Multimedia Appendix 4 for tables detailing comparisons across combinations of number of decimals.) The minority of participants who thought the two numbers were different on one or more measures judged estimates with fewer decimals as more believable, but also less accurate, less precise, less exact, less scientific, and more uncertain.

Table 4. Comparisons of two risk estimates

Which number is more	Percentage of participants who chose Number with fewer decimals	Both numbers equal	Number with more decimals	Significance of observed proportion of participants choosing fewer vs more decimals
Believable[a]?	11%	80%	9%	$P = .006$
Accurate[b]?	13%	70%	17%	$P < .001$
Precise[b]?	13%	62%	25%	$P < .001$
Exact[b]?	13%	63%	24%	$P < .001$
Scientific[b]?	11%	69%	20%	$P < .001$
Likely to be wrong[b]?	13%	74%	14%	$P = .28$
Uncertain[b]?	15%	72%	13%	$P = .002$

[a] Primary comparison outcome, question presented first on its own survey page.

[b] Secondary comparison outcomes, questions presented together on one page in random order.

After completing the questions comparing the two risk estimates, participants spent a median of 9.6 minutes (interquartile range 6.5 minutes) answering an unrelated survey before reaching the recall task, in which they were asked to recall both risk estimates they had been given earlier. Participants were not warned that they would be asked to recall the numbers.

The proportions of participants with correct recall are shown in Table 5 for both exact and approximate recall error. Errors in exact recall increased quickly with precision. The majority of participants had correct approximate recall across all levels of precision, but errors remained more common in estimates with decimal places than in those with zero decimals. The effects of precision were significant on both exact and approximate recall.

Table 5. Participants with correct recall

Precision (number of decimals)	Exact recall Correct	Odds ratio (95% CI)	Approximate recall Correct	Odds ratio (95% CI)
0	93%	Reference	96%	Reference
1	83%	0.36 (0.29–0.44)	94%	0.65 (0.49–0.86)
2	70%	0.17 (0.14–0.21)	95%	0.70 (0.53–0.94)
3	43%	0.06 (0.05–0.07)	94%	0.61 (0.45–0.81)
	Wald $\chi^2_3 = 1014$, $P < .001$		Wald $\chi^2_3 = 12.1$, $P = .007$	

Discussion

Principal Results

This study suggests that risk calculators that produce risk estimates with different levels of precision can result in different perceptions of those estimates in terms of believability and risk magnitude, as well as differences in recall. In this experiment, risk estimates with zero decimals were judged as the most believable. People may find integers somewhat more believable than numbers with decimals simply because integers are easier to understand. As evidenced by, for example, jokes and confusion about an average American family having 2.2 children, it is challenging for people to map population-based statistics onto individual circumstances. Indeed, many people, even those who are well educated, have trouble with probabilities and percentages [36,37]. This is particularly true of people with poor numeracy skills [38]. The fact that people have trouble with this concept is perhaps not surprising given that a patient doesn't experience the probability of an event occurring; she or he either experiences the event or not [39]. Therefore, truly understanding a risk estimate requires conceptually mapping a probability onto a binary outcome. Adding decimals to the risk estimate only makes this task more challenging. Simplifying the risk estimate might therefore increase understanding and, hence, believability.

Risk estimates with the least precision (zero decimals) also felt smaller on average than estimates with greater precision. This finding parallels previous research on ratio bias, in which statistical frequencies presented using smaller denominators felt smaller than those that used larger denominators [15,16]. We speculate that seeing fewer numbers evokes in people a smaller degree of number sense and hence lower risk magnitudes.

Lower precision was also associated with better recall of the given risk estimates. It is not particularly surprising that people found numbers with more decimals places more difficult to remember perfectly. Recalling four digits takes considerably more cognitive capacity than recalling one. It is more notable that, even when allowing for a very generous margin of error in a recall task that took place shortly after the estimate was provided, there were statistically significant differences in approximate recall between estimates with zero decimals and all three estimates with decimals. This means that using decimals in a risk estimate not only reduces the chances that users will be able to recall the number exactly, but also reduces the likelihood that they will be able to remember it even approximately. This may be partially attributable to a lack of understanding about the meaning of decimals, because if people are unable to comprehend the data that they have been given, they will not be able to turn it into information that can later be recalled.

This study also suggests that the number of questions asked in a risk calculator may have an effect on perceived risk magnitude. People who completed a longer questionnaire judged the risk estimate as smaller. Although our study found no statistical effect of the number of questions in the calculator on believability, this may have been because even our version with fewer pages of questions was sufficient to be over a threshold of believability. Further research will be required to explore the effects of very brief questionnaires on people’s assessments of risk calculators, but it is worth noting that, even with a very detailed questionnaire, the estimate with zero decimals still garnered the highest believability scores.

Limitations

There are three main limitations to this study. First, this experiment was based on a hypothetical scenario with artificial risk estimates all around 2%. We do not know whether similar effects would be found in situations in which numbers are real and individualized for the user, people are self-motivated to seek out the risk information, and/or numbers are larger or smaller. However, our mock risk calculator was modeled after real-world examples, and thus we have no reason to believe that patient behavior would differ when using an actual risk calculator to which he or she was directed, for example, in a routine monthly email from his or her health care group or system. We acknowledge that it is more difficult to predict how people might respond in a similar situation in which they are deliberately seeking out the information. However, conducting a controlled experiment in which the only variation was random assignment of the number of decimals in the risk estimate allowed us to control for many of the complexities of how people decide whether a piece of online health information is trustworthy, thereby isolating the unique effects of the precision of the risk estimate. Findings regarding real-world use of risk calculators will depend to some extent on users’ prior expectations regarding their risk; for example, people may be resistant to accepting risk estimates that are higher than their prior expectations [40]. We did not assess prior expectations in this study because we wished to avoid biasing responses to our questionnaire [41], so we cannot speculate on how relationships between prior expectations and assigned estimates might have influenced participants’ responses. We believe our findings will remain applicable to real-world risk calculators that display risks of varying magnitudes, but confirming this belief requires further research.

Second, all of the statistically significant findings in this study have small effect sizes. Single-digit *F* statistics, small odds ratios, and modest findings in post hoc tests suggest only small differences in the way people interpret and recall risk estimates with varying levels of decimal precision. However, online risk calculators aim to reach large numbers of people amid a cacophony of conflicting and confusing health information. Therefore, as with other challenges in the complex field of health communication, small effects may be worth attention, especially when they can be brought about by design modifications as easily implemented as rounding risk estimates to the nearest integer. If developers of a risk calculator would like users to judge received estimates as highly believable, this research suggests that expressing results as integers may help with this goal. The simplicity of this design change implies a high benefit-cost ratio.

Third, although our study included some secondary outcomes selected in the hopes that these might help unpack any differences found in the primary outcomes, effects on the secondary measures were largely absent. This may be partly attributable to the small effect sizes on the primary outcomes—it can be harder to unpack a small box. Nonetheless, it would be useful to better understand the mechanisms behind any differences in how people perceive risk estimates expressed as integers versus those with decimal places. Further research will be required to achieve this understanding.

Comparison With Prior Work

To our knowledge, there has been no prior work examining the effect of decimal precision in risk estimates.

However, our finding that increased precision leads to lower believability is in line with previous qualitative research suggesting that ranges of risk estimates may be seen as more credible than point estimates [26]. On the other hand, our finding is in contrast with experimental work that followed this qualitative study, which found no main effect on credibility [27]. This lack of agreement may be due to the inherent difference between the way people interpret ranges versus point estimates and the way they interpret numbers of decimal places. It may also be due to measurement differences. In the previous study, credibility was defined by an ad hoc scale that combined two items—ratings of trust in the results of the computer program and perceptions of accuracy of the risk estimate—whereas in our study, we assessed believability by asking participants to rate the believability of the estimate directly.

Our finding that increased precision also leads to perception of lower risk magnitude is in contrast with previous research in which more ambiguous risk estimates, meaning those expressed as ranges rather than point estimates, led to increased risk perceptions [26,27]. We believe that this difference is another example of differences in interpretation between different ways of expressing precision—that is, decimals places, or ranges and point estimates. Both findings (those in previous work and ours) support our speculation that seeing fewer numbers may elicit a lower overall sense of magnitude.

Research in consumer pricing suggests that prices with decimal places may be interpreted by simply ignoring numbers after the decimal place. If this were to also occur in health risks, we would expect to observe an interaction between precision and direction on perceived risk magnitude. That is, we would expect risk magnitude scores for the rising condition (2, 2.1, 2.13, and 2.133) to remain consistent regardless of precision, while those for the falling condition (2, 1.9, 1.87, and 1.867) would decrease between the first estimate and the other 3. We did not observe any such interaction, and we suggest that this is likely because, even if this effect exists in the health context, it may be significantly smaller and thus not detectable in this study. In other words, a price of $1 may feel different from $2, but a 1% risk may not feel meaningfully smaller than a 2% risk.

Our finding that fewer decimal places leads to better recall is consistent with research reporting that health communications that provide less detail lead to higher comprehension than those that provide more detail [42,43].

Conclusions

There are subtle but significant differences in how people interpret risk estimates of varying precision. Increasing precision in the form of decimal places shows no clear benefit and suggests small but significant costs. Results from our experiment suggest that, in general, rounding to the nearest integer is preferable for communicating small risk estimates so that they may be judged as believable and remembered correctly. Given these findings, we recommend that risk calculator designers structure their algorithms to express risk in integers, though expressions to 1 decimal place may also be acceptable in situations when user recall of the number is not an important consideration or when greater precision is necessary to show differences between two or more numbers.

Acknowledgments

Financial support for this study was provided by a grant from the U. S. National Institutes for Health (R01 CA087595). Dr. Zikmund-Fisher is supported by a Mentored Research Scholar Grant from the American Cancer Society (MRSG-06-130-01-CPPB). The funding agreements ensured the authors' independence in designing the study, interpreting the data, and publishing the report.

Three photos were included in the mock websites shown in Figures 1 and 2 in Multimedia Appendix 2. Two photos – the first of some chemistry lab equipment, the second of an ivy-covered building– were selected from the photostreams of Flickr users Horia Varlan and Mark Sardella, respectively, and were used in this study under Creative Commons licenses. The third photo, of a mother and daughter at a graduation ceremony, is a family photo from one of the researchers and was used with permission of the researcher and family members.

The authors gratefully acknowledge Peter Ubel's comments on the early experimental design of this study.

Conflicts of Interest

None declared

Multimedia Appendix 1

Breast Cancer Risk Calculator Example.

[PDF file (Adobe PDF File), 39 KB - jmir_v13i3e54_app1.pdf]

Multimedia Appendix 2

Detailed Methods.

[PDF file (Adobe PDF File), 600 KB - jmir_v13i3e54_app2.pdf]

Multimedia Appendix 3

Detailed Flow Diagram.

[PNG file (PNG Image), 28 KB - jmir_v13i3e54_app3.png]

Multimedia Appendix 4

Additional Details of Results.

[PDF file (Adobe PDF File), 45 KB - jmir_v13i3e54_app4.pdf]

References

1. AIDSGAME.com. 2006. Free HIV test URL: http://www.aidsgame.com/quick_hiv_test.aspx [accessed 2010-08-27] [WebCite Cache ID 5sJFwZRw5]
2. National Cancer Institute. US National Institutes of Health. 2008. Breast Cancer Risk Assessment Tool URL: http://www.cancer.gov/bcrisktool/Default.aspx [accessed 2010-08-27] [WebCite Cache ID 5sJI5gxyx]
3. Mayo Clinic staff. MayoClinic.com. 2010. Heart Disease Risk Calculator URL: http://www.mayoclinic.com/health/heart-disease-risk/HB00047 [accessed 2010-08-27] [WebCite Cache ID 5sJI9IzJ4]
4. National Cholesterol Education Program. NCEP. Risk Assessment Tool for Estimating Your 10-Year Risk of Having a Heart Attack URL: http://hp2010.nhlbihin.net/atpiii/calculator.asp [accessed 2011-01-11] [WebCite Cache ID 5verTeAnJ]
5. American Diabetes Association. ADA. 2009. My Health Advisor URL: http://www.diabetes.org/diabetes-basics/prevention/my-health-advisor/ [accessed 2011-01-11] [WebCite Cache ID 5verMk5A1]
6. American Diabetes Association. ADA. 2010. Diabetes Risk Test URL: http://www.diabetes.org/diabetes-basics/prevention/diabetes-risk-test/ [accessed 2010-08-27] [WebCite Cache ID 5sJIBdASq]
7. Prostate Cancer Canada. PCATs to Predict the Odds a Biopsy Will Find Prostate Cancer. no date URL: http://www.prostatecancer.ca/Tools/local/files/english/calculator/calculator.html [accessed 2010-08-27] [WebCite Cache ID 5sJIYQstv]
8. Memorial Sloan-Kettering Cancer Center. 2011. Prostate Cancer Nomograms: Post-Radical Prostatectomy URL: http://nomograms.mskcc.org/prostate/PostRadicalProstatectomy.aspx [accessed 2011-07-26] [WebCite Cache ID 60THUhRZP]
9. Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine. 2007. Your Disease Risk URL: http://www.yourdiseaserisk.wustl.edu/ [accessed 2010-08-28] [WebCite Cache ID 5sKKj9DAr]
10. Holmberg C, Harttig U, Schulze MB, Boeing H. The potential of the Internet for health communication: the use of an interactive on-line tool for diabetes risk prediction. Patient Educ Couns 2011 Apr;83(1):106-112. [doi: 10.1016/j.pec.2010.04.021] [Medline: 20547029]
11. Waters EA, Sullivan HW, Nelson W, Hesse BW. What is my cancer risk? How internet-based cancer risk assessment tools communicate individualized risk estimates to the public: content analysis. J Med Internet Res 2009;11(3):e33 [FREE Full text] [doi: 10.2196/jmir.1222] [Medline: 19674958]
12. Halls SB. Detailed Breast Cancer Risk Calculator. 2008. URL: http://www.halls.md/breast/risk.htm [accessed 2010-09-20] [WebCite Cache ID 5ssU7pe8d]
13. GenneX Healthcare Technologies, Inc. 1999. Breast Cancer Risk URL: http://www.estronaut.com/a/breastInteractive2.htm [accessed 2010-09-20] [WebCite Cache ID 60TF7An03]
14. Nelson W, Reyna VF, Fagerlin A, Lipkus I, Peters E. Clinical implications of numeracy: theory and practice. Ann Behav Med 2008 Jun;35(3):261-274. [doi: 10.1007/s12160-008-9037-8] [Medline: 18677452]
15. Woloshin S, Schwartz LM, Byram S, Fischhoff B, Welch HG. A new scale for assessing perceptions of chance: a validation study. Med Decis Making 2000;20(3):298-307. [Medline: 10929852]
16. Weinstein ND. What does it mean to understand a risk? Evaluating risk comprehension. J Natl Cancer Inst Monogr 1999(25):15-20. [Medline: 10854451]
17. Yamagishi K. When a 12.86% mortality is more dangerous than 24.14%: implications for risk communication. Appl Cogn Psychol 1997;11(6):495-506. [doi: 10.1002/(SICI)1099-0720(199712)11:6<495::AID-ACP481>3.0.CO;2-J</495::AID-ACP481>]
18. Denes-Raj V, Epstein S. Conflict between intuitive and rational processing: when people behave against their better judgment. J Pers Soc Psychol 1994 May;66(5):819-829. [Medline: 8014830]
19. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Making 2007;27(5):696-713. [doi: 10.1177/0272989X07307271] [Medline: 17873259]
20. Cuite CL, Weinstein ND, Emmons K, Colditz G. A test of numeric formats for communicating risk probabilities. Med Decis Making 2008;28(3):377-384. [doi: 10.1177/0272989X08315246] [Medline: 18480036]
21. Janiszewski C, Uy D. Precision of the anchor influences the amount of adjustment. Psychol Sci 2008 Feb;19(2):121-127. [doi: 10.1111/j.1467-9280.2008.02057.x] [Medline: 18271859]
22. Zhang CYZ, Schwarz N. The granularity effect. In: Proceedings. 2011 Presented at: The Society for Consumer Psychology 2011 Winter Conference; Feb 24-26, 2011; Atlanta, GA, USA.
23. Schindler RM, Kirby PN. Patterns of rightmost digits used in advertised prices: implications for nine-ending effects. Journal of Consumer Research 1997;24(2):192-201. [doi: 10.1086/209504]
24. Stiving M. Price-endings when prices signal quality. Management Science 2000;46(12):1617-1629. [doi: 10.1287/mnsc.46.12.1617.12078]
25. Sonnemans J. Price clustering and natural resistance points in the Dutch stock market: a natural experiment. European Economic Review 2006;50(8):1937-1950. [doi: 10.1016/j.euroecorev.2005.09.001]
26. Han PK, Klein WM, Lehman TC, Massett H, Lee SC, Freedman AN. Laypersons' responses to the communication of uncertainty regarding cancer risk estimates. Med Decis Making 2009;29(3):391-403. [doi: 10.1177/0272989X08327396] [Medline: 19470720]
27. Han PK, Klein WM, Lehman T, Killam B, Massett H, Freedman AN. Communication of uncertainty regarding individualized cancer risk estimates: effects and influential factors. Med Decis Making 2011;31(2):354-366. [doi: 10.1177/0272989X10371830] [Medline: 20671211]

28. Weinstein ND. Perceived probability, perceived severity, and health-protective behavior. Health Psychol 2000 Jan;19(1):65-74. [Medline: 10711589]
29. Altekruse SF, Kosary CL, Krapcho M, Neyman N, Aminou R, Waldron W, et al, editors. SEER Cancer Statistics Review 1975-2007. Bethesda, MD: National Cancer Institute; 2010. Table 1.14, Lifetime Risk URL: http://seer.cancer.gov/csr/1975_2007/ [accessed 2010-09-20] [WebCite Cache ID 60TFi28C3]
30. Fuhrel-Forbis A, Zikmund-Fisher BJ, Exe N, Kahn V, Witteman H, Jankovic A, et al. Imprecise definitions of "precise": preferences for precision depend on context. 2011 Presented at: Annual Meeting of the Society of Medical Decision Making; October 24-27, 2010; Toronto, ON, Canada. [doi: 10.1177/0272989X10393877]
31. SSI Survey Sampling International. SSI. 2010. SSI Verify URL: http://www.surveysampling.com/en/ssi-verify [accessed 2010-11-11] [WebCite Cache ID 5uAWKnAYa]
32. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
33. Champion VL, Skinner CS, Menon U, Rawl S, Giesler RB, Monahan P, et al. A breast cancer fear scale: psychometric development. J Health Psychol 2004 Nov;9(6):753-762. [doi: 10.1177/1359105304045383] [Medline: 15367754]
34. Miles A, Voorwinden S, Mathews A, Hoppitt LC, Wardle J. Cancer fear and the interpretation of ambiguous information related to cancer. Cognition & Emotion 2009;23(4):701-713. [doi: 10.1080/02699930802091116]
35. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. Journal of Research in Personality 2003;37:504-528. [doi: 10.1016/S0092-6566(03)00046-1]
36. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001;21(1):37-44. [Medline: 11206945]
37. Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Communicating statistical information. Science 2000 Dec 22;290(5500):2261-2262. [Medline: 11188724]
38. Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. Numeracy and decision making. Psychol Sci 2006 May;17(5):407-413. [doi: 10.1111/j.1467-9280.2006.01720.x] [Medline: 16683928]
39. Walker VR. Direct Inference, Probability, and a Conceptual Gulf in Risk Communication. Risk Analysis 1995;15:603-609. [doi: 10.1111/j.1539-6924.1995.tb00756.x]
40. Harle CA, Downs JS, Padman R. A clustering approach to segmenting users of internet-based risk calculators. Methods Inf Med 2011 May 13;50(3):244-252. [doi: 10.3414/ME09-01-0080] [Medline: 20300681]
41. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. How making a risk estimate can change the feel of that risk: shifting attitudes toward breast cancer risk in a general public survey. Patient Educ Couns 2005 Jun;57(3):294-299. [doi: 10.1016/j.pec.2004.08.007] [Medline: 15893211]
42. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. A demonstration of "less can be more" in risk graphics. Med Decis Making 2010;30(6):661-671. [doi: 10.1177/0272989X10364244] [Medline: 20375419]
43. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr;64(2):169-190. [doi: 10.1177/107755870640020301] [Medline: 17406019]

Abbreviations

MANOVA: multivariate analysis of variance
OR: odds ratio

Edited by G Eysenbach; submitted 29.09.10; peer-reviewed by C Harle, P Wicks; comments to author 28.10.10; revised version received 16.05.11; accepted 17.05.11; published 07.09.11.

Please cite as:
Witteman HO, Zikmund-Fisher BJ, Waters EA, Gavaruzzi T, Fagerlin A
Risk Estimates From an Online Risk Calculator Are More Believable and Recalled Better When Expressed as Integers
J Med Internet Res 2011;13(3):e54
URL: http://www.jmir.org/2011/3/e54/
doi:10.2196/jmir.1656
PMID:21908265

Original Paper

An Online Resource of Digital Stories About Cancer Genetics: Qualitative Study of Patient Preferences and Information Needs

Rachel Iredale[1], BSocSc, MA, PhD; Lisa Mundy[1], BSC(Hons), PhD; Jennifer Hilgart[1], BSC(Hons)

[1] Institute of Medical Genetics, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
Rachel Iredale, BSocSc, MA, PhD
Institute of Medical Genetics
Cardiff University
Heath Park
Cardiff, CF14 4XN
United Kingdom
Phone: 44 (0)2920 748925
Fax: 44 (0)2920 743846
Email: Rachel.Iredale@wales.nhs.uk

Abstract

Background: The Cancer Genetics Service for Wales (CGSW) was established in 1998 as an all-Wales service for individuals with concerns about their family history of cancer. CGSW offers a range of services such as risk assessment, genetic counseling, and genetic testing. Individuals referred to cancer genetics services often have unmet information and support needs, and they value access to practical and experiential information from other patients and health professionals. As a result of the lifelong nature of genetic conditions, a fundamental challenge is to meet the ongoing needs of these patients by providing easily accessible and reliable information.

Objectives: Our aims were to explore how the long-term information and support needs of CGSW patients could be met and to assess whether an online bank of digital stories about cancer genetics would be acceptable to patients.

Methods: In 2009, CGSW organized patient panels across Wales. During these events, 169 patients were asked for their feedback about a potential online resource of digital stories from CGSW patients and staff. A total of 75 patients registered to take part in the project and 23 people from across Wales agreed to share their story. All participants took part in a follow-up interview.

Results: Patient preferences for an online collection of cancer genetics stories were collected at the patient panels. Key topics to be covered by the stories were identified, and this feedback informed the development of the website to ensure that patients' needs would be met. The 23 patient storytellers were aged between 28 and 75 years, and 19 were female. The digital stories reflect patients' experiences within CGSW and the implications of living with or at risk of cancer. Follow-up interviews with patient storytellers showed that they shared their experiences as a means of helping other patients and to increase understanding of the cancer genetics service. Digital stories were also collected from 12 members of staff working at CGSW. The digital stories provide reliable and easily accessible information about cancer genetics and are hosted on the *StoryBank* website (www.cancergeneticsstorybank.co.uk).

Conclusions: The Internet is one mechanism through which the long-term information and support needs of cancer genetics patients can be met. The *StoryBank* is one of the first places where patient and staff stories have been allied to every aspect of a patient pathway through a service and provides patients with an experiential perspective of the cancer genetics "journey." The *StoryBank* was developed in direct response to patient feedback and is an innovative example of patient involvement in service development. The stories are a useful resource for newly referred patients, current patients, the general public, and health care professionals.

(J Med Internet Res 2011;13(3):e78) doi:10.2196/jmir.1735

KEYWORDS

Consumer health information; familial cancer; Internet; narrative medicine; patient involvement.

Introduction

Health Information on the Internet

Internet use is growing rapidly throughout the world and individuals are increasingly turning to the Internet as a source of health information [1]. Indeed, health websites are some of the most viewed sites on the Internet [2]. In the United Kingdom, the National Health Service (NHS) Cancer Plan advises that the Internet should be promoted as a source of information for cancer patients [3]. It has been shown that people with cancer use the Internet for a wide range of information and support needs, throughout their illness [4,5]. For example, cancer patients use the Internet to interpret symptoms; to help understand consultations; to find out about tests and treatments; as a source of support; and to hear about the experiences of other patients [5-7]. Research suggests that many patients who use the Internet are satisfied with the information they receive and that their subsequent health care decisions are influenced by the information they gain [8,9].

Despite concerns about the accuracy of health information on the Internet [1,5,10], studies have shown that when searching the Internet, people assess the credibility of a website in several ways [10-12]. For example, participants in out study looked for the source of the website and prefer noncommercial sites, which are attached to recognized centers of excellence, such as universities or well-known health centers [12]. Overall, research suggests that the Internet is a useful resource for health information and individuals continue to use the Web to find this information. Research also shows that individuals want to hear from other people who have been through similar situations [13]. This is reflected in the increasing use of online communities and social networking sites for health purposes. Bender and colleagues identified a total of 1,090,397 Facebook users who were members of 620 breast cancer groups [14]. Cancer patients also benefit from having access to online stories from other patients, especially when the stories cover their specific information needs [15].

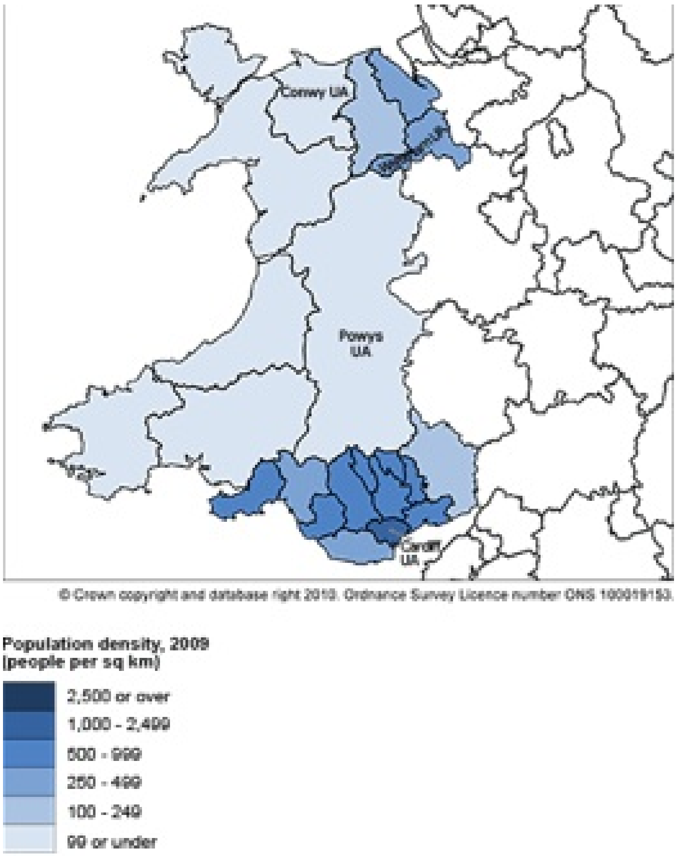
The Role of Stories

Health practitioners have traditionally used didactic approaches when giving information to patients. However, there is growing recognition that using stories to transmit health information may have many benefits. Narratives have been shown to improve the processing of information by capturing attention, enhancing understanding, and facilitating recall [16]. People with serious illnesses welcome the opportunity to hear from others who have had a similar experience [12,13]. One study showed that patients valued being able to access the experiences of other patients because it gave them reassurance and provided practical information that health professionals may not have perceived to be relevant [12]. Patients reported that hearing the experiences of others who had been through a similar situation would have reduced feelings of fear and isolation during their illness [12]. Stories have also been shown to be useful when presenting information to certain population groups [16,17]. For example, African American women from low-income neighborhoods who watched a video narrative from another patient describing a mammogram showed greater confidence with the procedure and had more conversations about breast cancer with family members than did women who watched a content-equivalent informational video [16].

The Cancer Genetics Service for Wales

The population of the United Kingdom is approximately 61 million, of whom 3 million live in Wales (area approximately 8000 square miles). A majority of the Welsh population reside in South Wales (see Figure 1) [18]. The Cancer Genetics Service for Wales (CGSW) was established in 1998 for individuals with concerns about their family history of cancer. It is an all-Wales service with three sites across the country in South East Wales, South West Wales, and North Wales. To date over 26,000 people have been referred to the service. To process referrals, CGSW uses a triage system [19]. First, patients are required to complete and return a detailed family history questionnaire. This information is then used to make a cancer genetic risk assessment, with an individual classified as having an average, moderate, or high risk of inheriting cancer. Patients found to be at moderate or high risk may be offered genetic counseling, screening, and if appropriate, genetic testing. Individuals categorized as being at average risk are advised that their risk is the same as that of the general population, and in most cases they are not offered a clinic appointment. Approximately 30% of all patients referred to CGSW are classified as being at high risk [20]. Unlike at many other health care services, once a person is referred to CGSW they can remain with the service for life because of the lifelong nature of genetic conditions. In the field of genetics, most information is based on probability and uncertainty and, as a result, living with the risk of inheriting cancer is a complex experience that can affect cognitive, emotional, and social functioning [21]. Several studies have shown that individuals undergoing genetic risk assessments show increased levels of psychological distress [22,23].

Figure 1. Population density map of Wales.



© Crown copyright and database right 2010. Ordnance Survey Licence number ONS 100019153.

Identifying the Need for the Cancer Genetics StoryBank

It should be recognized that the information needs of cancer genetics patients are diverse and can change over time [24-27]. Individuals referred to CGSW often have not been affected with cancer themselves but have a history of cancer in their families. For these people, unlike those who have a cancer diagnosis, their status is ambiguous because they maybe at increased risk of cancer but they do not have that illness now [28]. Furthermore, these individuals may have children who might also be at increased risk of hereditary cancer. The combination of these factors can make it difficult for families involved with cancer genetics services to make sense of their situation. It is imperative to meet patients' varying information needs because improved patient understanding may increase uptake of recommended screening and preventive measures, as well as help an individual cope with the risk assessment process, their risk categorization, and living with the risk of cancer [29,30].

Previous studies have found that the information needs of many patients living with a genetic risk of cancer are not being met [20,24,29] and that individuals undergoing genetic risk assessments have poor knowledge of genetics and the processes involved [30-33]. Many resources offer Web-based support to individuals with cancer, including online communities on UK cancer charity websites such as Macmillan and Breast Cancer Care. However, there are limited resources for individuals living at risk of cancer. One study examining the needs of patients referred to CGSW found that patients wanted improved access to information [20]. For example, patients sought general information about services, new developments in genetics, and details of other charities and organizations that could provide more information. Patients also wanted specific information, such as descriptions of a clinic consultation and the procedures involved in a DNA test. Patients suggested this information could be provided on a public website [20]. Another study exploring the information needs of women carrying a *BRCA1/2* gene mutation, which increases their risk of breast and ovarian cancer, found that participants felt that they and their families would benefit from an ongoing support network, which incorporated some Internet-based support [34]. Previous research with CGSW patients has shown that patients want up-to-date and accurate information about cancer genetics, to be reminded that they are never discharged from the service, and the opportunity to hear from other patients [20,35]. The *StoryBank* was therefore developed within CGSW in response to patient feedback, to address the demand for a Web-based resource for cancer genetics patients, which would meet their information and support needs on a long-term basis.

Involving Patients in Service Development

Patient involvement in health service development is imperative because in most Western countries there is increasing recognition that involving patients in health care is important for improving the quality of care provided [36,37]. In England and Wales health policies over the last 10 years have aimed to increase patient involvement in the NHS [38,39]. In relation to cancer, the Calman-Hine report and the NHS Cancer Plan have long emphasized that cancer services should be more patient centered and should be developed with consideration of patient views [3,40].

This paper describes how patients were involved in the development of the *StoryBank* [41], a website hosting a collection of digital stories from CGSW patients and staff about the cancer genetics service and associated clinical and psychosocial issues. Digital stories are the fusion of a narrative, in which the individual tells his or her story, with a series of images to illustrate the story. The aim of the *StoryBank* is to provide information and support to current and future patients, in a form that is easily accessible, reliable, engaging, and perhaps most important, allows patients to hear about the experiences of other patients. Patients have been involved at every stage of the project, from the initial conception to the launch of the website.

Methods

Patient Panels

The *StoryBank* has evolved as part of a 5-year program of work focusing on increasing patient involvement in service improvement and delivery at CGSW [35]. In June 2009, a newsletter was issued inviting CGSW patients to attend a patient panel. The newsletter was sent to 5906 patients, who had expressly asked to be kept informed about new initiatives. Of these patients, 28.46% (1681/5906 patients) returned the invitation, with 25.3% (425/1681 patients) expressing an interest in attending a patient panel. In total three patient panels were held across Wales during autumn 2009, with 169 participants attending (South East Wales n = 83; South West Wales n = 43; North Wales n = 43). The aims of the patient panels were for patients to decide which issues should be covered by the stories and how the Web-based resource should be developed. At the patient panels, patients were shown examples of digital stories. In groups of 8–10, patients were asked to discuss how digital stories could best meet CGSW patients' information and support needs and their preferences regarding how they would like to tell their stories. Each group was given a feedback sheet to record their comments. All the feedback sheets were collected at the end of each patient panel. We collected 359 comments, 111 of which were excluded as they related to comments about the digital stories that were shown or about the patient panel itself. Therefore, we analyzed the content of 248 comments about patient preferences for the *StoryBank* [42]. After detailed review of the data, two authors (RI and JH) developed the standardized coding categories. RI and JH then independently coded the comments, with any discrepancies resolved through discussion. Patient preferences for the digital stories were examined by calculating simple frequencies on all coding categories. Information obtained during this task was used to plan the *StoryBank* project, and patients were also invited to register to share their own cancer genetics story.

Participants

In total, 75 patients at the patient panels said they would like to take part in the *StoryBank*. In spring 2010, further information about the project and a consent form were sent to these patients. Patients were asked to return the consent form to confirm they wanted to share their story and were happy to be contacted. From across Wales, 24 patients returned a consent form, although one individual did not progress past the phone call stage. The 23 patients who shared their story ranged in age from 28 to 75 years and 19 were female; 12 (52%) were classified as being at high risk, 6 (26%) as moderate risk, and 3 (13%) as average risk; 2 (9%) were relatives of patients who had been referred to the service.

Procedure

Once a consent form had been received, each participant was contacted by telephone to arrange a convenient date and location to record their story. The majority of participants (20/23, 87%) chose to record their story at home.

During the appointments, which took the form of an unstructured interview, participants were asked to describe their experience of the cancer genetics service. The researchers' role was simply to facilitate the telling of the story and to ask patients for more information when clarification was needed. All recordings were made using a high-quality audio recorder. The audio recordings lasted between 17 and 60 minutes. Participants were asked to provide photographs to illustrate their story, if they desired. Only one patient did not want to use personal photographs and another patient wished to remain anonymous and so provided photographs that did not include family members or friends.

Stories were also collected from 12 members of staff working in, or with, CGSW. Staff included specialists in clinical genetics, a family history coordinator, and a clinical scientist at the All Wales Molecular Genetics Laboratory. Staff provided factual stories about cancer, genetics, CGSW, and the processes involved. Recordings of staff stories lasted between 10 and 45 minutes.

All interviews were transcribed verbatim. Both patient and staff stories were edited to short clips, between 40 and 300 seconds in duration. More than 1 story was edited for 12 of the storytellers. All stories were seen and verified by every storyteller before being uploaded onto the website. The 53 stories from patients and staff are hosted on Vimeo (Vimeo, LLC, New York, NY, USA), a video hosting website, and the stories are embedded within the *StoryBank* website [41].

Evaluation

The *StoryBank* was evaluated on two levels: first, a brief telephone follow-up interview was conducted with all 23 patient storytellers approximately 2 weeks after the storytelling appointment, to explore their reasons for taking part in the *StoryBank* and their experiences of telling their story. All interviews were recorded and transcribed verbatim. The content of the follow-up interview was analyzed to summarize patients' evaluation of sharing their story and their views about the *StoryBank* [42]. Each transcript was read carefully and all relevant text was highlighted. All highlighted text from three to four transcripts was used to develop preliminary categories. The remaining transcripts were coded using these categories and by adding new categories where data did not fit into an existing category [43]. The data in all the categories were then examined independently by two authors (RI and JH), which led to some categories being combined. This resulted in three main categories of storytellers' reasons for contributing to the *StoryBank* and how they hoped their stories would benefit website users. For the second level of evaluation, a brief online survey was incorporated into the website to capture data about who was using the website and to give users the opportunity to leave feedback about the site. As this evaluation is ongoing, data from the online survey will be presented elsewhere.

Results

Patient Preferences for the StoryBank

At the patient panels, patients suggested topics that should be included in the *StoryBank*. These included tracing family histories; living with the risk of cancer; and telling family members, such as children, about their risk of inherited cancer. On the feedback forms, patients commented that the stories should convey personal experiences and practical information about the service and cancer genetics. Patients also expressed a preference for the *StoryBank* to include stories about a range of cancers and different levels of risk. Patient comments about the *StoryBank* and how they were used to inform the development of the website are shown in Table 1. Feedback about the project at the patient panels was generally positive with comments such as *"I would have liked to have seen a digital story when I was going through the journey, for a better understanding"* and *"a great idea and a valuable service"*.

Table 1. Patient preferences for the StoryBank (N = 248) and how they have been considered during website development

Coding category	Patient comments	Number of comments	% of total comments (N = 248)	How suggestions informed website development
Stories about personal experiences	*A good digital story would be how a patient dealt with their journey through the genetic history process.*	74	30	23 patient stories were collected and hosted on the website about their experiences of cancer and cancer genetics
Stories providing service information	*Initial digital stories need to be informative about the service including contact information.*	69	28	12 staff stories containing factual information about the service were collected and hosted on the website
The importance of including a variety of stories (eg, different age groups, different cancers)	*Variety of stories, ages, backgrounds. Different stories for different cancers.*	28	11	Male and female storytellers are a range of ages, affected by different cancers and at various levels of risk
Preference for positive stories	*In general stories need to be more positive. People don't want to hear the bad, they want encouragement and support.*	26	11	Many of the stories hosted on the website relate to the benefits of being referred to the cancer genetics service
Stories about the impact of genetics on families	*What it entails for partners of people affected .*	13	5	Storytellers shared their personal experiences of cancer genetics and the impact this had on their families, especially their children
Information about other support services	*Relevant information about various different counseling bodies available from Samaritans to grief counselors*	12	5	The website contains links to external sources of support including charities such as Tenovus, the cancer charity in Wales.
Stories from family members of patients	*Show the whole family's feelings, including older and younger members.*	10	4	2 storytellers are related to CGSW patients and their stories are about the impact upon them
Would like a discussion forum	*Forum to discuss details on genetics — knowing you have a faulty gene which has yet to be identified and how you cope with it.*	2	1	A discussion forum is not currently feasible, although users of the website are encouraged to contact CGSW if they have any queries or concerns
Other comments (eg, relating to own experience of cancer)		14	6	Not applicable

Developing the Website

A total of 53 individual stories about the various aspects of the cancer genetics service were produced. Of these, 37 clips are from patients about their personal experiences and 16 clips are from staff providing more practical information about the service (see Multimedia Appendix 1). The content of the *StoryBank* was developed using the information provided directly from service users at the patient panels. First, the routine patient journey through CGSW was identified and divided into key topics. These topics were as follows:

- What is the Cancer Genetics Service for Wales?
- Completing the family history questionnaire
- Waiting for the results of a risk assessment
- Being told if you are at low, moderate, or high risk of cancer
- Attending genetic counseling
- Having a genetic test
- Making decisions about screening and surgery

- Living with cancer
- Information about rare cancers
- Sources of support.

Next, the main themes of each patient and staff story were identified and mapped onto the key topics. The stories are presented on the website under these key topics (see Figure 2). However, there is also a page on the site that lists all the storytellers and their stories, so that users can also search for stories via the storyteller in order to see the journey one particular person might have gone through [41]. All the stories were transcribed and the transcripts are available to download from the website as PDF files (see Figure 3 and Multimedia Appendix 2).

The website was designed with the aid of a professional Web designer, and patients and staff at CGSW were invited to give their feedback about the website before it was launched. A special event was held to mark the launch of the website on November 25, 2010. The event was opened by the First Minister of Wales, Rt Hon Carwyn Jones AM, and was attended by 105 guests, including the storytellers and their families; staff at CGSW; and others interested in cancer, genetics, and promoting patient-centered service delivery across NHS Wales. In the first month (November 25, 2010 to December 25, 2010) since the website was launched, the stories were viewed 494 times and were played in 15 countries across the world, including the USA, Australia, South Korea, the Netherlands, and Spain.

Figure 2. Screenshot from the *StoryBank* (www.cancergeneticsstorybank.co.uk).

Welcome to the StoryBank
Annie Procter, Clinical Director of the All Wales Medical Genetics Service, welcomes you to the StoryBank

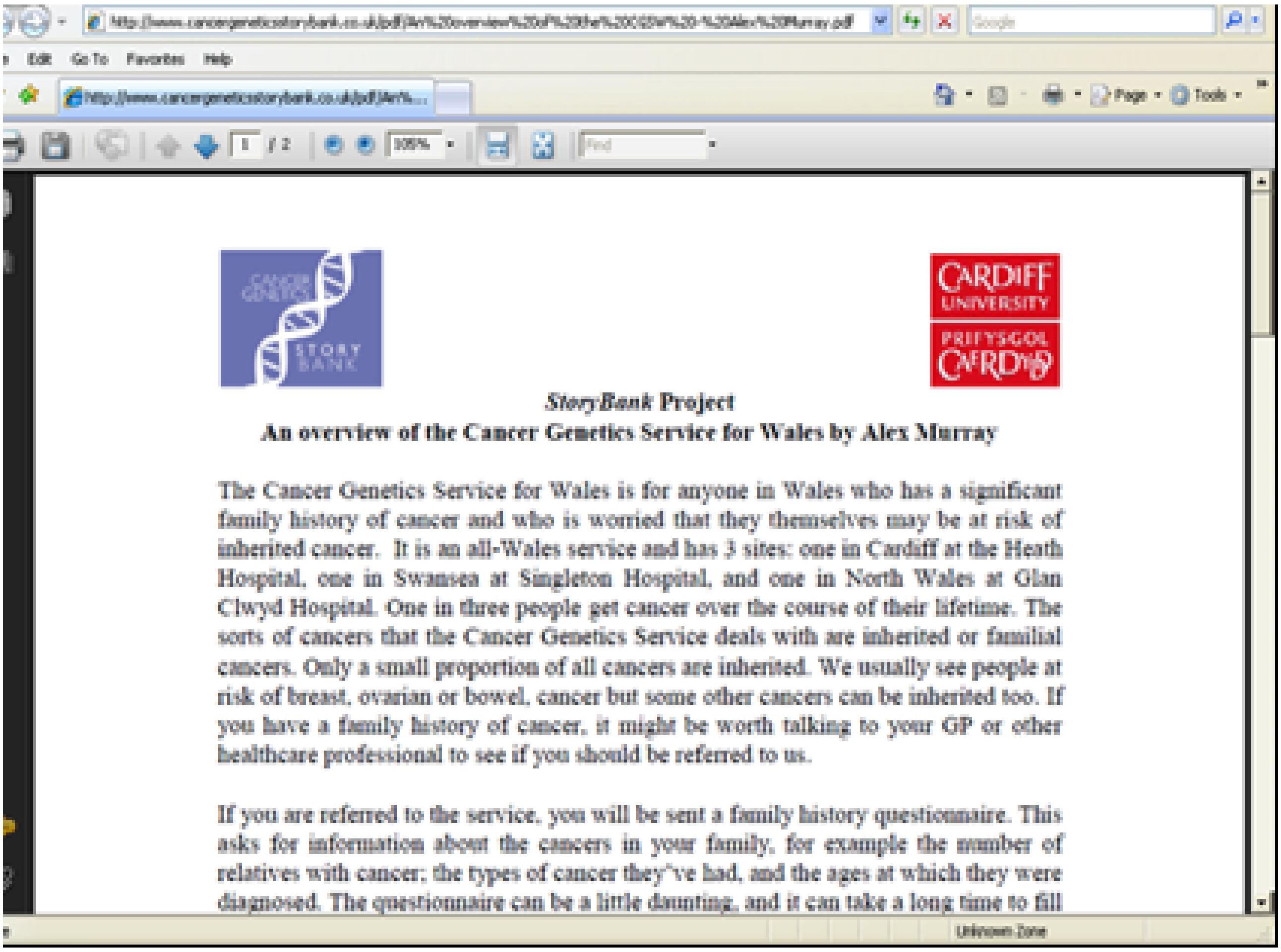


Figure 3. Screenshot of PDF story transcript from the *StoryBank* (www.cancergeneticsstorybank.co.uk).

***StoryBank* Project**

An overview of the Cancer Genetics Service for Wales by Alex Murray

The Cancer Genetics Service for Wales is for anyone in Wales who has a significant family history of cancer and who is worried that they themselves may be at risk of inherited cancer. It is an all-Wales service and has 3 sites: one in Cardiff at the Heath Hospital, one in Swansea at Singleton Hospital, and one in North Wales at Glan Clwyd Hospital. One in three people get cancer over the course of their lifetime. The sorts of cancers that the Cancer Genetics Service deals with are inherited or familial cancers. Only a small proportion of all cancers are inherited. We usually see people at risk of breast, ovarian or bowel, cancer but some other cancers can be inherited too. If you have a family history of cancer, it might be worth talking to your GP or other healthcare professional to see if you should be referred to us.

If you are referred to the service, you will be sent a family history questionnaire. This asks for information about the cancers in your family, for example the number of relatives with cancer; the types of cancer they've had, and the ages at which they were diagnosed. The questionnaire can be a little daunting, and it can take a long time to fill

Evaluation

Data from the follow-up interviews demonstrate that the storytellers were all positive about the *StoryBank* and the experience of sharing their story. Table 2 shows the three coding categories that were developed from the content analysis of the interview transcripts. A majority of the storytellers (19/23, 83%) stated that they wanted to share their story as a means of helping others who were going through a similar experience. For example, one patient said

> *I'd done a lot of Internet searching and I couldn't find a story similar to mine. So I thought if I get an opportunity to say my story then I will,because at some point there will be somebody going through the same thing as me.*

Another patient explained

> *I've got first hand experience of what it's like and hopefully somebody else can experience positive feelings from what I've said*

Table 2. Storytellers' reflections on sharing their story (N = 23)

Coding category	Examples	N (%)
Providing reassurance and support for other patients	*... to support people through it, to let them know there's light at the end of the tunnel and to know they're not on their own everybody feels that way or a lot of people feel that way.* *We have a very rare, it's a very rare cancer and I kind of think rare cancers need to be included you know because there's a lot written, lots of support and everything else for the common cancers and I think then that makes, well certainly made me with a rare cancer feel kind of more isolated if you like.*	19 (83)
Increasing knowledge and understanding of the cancer genetics service	*When we were involved with the genetics service, we'd not known anything about it, um and we just felt that if we could do something that would help others understand what might be in front of them, then it might be worth doing.* *I think that it's really quite important that people go with that in mind ... that they understand that the test isn't absolutely the end of the story. If you get a negative result there still might be something beyond that.*	9 (39)
Reducing the fear and uncertainty surrounding cancer genetics	*Well I hope people who listen or see it will find it encouraging to go forward and seek advice.* *I just hope that through my experience and through the way I've tried to explain things, it makes it less of a specter and it makes it less of an unknown quantity.*	6 (26)

Discussion

During 3 years of patient-centered research at CGSW we found that patients want access to factual information about cancer genetic services and to hear personal stories from a variety of patients who are similar to themselves in terms of cancer genetic risk level, age, and gender [20,35]. Patients were invited to attend a patient panel to suggest topics to be included in the *StoryBank*, a Web-based resource of digital stories from patients and staff at CGSW. These suggestions were used to ensure that the *StoryBank* met the information and support needs of current and future patients and that it was developed in accordance with patient preferences. Increasing patient involvement in service development is vital because it leads to improved communication and decision making, and ultimately to better health outcomes [37]. We suggest that the *StoryBank* project is an innovative example of involving patients in service development and delivery.

The success of websites such as healthtalkonline (formerly DIPEx) [13,44] and the use of patient stories on UK cancer charity websites such as Breast Cancer Care and Macmillan demonstrates that patients want, and indeed use, the opportunity to listen to individuals who have had similar experiences. Incorporating stories from CGSW patients with information from staff members, and mapping these onto every key stage of the cancer genetics journey, allows us to provide information in a form that is reliable, accessible, and engaging. Providing patients with the information they require allows individuals to have a better understanding of the risk assessment and associated processes [31]. This in turn results in several benefits such as more realistic patient expectations, and ultimately helps individuals cope with their risk assessment and risk categorization [29,30]. The *StoryBank* is one of the first places where patient stories have been aligned with the patient pathway through a health service. By providing narrative accounts from patients who have already been referred to CGSW, about how they have coped with the emotional challenges they faced, we hope that future patients will gain encouragement and support as they progress on their own journey through the service.

The *StoryBank* also informs patients about the psychosocial aspects of genetic testing and where to access additional information and support, which are often omitted from standard information leaflets provided in genetics clinics [45]. Presenting health information in a digital format can also be beneficial for patients with low literacy levels [46], and the availability of cancer genetic information before counseling may help patients with lower literacy to prepare for their clinic appointment [47]. Patient narratives have also been shown to improve understanding and aid recall of health-related information [16,36]. The data from the follow-up telephone interviews also show that participants wanted to share their story to increase other patients' understanding of the services provided by CGSW and to convey information about what to expect from genetic risk assessment, genetic counseling, and genetic testing.

By hosting patient and staff digital stories on a website, the *StoryBank* allows individuals to access information about CGSW and hear from other patients 24 hours a day, from anywhere in the world. This is particularly beneficial for CGSW, which, as an all-Wales service, receives referrals from patients across the country. Many people in Wales live in rural communities with poor transport links [48,49]. Thus, one of the challenges faced by CGSW was meeting the information and support needs of all its patients, while recognizing the geographical constraints of Wales. The *StoryBank* provides a solution to meeting these patient needs.

The challenge for the future is to ensure that health professionals are aware of the website and the information it provides. As individuals continue to turn to the Internet for health information, it is vital that health professionals are familiar with the information that is available and are able to direct patients to the relevant resources [50]. Staff at CGSW have been involved with the project from the patient panels to the development and launch of the website. Furthermore, all staff members with patient contact at CGSW have been asked to include the *StoryBank* website address on correspondence with patients. Our task is to continue to increase awareness of the website both within CGSW and more widely. For example, general practitioners are one of the main routes through which patients are referred to CGSW [19]. Thus, it is important that they understand the service provided and the area of cancer genetics more generally.

Conclusions

In conclusion, the *StoryBank* was developed in response to patient feedback to meet the long-term information and support needs of cancer genetics patients. The *StoryBank* provides a Web-based resource of reliable, accessible, and engaging information for current and future cancer genetics patients. The patients' stories communicate an experiential perspective of the cancer genetics service. The *StoryBank* is one of the first places where patient and staff stories have been mapped onto each stage of a service, which will help people gain a sense of the journey they too may experience. We suggest that the *StoryBank* could also be useful in raising awareness of cancer genetics and CGSW among health care professionals. The *StoryBank* is an innovative online resource, developed in collaboration with current patients, which could be applied to other genetic or other long-term chronic health conditions, such as diabetes and asthma, in the NHS and elsewhere.

Acknowledgments

We thank all the storytellers for contributing their stories to the StoryBank. The authors would like to thank Julie Hayward, Alex Murray, Rose Thompson and StoryWorks for their help with the StoryBank project. The work was funded by Tenovus, the Cancer Charity in Wales. The funders had no involvement in the research design or conduct or in the interpretation of results.

Conflicts of Interest

None declared

Multimedia Appendix 1

Digital story (video) from the StoryBank website about the Cancer Genetics Service for Wales.

[MP4 File (MP4 Video), 84MB - jmir_v13i3e78_app1.mp4]

Multimedia Appendix 2

PDF transcript of the digital story about the Cancer Genetics Service for Wales.

[PDF File (Adobe PDF File), 115KB - jmir_v13i3e78_app2.pdf]

References

1. Bernhardt JM, Lariscy RA, Parrott RL, Silk KJ, Felter EM. Perceived barriers to Internet-based health communication on human genetics. J Health Commun 2002;7(4):325-340. [doi: 10.1080/10810730290088166] [Medline: 12356290]
2. Eaton L. Europeans and Americans turn to internet for health information. BMJ 2002;325(7371):989.
3. UK Department of Health. 2000 Sep. The NHS Cancer Plan: A Plan for Investment, a Plan for Reform URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4014513.pdf [accessed 2011-06-15] [WebCite Cache ID 5zSRQPNhH]
4. Satterlund MJ, McCaul KD, Sandgren AK. Information gathering over time by breast cancer patients. J Med Internet Res 2003;5(3):e15 [FREE Full text] [doi: 10.2196/jmir.5.3.e15] [Medline: 14517106]
5. Ziebland S. The importance of being expert: the quest for cancer information on the Internet. Soc Sci Med 2004 Nov;59(9):1783-1793. [doi: 10.1016/j.socscimed.2004.02.019] [Medline: 15312914]
6. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Use of the Internet by women with breast cancer. J Med Internet Res 2002;4(2):E9 [FREE Full text] [doi: 10.2196/jmir.4.2.e9] [Medline: 12554556]
7. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Internet use and social support in women with breast cancer. Health Psychol 2002 Jul;21(4):398-404. [Medline: 12090683]
8. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371. [Medline: 15224975]
9. Raupach JC, Hiller JE. Information and support for women following the primary treatment of breast cancer. Health Expect 2002 Dec;5(4):289-301. [Medline: 12460218]
10. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
11. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 12020305]

12. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]
13. Herxheimer A, McPherson A, Miller R, Shepperd S, Yaphe J, Ziebland S. Database of patients' experiences (DIPEx): a multi-media approach to sharing experiences and information. Lancet 2000 Apr 29;355(9214):1540-1543. [doi: 10.1016/S0140-6736(00)02174-7] [Medline: 10801187]
14. Bender JL, Jimenez-Marroquin MC, Jadad AR. Seeking support on Facebook: a content analysis of breast cancer groups. J Med Internet Res 2011;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]
15. Overberg R, Otten W, de Man A, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the internet: an online randomized controlled experiment. J Med Internet Res 2010;12(1):e7 [FREE Full text] [doi: 10.2196/jmir.1215] [Medline: 20215101]
16. Kreuter MW, Holmes K, Alcaraz K, Kalesan B, Rath S, Richert M, et al. Comparing narrative and informational videos to increase mammography in low-income African American women. Patient Educ Couns 2010 Dec;81 Suppl:S6-14. [doi: 10.1016/j.pec.2010.09.008] [Medline: 21071167]
17. McQueen A, Kreuter MW. Women's cognitive and affective reactions to breast cancer survivor stories: a structural equation analysis. Patient Educ Couns 2010 Dec;81 Suppl:S15-S21. [doi: 10.1016/j.pec.2010.08.015] [Medline: 20850258]
18. UK Office for National Statistics. 2011 Jun 15. Wales Population Density: By Local or Unitary Authority, 2009 URL: http://www.statistics.gov.uk/cci/nugget.asp?id=1081 [accessed 2011-06-15] [WebCite Cache ID 5zSMhwHQT]
19. Elwyn G, Iredale R, Gray J. Reactions of GPs to a triage-controlled referral system for cancer genetics. Fam Pract 2002 Feb;19(1):65-71 [FREE Full text] [Medline: 11818352]
20. Iredale R, Brain K, Gray J, France E. The information and support needs of women at high risk of familial breast and ovarian cancer: how can cancer genetic services give patients what they want? Fam Cancer 2003;2(2):119-121. [Medline: 14649669]
21. Emery J. Is informed choice in genetic testing a different breed of informed decision-making? A discussion paper. Health Expect 2001 Jun;4(2):81-86. [Medline: 11359537]
22. Bennett P, Wilkinson C, Turner J, Edwards RT, France B, Griffith G, et al. Factors associated with intrusive cancer-related worries in women undergoing cancer genetic risk assessment. Fam Cancer 2009;8(2):159-165. [doi: 10.1007/s10689-008-9221-9] [Medline: 19011994]
23. Brain K, Norman P, Gray J, Rogers C, Mansel R, Harper P. A randomized trial of specialist genetic assessment: psychological impact on women at different levels of familial breast cancer risk. Br J Cancer 2002 Jan 21;86(2):233-238 [FREE Full text] [doi: 10.1038/sj.bjc.6600051] [Medline: 11870512]
24. Metcalfe KA, Liede A, Hoodfar E, Scott A, Foulkes WD, Narod SA. An evaluation of needs of female BRCA1 and BRCA2 carriers undergoing genetic counselling. J Med Genet 2000 Nov;37(11):866-874 [FREE Full text] [Medline: 11073541]
25. Brain K, Gray J, Norman P, Parsons E, Clarke A, Rogers C, et al. Why do women attend familial breast cancer clinics? J Med Genet 2000 Mar;37(3):197-202 [FREE Full text] [Medline: 10699056]
26. Halbert C, Kessler L, Collier A, Paul Wileyto E, Brewster K, Weathers B. Psychological functioning in African American women at an increased risk of hereditary breast and ovarian cancer. Clin Genet 2005 Sep;68(3):222-227. [doi: 10.1111/j.1399-0004.2005.00483.x] [Medline: 16098010]
27. Hawkins RP, Pingree S, Shaw B, Serlin RC, Swoboda C, Han JY, et al. Mediating processes of two communication interventions for breast cancer patients. Patient Educ Couns 2010 Dec;81 Suppl:S48-S53. [doi: 10.1016/j.pec.2010.10.021] [Medline: 21081261]
28. Scott S, Prior L, Wood F, Gray J. Repositioning the patient: the implications of being 'at risk'. Soc Sci Med 2005 Apr;60(8):1869-1879. [doi: 10.1016/j.socscimed.2004.08.020] [Medline: 15686817]
29. Metcalfe A, Werrett J, Burgess L, Chapman C, Clifford C. Cancer genetic predisposition: information needs of patients irrespective of risk level. Fam Cancer 2009;8(4):403-412. [doi: 10.1007/s10689-009-9256-6] [Medline: 19517270]
30. Phelps C, Wood F, Bennett P, Brain K, Gray J. Knowledge and expectations of women undergoing cancer genetic risk assessment: a qualitative analysis of free-text questionnaire comments. J Genet Couns 2007 Aug;16(4):505-514. [doi: 10.1007/s10897-007-9086-3] [Medline: 17318449]
31. Hallowell N, Murton F, Statham H, Green JM, Richards MP. Women's need for information before attending genetic counselling for familial breast or ovarian cancer: a questionnaire, interview, and observational study. BMJ 1997 Jan 25;314(7076):281-283 [FREE Full text] [Medline: 9022494]
32. Henderson BJ, Maguire BT. Lay representations of genetic disease and predictive testing. J Health Psychol 1998;3:233-241. [doi: 10.1177/135910539800300207]
33. Pieterse AH, Ausems MG, Van Dulmen AM, Beemer FA, Bensing JM. Initial cancer genetic counseling consultation: change in counselees' cognitions and anxiety, and association with addressing their needs and preferences. Am J Med Genet A 2005 Aug 15;137(1):27-35. [doi: 10.1002/ajmg.a.30839] [Medline: 16007615]
34. Hughes L, Phelps C. "The bigger the network the bigger the bowl of cherries...": exploring the acceptability of, and preferences for, an ongoing support network for known BRCA 1 and BRCA 2 mutation carriers. J Genet Couns 2010 Oct;19(5):487-496. [doi: 10.1007/s10897-010-9300-6] [Medline: 20411314]

35. Iredale R, Mundy L, Marsden D, Murray A. Patient involvement at the Cancer Genetics Service for Wales: meeting the long-term information and support needs of people at risk of inherited cancer. Patient Educ Couns 2011 Aug;84(2):280-281. [doi: 10.1016/j.pec.2011.03.004] [Medline: 21482062]
36. Wilcock PM, Brown GC, Bateson J, Carver J, Machin S. Using patient stories to inspire quality improvement within the NHS Modernization Agency collaborative programmes. J Clin Nurs 2003 May;12(3):422-430. [Medline: 12709117]
37. Crawford MJ, Rutter D, Manley C, Weaver T, Bhui K, Fulop N, et al. Systematic review of involving patients in the planning and development of health care. BMJ 2002 Nov 30;325(7375):1263 [FREE Full text] [Medline: 12458240]
38. UK Department of Health. 2005 Mar 17. Creating a Patient-Led NHS: Delivering the NHS Improvement Plan URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4106507.pdf [accessed 2011-06-15] [WebCite Cache ID 5zSRcdtLQ]
39. Welsh Assembly Government. 2003 Sep. Signposts Two: Putting Public and Patient Involvement Into Practice in Wales URL: http://wales.gov.uk/docrepos/40382/dhss/reportsenglish/signposts2-e.pdf;jsessionid=qGTGN4McT9b6dwPjC0Trh1dY5jyFjbJ6NxLfhpl0nLpxXjbCh969!-342331487?lang=en [accessed 2011-06-15] [WebCite Cache ID 5zSRoRvhQ]
40. Expert Advisory Group on Cancer to the Chief Medical Officers of England and Wales. UK Department of Health, Welsh Office. 1995 Apr. A Policy Framework for Commissioning Cancer Services URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4014366.pdf [accessed 2011-06-15] [WebCite Cache ID 5zSRv79QU]
41. Cancer Genetics Research Team. Institute of Medical Genetics, Cardiff University School of Medicine. 2010. Cancer Genetics StoryBank URL: http://www.cancergeneticsstorybank.co.uk/storytellers.htm [accessed 2011-06-15] [WebCite Cache ID 5zSN1ghoU]
42. Weber RP. Basic Content Analysis. 2nd edition. Newbury Park, CA: Sage Publications; 1990.
43. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
44. DIPEx. 2008. Healthtalkonline URL: http://www.healthtalkonline.org/ [accessed 2011-06-15] [WebCite Cache ID 5zSN5rDHs]
45. Lewis C, Mehta P, Kent A, Skirton H, Coviello D. An assessment of written patient information provided at the genetic clinic and relating to genetic testing in seven European countries. Eur J Hum Genet 2007 Oct;15(10):1012-1022 [FREE Full text] [doi: 10.1038/sj.ejhg.5201874] [Medline: 17568389]
46. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]
47. Joseph G, Beattie MS, Lee R, Braithwaite D, Wilcox C, Metrikin M, et al. Pre-counseling education for low literacy women at risk of hereditary breast and ovarian cancer (HBOC): patient experiences using the Cancer Risk Education Intervention Tool (CREdIT). J Genet Couns 2010 Oct;19(5):447-462. [doi: 10.1007/s10897-010-9303-3] [Medline: 20490636]
48. Iredale R, Gray J, Murtagh G. Telegenetics: a pilot study of videomediated genetic consultations in Wales. J Med Marketing 2002;2:130-135. [doi: 10.1057/palgrave.jmm.5040066]
49. Statistics for Wales. Welsh Assembly Government. 2009. Wales's Population: A Demographic Overview, 2009 URL: http://wales.gov.uk/docs/statistics/2009/090326walespop09en.pdf [accessed 2011-06-15] [WebCite Cache ID 5zSS3xENB]
50. Alexander J, Zeibland S. The Web: bringing support and health information into the home: the communicative power of qualitative research. Int J Nurs Stud 2006 May;43(4):389-391. [doi: 10.1016/j.ijnurstu.2005.10.012] [Medline: 16438973]

Abbreviations

CGSW: Cancer Genetics Service for Wales
NHS: National Health Service

Edited by G Eysenbach; submitted 24.01.11; peer-reviewed by G Dolan, H Skirton; comments to author 25.04.11; revised version received 16.06.11; accepted 16.06.11; published 30.09.11.

Please cite as:
Iredale R, Mundy L, Hilgart J
An Online Resource of Digital Stories About Cancer Genetics: Qualitative Study of Patient Preferences and Information Needs
J Med Internet Res 2011;13(3):e78
URL: http://www.jmir.org/2011/3/e78/
doi:10.2196/jmir.1735
PMID:22057223

Original Paper

Two Complementary Personal Medication Management Applications Developed on a Common Platform: Case Report

Stephen E Ross[1], MD; Kevin B Johnson[2], MS, MD; Katie A Siek[3], PhD; Jeffry S Gordon[4], PhD; Danish U Khan[3], MS; Leah M Haverhals[5], MA

[1]Division of General Internal Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States
[2]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, United States
[3]Wellness Innovation and Interaction Lab, Department of Computer Science, University of Colorado Boulder, Boulder, CO, United States
[4]School of Nursing, Vanderbilt University, Nashville, TN, United States
[5]Colorado REAP (Research Award Enhancement Program) to Improve Care Coordination, Veterans Affairs Medical Center, Denver, CO, United States

Corresponding Author:
Stephen E Ross, MD
Division of General Internal Medicine
School of Medicine
University of Colorado Anschutz Medical Campus
Mail Stop B180
12631 E. 17th Avenue
Aurora, CO, 80045
United States
Phone: 1 303 724 2267
Fax: 1 303 724 2270
Email: Steve.Ross@ucdenver.edu

Abstract

Background: Adverse drug events are a major safety issue in ambulatory care. Improving medication self-management could reduce these adverse events. Researchers have developed medication applications for tethered personal health records (PHRs), but little has been reported about medication applications for interoperable PHRs.

Objective: Our objective was to develop two complementary personal health applications on a common PHR platform: one to assist children with complex health needs (MyMediHealth), and one to assist older adults in care transitions (Colorado Care Tablet).

Methods: The applications were developed using a user-centered design approach. The two applications shared a common PHR platform based on a service-oriented architecture. MyMediHealth employed Web and mobile phone user interfaces. Colorado Care Tablet employed a Web interface customized for a tablet PC.

Results: We created complementary medication management applications tailored to the needs of distinctly different user groups using common components. Challenges were addressed in multiple areas, including how to encode medication identities, how to incorporate knowledge bases for medication images and consumer health information, how to include supplementary dosing information, how to simplify user interfaces for older adults, and how to support mobile devices for children.

Conclusions: These prototypes demonstrate the utility of abstracting PHR data and services (the PHR platform) from applications that can be tailored to meet the needs of diverse patients. Based on the challenges we faced, we provide recommendations on the structure of publicly available knowledge resources and the use of mobile messaging systems for PHR applications.

(J Med Internet Res 2011;13(3):e45) doi:10.2196/jmir.1815

KEYWORDS

Health records, personal; drug information services; medical informatics

Introduction

Medication management accounts for the majority of medical errors in ambulatory care [1,2]. Errors in home administration account for many of these errors [3-5], particularly after care transitions, such as being discharged home from the hospital [6,7]. Some errors occur because patients commonly have difficulty maintaining an accurate, current list of increasingly complex medication regimens [8,9]. Lay persons find generic and proprietary medication names to be inscrutable and redundant. Patients also lack critical information about the medicines they take. In busy practices, clinicians only inconsistently review medication regimens and warn about the potential side effects of new medications [10,11]. Even with assistance from pharmacists and resources such as medication information sheets, many patients remain uncertain about key medication questions [12].

The Institute of Medicine recommends patient-centered approaches to address these deficiencies: fostering a strong consumer–provider partnership in medication management, enhancing communication, and developing tools for “patient (or surrogate) self-management support” [2]. Groups such as the Markle Foundation [13] and the Commission for Systemic Interoperability [14] note the role personal health records (PHRs) can play in improving medication management. Paper PHRs are familiar in pediatric practice [15,16] and improve medication management in adult care transitions [17,18]. Electronic PHRs have also shown promise as aids to medication management [2,19]. These include *standalone* PHRs such as MyMedicationList, which links patient-entered medication data to consumer health information [20], and *tethered* PHRs such as the Patient Gateway medications module, which allows patients to review, track, and communicate with physicians about the medication list derived from a leading institution’s electronic medical record [19].

Interoperable PHRs promise to empower patients even more. For medication management, an interoperable platform for PHR data could improve coordination of care by consolidating multiple sources of prescribing data (from the electronic medical records of multiple independent practices) and fulfillment or dispensing data (from pharmacy records and claims), allowing patients to share these data at their discretion [14,21,22]. An application layer could enrich these data with consumer health information [23], tools that identify drug interactions and duplications, and scheduling applications. Mobile applications could support reminders to take medications and facilitate communication among patients and caregivers. However, with these advantages also come the daunting challenges of designing devices and user interfaces that are reliable, are easy to use, and present complex information in ways that consumers find straightforward and helpful.

As part of Project HealthDesign [24] we explored these opportunities and challenges. Project HealthDesign was launched in December 2006 (before the availability of commercial PHR platforms such as Indivo, Microsoft HealthVault, and Google Health) to “demonstrate the power and potential” of interoperable PHRs. Nine teams participated in the project, each representing a different target user and use case. Our two teams had a complementary focus on medication management for patients with chronic diseases. The Vanderbilt University team developed MyMediHealth (MMH) for children with complex illnesses [25]. The University of Colorado team developed the Colorado Care Tablet (CCT) for older adults prone to care transitions [26]. While each application tailored its user interface for its target population, each used a common interoperable PHR platform [27]. Here we report how the two applications shared common services and how we addressed key informatics and user interface challenges related to ambulatory medication self-management.

Methods

Iterative Development Process

For all Project HealthDesign grantees, the primary objective was to create a personal health application that would be compelling for the targeted user group. The target users of MMH were children with complex diseases such as cystic fibrosis and their caregivers (parents or guardians). The target users of CCT were older patients with multimorbidity (2 or more chronic diseases such as diabetes, hypertension, or heart failure) taking multiple medications. These adults are prone to fragmentation of care through minor care transitions (seeing doctors with separate medical records systems) and major care transitions (transitions to and from the hospital).

During a 6-month design phase, project teams developed functional requirements based on a series of individual in situ interviews (eg, home, school, day care) and facilitated group discussions with target users. While a detailed description of the data collected and the analysis methods used is beyond the scope of this report, a brief description of the interviews and settings used is provided in Table 1. During the 12-month prototype phase, both groups employed iterative, user-centered design techniques to evaluate prototypes and provide the target population with a continuous voice in the design cycle. The methods employed during all phases of the project were approved by the Institutional Review Boards at the University of Colorado and Vanderbilt University.

Table 1. Users and settings studied in user-centered design process

Vanderbilt: MyMediHealth	Colorado: Colorado Care Tablet
Design phase	Design phase
3 group sessions with parents of children with cystic fibrosis, school officials, before/after school care staff, and school nurses	12 individual interviews in home setting, 1 in hospital setting, with 15 primary users aged 73-90 years (mean 82) and 2 family caregivers aged 48 and 57
3 day care site visits	4 group interviews with 27 primary users over age 65 years
4 school site visits	
Prototype phase	Prototype phase
2 group sessions with parents, school officials, before/after school care staff, and school nurses	Review of storyboard
1-month pilot of paging system with 20 children who had cystic fibrosis	6 individual sessions with 7 participants, 5 older adults from the target user group aged 70–85 years (mean 75) and 2 caregivers aged 75 and 82 years
Storyboard review by 200 families of children with daily chronic medication needs	2 group sessions with 9 older adults from the target user group aged 80–88 years (mean 83) and 3 caregivers aged 48–59 years (mean 53)
	6 rapid iterative testing and evaluation sessions [28,29] with a total of 22 primary users aged 61–86 years (mean 76) and 9 caregivers aged 41–61 years (mean 53)

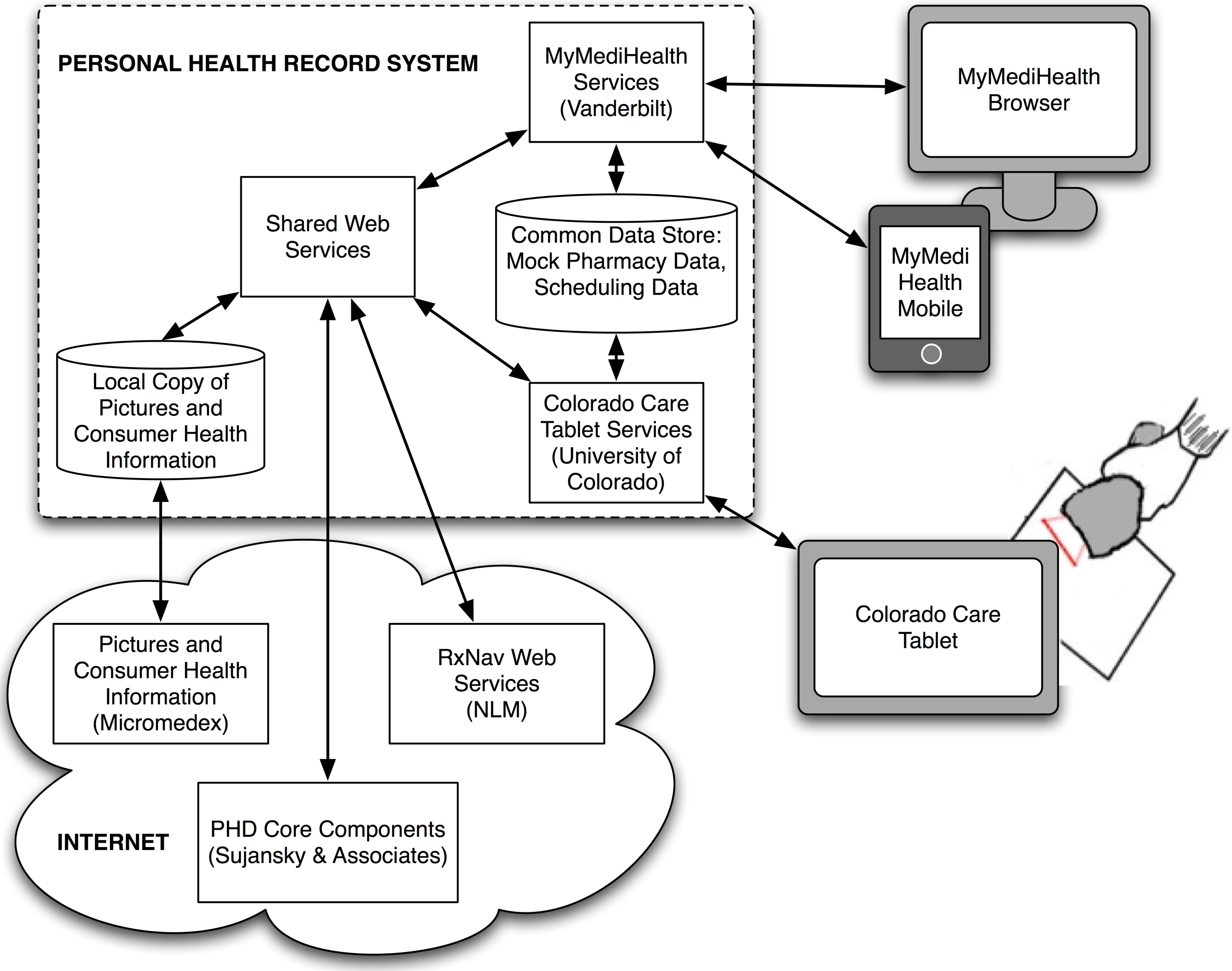
Architecture: Shared Components

Early in the development process, it was clear that both applications would need a common data store and shared functions to (1) normalize medication identities (ie, translating between National Drug Codes [NDCs], RxNorm concept unique identifiers [RXCUIs], and proprietary identifications to identify duplicate medications and ingredients in medication lists), (2) link medications to consumer health information, and (3) link medications to images wherever possible. We employed a shared knowledge approach to take advantage of efficiencies in development and to provide users the ability to switch between the two applications (eg, to use CCT to build a medication list and MMH to schedule and prompt medication use). The client-server architecture, shown in Figure 1, used thin clients (off-loading the computing software to the PHR system) to make for robust, flexible, and scalable prototypes.

The *PHR platform* (PHD Core Components) used a service-oriented architecture to authenticate users and to store and retrieve various data types [27]. Of note, these prototypes did not receive, transmit, or store medication data for actual patients. MMH and CCT used simulations of electronic health record-based prescribing data and (in the case of CCT) dispensing data (ie, data available from pharmacy and claims) to test user interface scenarios. MMH also used its own local storage for timing of alerts and recording medication administration events.

In addition to the common platform, CCT and MMH used RxNav Web services [30,31] from the National Library of Medicine for normalization of medications. A commercial medication knowledge base (Micromedex; Thomson Reuters, Greenwood Village, CO, USA) was used to supply images of medications and consumer health information. Although we endeavored to use open-source tools wherever possible, no publicly available content was available for these items at the time of development.

Figure 1. Architecture of two personal medication management applications



Medication Identity: Representation and Linkage to Knowledge Bases

The PHR platform allowed for flexibility in representations. Each medication record could be represented by a coding scheme (eg, “RXCUI”) and a code (eg, “20610”). Text entries such as generic names, trade names, and free text entries (eg, “blue pill”) were also allowed, since the target populations commonly thought of their medications by color and context. However, only 1 code could be associated with each record.

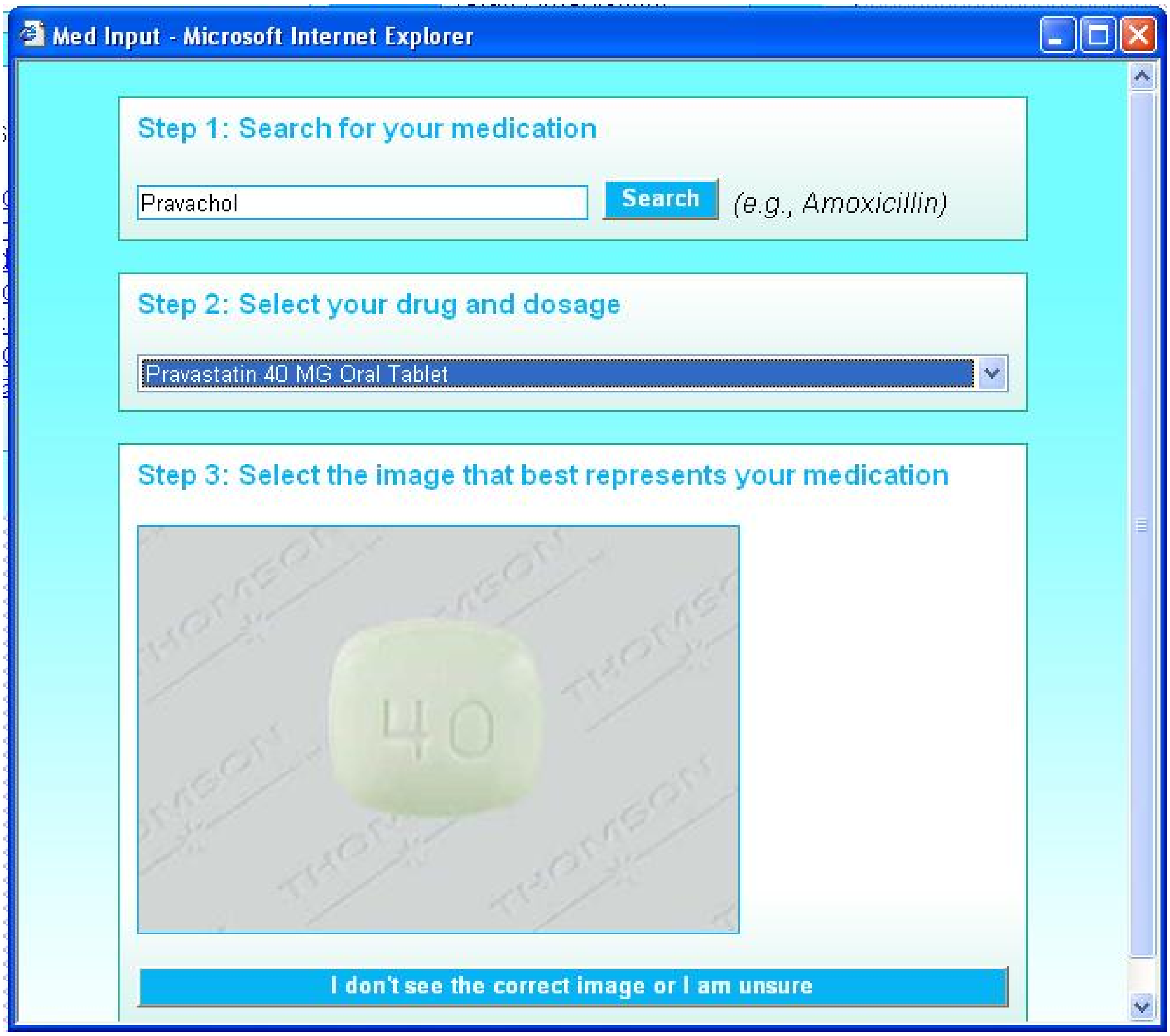
Ultimately, whenever possible we stored medication identities as NDCs. The NDC is widely used for representing medications in electronic prescribing and fulfillment data (eg, data from Surescripts medication history) and allows images to be associated with medication identities. We were able to use RxNav Web services to normalize NDC representations whenever a more abstract concept (such as the medication ingredient) was needed, but this required additional processing for normalization. In the future, storing multiple representations in each medication record (eg, storing various RXCUIs in addition to the NDC, as is possible with the ASTM Continuity of Care Record and HL7 Continuity of Care Document schema) in advance would reduce processing cycles when the application is run.

Recognizing that some medications would need to be entered manually by users, we developed systems to assist capture of codified data, rather than simple free text. CCT and MMH employed parallel functionality for this purpose, as shown in Figure 2. When users typed in part or all of a medication name, the application displayed a list of candidate medications. When corresponding images existed, they were presented to the patient. When the user confirmed the image of the medication to be entered, the associated NDC was stored as the medication identity. If the name of the medication matched, but none of the images matched, then the application could not derive an NDC, and the RXCUI associated with the semantic clinical drug name was stored. This functionality was supported using RxNav services and the Micromedex drug image database:

- The text string entered by the user was processed by the RxNav spell check function. If the name was not recognized, alternative spellings were suggested.
- RxNav linked the text string to a semantic clinical drug name.
- Putative NDCs were derived.
- Images for each putative NDC were retrieved from the Micromedex drug image database (which was indexed by NDC).
- Images were displayed for user selection.

The RxNav Web services proved well suited for this function. Response time was typically 1–2 seconds, and the services were consistently available. RxNav was considerably easier to implement than downloading RxNorm tables and updating them when new versions were released. The Micromedex drug image database also generally performed well; however, it often lacked entries for putative NDCs that RxNav generated for generic medications. Images were most commonly available for solid medication forms (eg, capsules, tablets), less frequently for inhaled forms, and rarely for liquid forms.

Figure 2. Medication selection user interface for both applications



Step 1: Search for your medication

Pravachol

(e.g., Amoxicillin)

Step 2: Select your drug and dosage

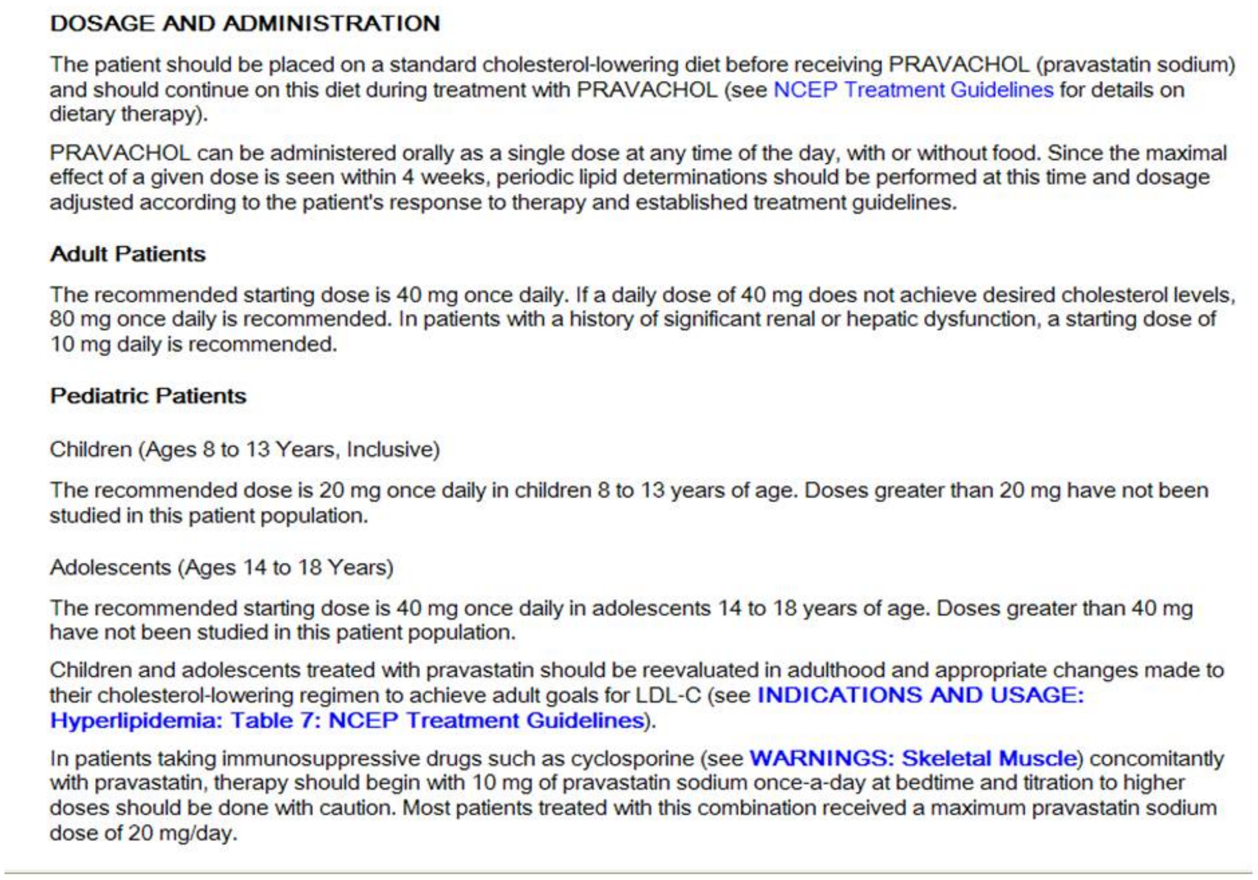
Pravastatin 40 MG Oral Tablet

Step 3: Select the image that best represents your medication

Dosing Frequency: Representation and Linkage to Knowledge Databases

Representation of dosing frequency proved more challenging than representation of the medication identity. We endeavored to include representations that would facilitate machine-actionable decision support to assist users with scheduling, but resources available at the time were inadequate in several ways. One challenge was making guidelines computable. Guidelines on frequency of administration from drug knowledge databases (such as DailyMed) are available only in descriptive form, not in a codified, computable form (Figure 3). While it is reasonably easy to convert descriptions of simple frequencies into machine-actionable representations, it is much more difficult to encode important additional descriptive constraints on dosing, particularly in relation to food consumption. For example, tetracycline, a medicine commonly given to children in the MMH target user group, should be taken with a glass of water on an empty stomach, half an hour before or 2 hours after meals, and never at the same time as antacids or iron. Another challenge was capturing dosing information from prescribing data. While the National Council on Prescription Drug Programs SCRIPT standard for structured and codified SIG (dosing instructions) includes the necessary structure for basic instructions, it does not support complex instructions or timing of doses [32]. Thus, complex dosing instructions in prescribing data were embedded in noncomputable text strings.

Figure 3. DailyMed entry for pravastatin



DOSAGE AND ADMINISTRATION

The patient should be placed on a standard cholesterol-lowering diet before receiving PRAVACHOL (pravastatin sodium) and should continue on this diet during treatment with PRAVACHOL (see NCEP Treatment Guidelines for details on dietary therapy).

PRAVACHOL can be administered orally as a single dose at any time of the day, with or without food. Since the maximal effect of a given dose is seen within 4 weeks, periodic lipid determinations should be performed at this time and dosage adjusted according to the patient's response to therapy and established treatment guidelines.

Adult Patients

The recommended starting dose is 40 mg once daily. If a daily dose of 40 mg does not achieve desired cholesterol levels, 80 mg once daily is recommended. In patients with a history of significant renal or hepatic dysfunction, a starting dose of 10 mg daily is recommended.

Pediatric Patients

Children (Ages 8 to 13 Years, Inclusive)

The recommended dose is 20 mg once daily in children 8 to 13 years of age. Doses greater than 20 mg have not been studied in this patient population.

Adolescents (Ages 14 to 18 Years)

The recommended starting dose is 40 mg once daily in adolescents 14 to 18 years of age. Doses greater than 40 mg have not been studied in this patient population.

Children and adolescents treated with pravastatin should be reevaluated in adulthood and appropriate changes made to their cholesterol-lowering regimen to achieve adult goals for LDL-C (see **INDICATIONS AND USAGE: Hyperlipidemia: Table 7: NCEP Treatment Guidelines**).

In patients taking immunosuppressive drugs such as cyclosporine (see **WARNINGS: Skeletal Muscle**) concomitantly with pravastatin, therapy should begin with 10 mg of pravastatin sodium once-a-day at bedtime and titration to higher doses should be done with caution. Most patients treated with this combination received a maximum pravastatin sodium dose of 20 mg/day.

Therefore, we needed to supplement drug knowledge databases with metadata about frequencies (Table 2) and needed to create custom rules to automate scheduling (Table 3). For the MMH prototype, primitive knowledge bases were constructed for medications commonly used in cystic fibrosis, a prototypical pediatric disease that requires complex medication regimens.

Table 2. Encoded dosing frequencies and metadata

Frequency	Translation	Doses per day	Spacing of doses	Comments
Q8h	Every 8 hours	3	8 hours	
TID	Three times a day (during waking hours)	3	Variable, but with doses spaced as evenly as possible during waking hours	
QHS	At bedtime	1		Some medications dosed QHS should be given in the morning for patients working a night shift
QAC	With each meal	3, but may be as needed	Variable	Dosing dependent on planned meal times

Table 3. Custom rules for automated scheduling

Rule	Example
“All doses of this medication have been placed on the schedule.”	A user schedules at 7:00 AM an antibiotic that is to be given every 8 hours. Doses are automatically added at 3:00 PM and 11:00 PM.
“These two doses are too close in time.”	The user above tries to move the 11:00 PM dose to 7:00 PM.
“[Medication] should be taken with every meal.”	The user is taking digestion enzymes and schedules a snack. A dose of digestion enzymes is automatically added to that time.

Form Factors, Functions, and User Interfaces

The primary goals of the CCT, derived from the Care Transitions Intervention [17], were to assist older patients with multimorbidities by (1) helping to create medication lists using diverse prescribing and dispensing data, (2) providing easy access to authoritative consumer health information, (3) helping identify discrepancies between their personal medication list and medication lists from clinicians, and (4) preparing for visits with clinicians. A tablet PC was chosen as the primary form factor for several reasons. We sought to accommodate mobility, since medications are often stored in multiple locations in the home [33]. Touch-screen input was chosen to minimize computer anxiety [34] and to decrease input problems associated with mapping horizontal input (mouse or track pad) to vertical visualization (computer screen) [35]. We incorporated bar-code input based on acceptance of this technology by adults with complex conditions in previous work [36]. The system was implemented on a platform consisting of a tablet PC (ThinkPad X60; Lenovo, Inc, Morrisville, NC, USA) running the Windows XP tablet operating system (Microsoft Corporation, Redmond, WA, USA) with a bar-code scanner (Bluetooth Cordless Hand Scanner Series 7; Socket Mobile, Newark, CA, USA) to scan bar-codes that may be available on medication labels. The Web-accessible user interface was developed using HyperText Markup Language, PHP Hypertext Preprocessor [37], and cascading style sheets (CSS), for high performance and stylistic consistency.

The goals of MMH were to provide interconnected Web and mobile applications that would allow (1) caregivers to create a medication schedule, (2) caregivers to select medications for which use should be prompted, (3) patients to receive medication prompts on a mobile device, (4) patients to confirm that a dose was taken, and (5) caregivers to track medication-taking behavior. For the Web component, MMH was constructed to operate on any standard Web browser. It was developed using PHP with asynchronous JavaScript and XML and Flash (Adobe Systems Incorporated, San Jose, CA, USA) components for the user interface. An alert/notification system used PHP 5 and communicated with mobile phone devices using the short message service (SMS) messaging protocol.

Results

User Feedback during Development

While a comprehensive discussion of the iterative development process for CCT [26] and MMH [25] is beyond this discussion, a number of findings from interim user feedback sessions were particularly notable.

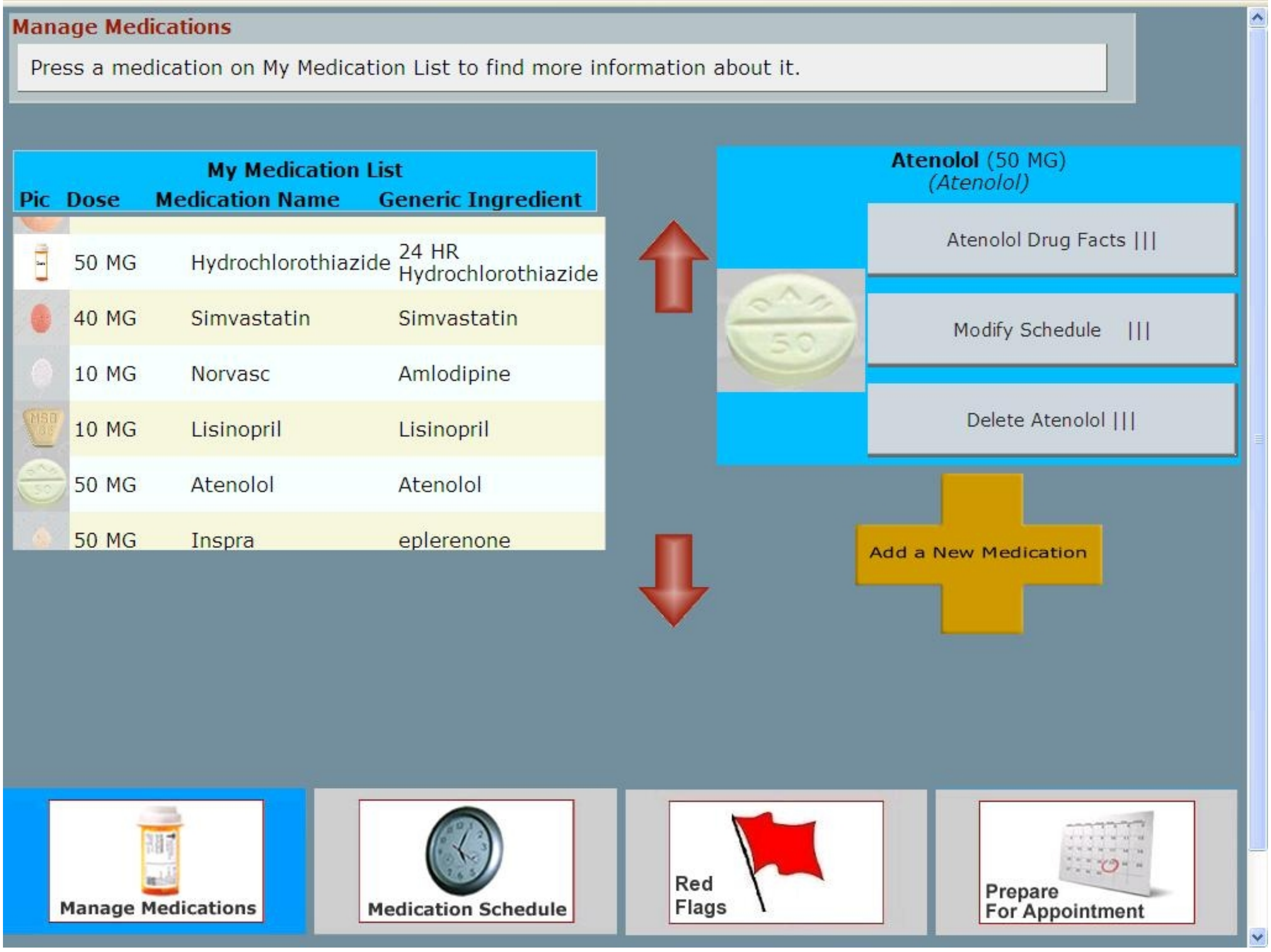
In general, the form factors proved to be appropriate. For CCT, older adults into their late 70s liked the concept of a mobile touch-screen tablet with large, readable text. In addition, they liked the concept of using a bar-code scanner to enter medication information from the prescription label, rather than entering the information by typing. However, the oldest users—those over 80—were averse to using any computerized interface for medication management, even when we took pains to refer to the tablet as an “appliance” rather than a computer. For MMH, children and their parents felt it was appropriate for school-aged children to carry and use a mobile device to assist in medication management. However, a proposal to embed the mobile device in a toy (such as a teddy bear) for younger school-aged children was not well received. Rather than making the device friendlier, younger children felt that carrying the toy would be stigmatizing.

Incorporating images of medications into the user interface was also greatly appreciated by both children and older adults. Both groups wanted medication images to be displayed on their respective Web interfaces. When the MMH mobile device sent medication prompts, the use of both text and medication images was greatly preferred to prompts with text alone. At the time of development, multimedia messaging service image messages were typically offered only on mobile plans at additional cost and were not integrated with SMS text messages, so the MMH prototype accommodated the desire for images by embedding URL links to images in SMS text messages.

Unlike younger users, older adults encountered unexpected difficulties with common user interface metaphors for navigation and actions. For navigation, older users preferred a dock of 4 key functions identified by an icon and text (Figure 4) instead of typical Web navigation structures with expanding top-horizontal and left-vertical action links. Within each core function, activities followed a linear “wizard” structure. Older adults also had problems with drag-and-drop actions when they had to schedule medications. In contrast to children and their parents, who found it very intuitive to drag medications from a personal medication list and drop them on to a calendar for scheduling in MMH (Figure 5), some older adults thought that dragging a medication from one list (for instance, one of their doctors’ medication lists) to their own list would corrupt the source (ie, would result in the medication being removed from the doctor’s list). Potential corruption of information maintained by medical professionals was a major point of concern.

As a whole, older adults consistently desired simplifications in CCT, even when this limited the application’s functionality. For instance, while it was expected that older adults would be interested in building their personal medication lists by referring to medication lists kept by their doctors, the older adults preferred *not* to be presented with multiple doctors’ medication lists. Instead, they wanted to view an aggregated list of all the medications that had been filled in the last year, from which they could select which ones were still being taken. Similarly, they found the user interface busy and confusing when their personal medication list was compared side by side with one of their doctor’s medication lists. In fact, they had little interest in ad hoc medication reconciliation at home. Instead, they felt that it was more appropriate for a medical professional to handle medication reconciliation at the time of appointments. To accommodate these preferences, the ultimate design of the CCT prototype could send a simple report of medication list discrepancies to a provider in advance of appointments.

Figure 4. Dock navigation for Colorado Care Tablet



Manage Medications

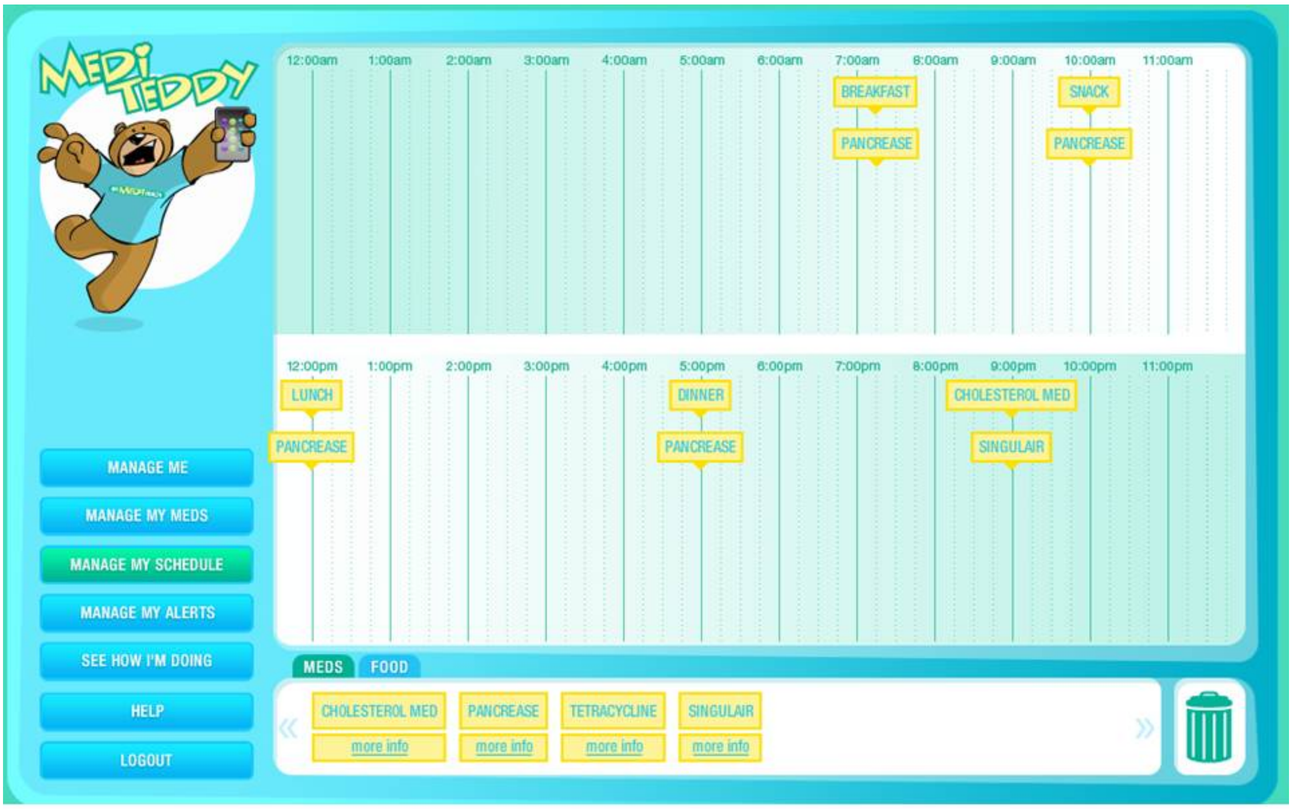
Press a medication on My Medication List to find more information about it.

My Medication List

Pic	Dose	Medication Name	Generic Ingredient
	50 MG	Hydrochlorothiazide	24 HR Hydrochlorothiazide
	40 MG	Simvastatin	Simvastatin
	10 MG	Norvasc	Amlodipine
	10 MG	Lisinopril	Lisinopril
	50 MG	Atenolol	Atenolol
	50 MG	Inspra	eplerenone

Atenolol (50 MG)
(Atenolol)

Figure 5. Drag-and-drop medication scheduling in MyMediHealth



Application

The development process resulted in the construction of working high-fidelity prototypes for user testing (see Multimedia Appendix 1 for CCT and Multimedia Appendix 2 for MMH). CCT was evaluated by 7 users in a final videotaped task-based usability test where participants created medication lists, a list of symptoms a participant should watch for during a care transition, and a memo in preparation for a doctor's visit. MMH was evaluated by 8 adult caregivers during a user study where participants created a medication list and developed a schedule. The MMH prototype was further subjected to evaluation by an online learning community [38]. Final testing of the high-fidelity prototypes confirmed the value of design choices made during iterative development, but also uncovered new practical issues in practice.

The ultimate design of CCT proved straightforward for users to navigate. Participants of all ages and computer skills were able to navigate CCT functions using the dock and linked wizards to build medication lists, to seek answers to common questions about individual medications, and to prepare for upcoming visits. However, users wanted CCT to answer additional questions about the medication list in general (whether there were drug interactions, whether it was dangerous to take "so many" medications, and whether some medications could be dropped). While the concept of the touch screen was well liked, many users found the touch screen insufficiently sensitive to their finger motions and required a stylus for certain tasks (such as using an on-screen keyboard). The bar-code scanner also performed inconsistently for bar-codes associated with prescription labels.

The functional MMH prototype was also well received. Overall, testers said that the scheduler prototype was generally easy to use, helpful for the family, and helpful for communication among family, school, and providers. Some members identified some important missing features, including support for dosing that varies by day or by degree of symptom, prompting about ideal locations on the schedule for a particular medication dose, support for as-needed dosing administration and dosing given less frequently than daily, and a more intuitive set of tools to create a medication list.

Other practical issues were uncovered for the mobile device. First, while the method of embedding a URL to provide both a message and a medication image was successful, it required at least 3 steps to manage an alert (receive a message alert, retrieve the message, and select the hyperlink). Second, since cell phone messages cannot be prioritized and are given bandwidth after cell phone audio calls are taken care of, there is the potential for message latency. While most messages are delivered within seconds of when the message is scheduled, some cell systems can hold a message for hours, or even not deliver the message at all. We experienced this latency intermittently during pilot testing. Although the system can be programmed to keep retrying a page until the patient acknowledges that they have either taken or not taken the medication, a page outside the correct timeframe may result in a missed dose.

We also tested integration of the two applications, using a scenario where the user used CCT to build a medication list and answer common questions about individual medications, then used MMH to set up a medication schedule. This scenario proved successful: medications entered in CCT were visible in MMH and vice versa, and it was possible to switch from one application to the other and back easily. However, due to differences in color and font, screen sizes, and user interface paradigms (touch screen vs point-and-click), further user interface development would be required to make transitions between the applications truly seamless.

Demonstrations of CCT and MMH, as well as source code for these applications (which is available under the Creative Commons license), are available from the Project HealthDesign website [39].

Discussion

Principal Findings

In this project, we succeeded in creating working prototypes of an interoperable PHR that accommodated fragmentation of care (using medication information from a variety of sources) and provided practical assistance in medication self-management. Employing a service-oriented architecture with shared components for data storage and information retrieval facilitated the development of complementary applications that could be tailored to different target users. Our user-centered design process allowed us to refine and simplify user interfaces to maximize usability even for relatively computer-naïve users.

Implications of the Findings

Our development effort has implications for informatics resources supporting medication self-management applications. We found the National Library of Medicine's online service-oriented RxNav utility for normalization of medication identities to be very useful. Since similar open-source services to provide medication images and consumer health information would also be useful, National Library of Medicine's recent work on MedlinePlus Connect is particularly welcome. Ideally, these resources should support both prescription and common nonprescription medications. Enrichment of standards and resources related to medication regimens would also be welcome. To provide robust assistance in scheduling complex medication regimens, two areas of development are needed: (1) ongoing refinement of standards for encoding medication instructions for prescriptions, and (2) more comprehensive, codified, machine-actionable resources for dosing recommendations. Information of particular interest is shown in Table 4.

Although the working prototypes were well received, our development process also highlighted practical issues regarding appropriate form factors and user interfaces for the respective target populations. For older adults with limited computer experience, use of common metaphors (such as drag-and-drop and hyperlinked navigation) may not be appropriate. Older adults are also willing to trade off navigation flexibility and functionality if it allows for a simplified user interface. Form factors such as tablet devices and bar-code scanners can accommodate their visual and dexterity needs, but they need further refinement to be used reliably and consistently. For children, mobile phones are an appropriate vehicle for prompts and reporting, but limitations in the ability to deliver images and recognition of latency issues need to be taken into account.

Table 4. Desirable encoding of machine-actionable dosing recommendations

- Which tablets may be crushed, which capsules may be sprinkled, and which may not
- Which medications may be administered by routes other than the strictly oral route
- How or whether to reschedule missed doses
- Whether a medication should be taken away from or with meals
- Whether two medications can be taken together

Comparison With the Literature

Our project builds upon previous work outlining the core needs for personally entered medication data [40,41] and reports of tethered [19] and untethered [42] PHRs for medication management. It builds upon the growing literature supporting the utility of mobile phones for prompting and recording medication taking [43-47]. However, it also shows that enhancing self-entered medication lists (such as MyMedicationList [42]) with personal information from diverse sources (pharmacy aggregators and electronic health records) is far more challenging for patients than simply providing patients a view of the medication list stored in a single tethered electronic health record.

Limitations

The primary limitation of this project is that it was not possible to provide patients with their own medication information for testing. The common PHR platform was standards-based but was designed for rapid prototyping, not secure storage. Linking the applications to secure platforms and presenting users with real medication information would allow for more realistic testing, both in the laboratory and in the field. It would also be useful to confirm our findings with larger numbers of participants in more geographically diverse settings.

Call for Further Development

The open-source code available from the Project HealthDesign site is intended to facilitate and catalyze future development based on the concepts presented here. With the development of highly functional commercial PHR platforms such as Dossia, Google Health, and Microsoft HealthVault, each with an expanding “ecosystem” of partners sharing data, it is possible to develop functional prototypes of CCT and MMH that can be deployed in the field. MMH is being expanded to provide a suite of tools for medication management in asthma, including a patient-generated pictographic medication list, text message-based medication reminders, a printable medication administration record, and an inhaler dose counter to help ensure that refills are requested in a timely fashion. CCT could also be redeployed on the iPad, which has a clean, simple form factor and robust touch-screen interface that has been enthusiastically received by consumers. Testing its utility in the context of care transitions would be particularly valuable.

Acknowledgments

This project was supported by Robert Wood Johnson Foundation as part of Project HealthDesign, grant #59880 for CCT and grant #59881 for MMH, with supplemental funding from the California HealthCare Foundation. Dr. Eric Coleman provided critical expertise in care transitions for the CCT. We greatly appreciate the assistance the National Library of Medicine (Stuart Nelson, Kelly Zeng, Olivier Bodenreider, and Terry Yoo) provided in helping us employ RxNav APIs and in explaining NLM's current and proposed resources for consumer health information and medication images. We also appreciate Micromedex for allowing us to use its consumer health information databases and drug image database for this project at no cost for this project. We also appreciate the support and guidance from the Project HealthDesign National Program Office, and the Vanderbilt Center for Better Health, for providing a rich environment for collaboration.

Conflicts of Interest

None declared

Multimedia Appendix 1

Colorado Care Tablet walkthrough with annotated screen shots

[PDF file (Adobe PDF File), 5531 KB - jmir_v13i3e45_app1.pdf]

Multimedia Appendix 2

MyMediHealth walkthrough with annotated screen shots.

[PDF file (Adobe PDF File), 414 KB - jmir_v13i3e45_app2.pdf]

References

1. Gandhi TK, Weingart SN, Borus J, Seger AC, Peterson J, Burdick E, et al. Adverse drug events in ambulatory care. N Engl J Med 2003 Apr 17;348(16):1556-1564 [FREE Full text] [doi: 10.1056/NEJMsa020703] [Medline: 12700376]
2. Committee on Identifiying and Preventing Medication Errors, Aspden P, Wolcott J, Bootman JL, Cronewett LR. Preventing Medication Errors: Quality Chasm Series. Washington, DC: National Academies Press; 2007.
3. [No authors listed]. JCAHO's 2006 National Patient Safety Goals: handoffs are biggest challenge. Hosp Peer Rev 2005 Jul;30(7):89-93. [Medline: 16004155]
4. Aronson JK. Medication errors resulting from the confusion of drug names. Expert Opin Drug Saf 2004 May;3(3):167-172. [doi: 10.1517/eods.3.3.167.31069] [Medline: 15155145]
5. Johnson KB, Butta JK, Donohue PK, Glenn DJ, Holtzman NA. Discharging patients with prescriptions instead of medications: sequelae in a teaching hospital. Pediatrics 1996 Apr;97(4):481-485. [Medline: 8632932]
6. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Ann Intern Med 2003 Feb 4;138(3):161-167 [FREE Full text] [Medline: 12558354]
7. Coleman EA, Berenson RA. Lost in transition: challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004 Oct 5;141(7):533-536. [Medline: 15466770]
8. Benson DS, Van Osdol W, Townes P. Quality ambulatory care: the role of the diagnostic and medication summary lists. QRB Qual Rev Bull 1988 Jun;14(6):192-197. [Medline: 3138607]
9. Gleason KM, Groszek JM, Sullivan C, Rooney D, Barnard C, Noskin GA. Reconciliation of discrepancies in medication histories and admission orders of newly hospitalized patients. Am J Health Syst Pharm 2004 Aug 15;61(16):1689-1695. [Medline: 15540481]
10. Kerzman H, Baron-Epel O, Toren O. What do discharged patients know about their medication? Patient Educ Couns 2005 Mar;56(3):276-282. [doi: 10.1016/j.pec.2004.02.019] [Medline: 15721969]
11. Safran C. The collaborative edge: patient empowerment for vulnerable populations. Int J Med Inform 2003 Mar;69(2-3):185-190. [Medline: 12810122]
12. Britten N, Stevenson FA, Barry CA, Barber N, Bradley CP. Misunderstandings in prescribing decisions in general practice: qualitative study. BMJ 2000 Feb 19;320(7233):484-488 [FREE Full text] [Medline: 10678863]

13. Personal Health Working Group. Markle Foundation. 2003. Personal Health Working Group: Final Report URL: http://www.providersedge.com/ehdocs/ehr_articles/The_Personal_Health_Working_Group_Final_Report.pdf [accessed 2011-03-15] [WebCite Cache ID 5xD52Zykh]
14. Commission on Systemic Interoperability. Ending the Document Game: Connecting and Transforming Your Healthcare Through Information Technology. Washington, DC: U.S. Government Printing Office; 2005.
15. Hanousek L. The medical and vaccination card for children--a document for the promotion of primary preventive care in pediatrics. Cesk Zdrav 1989 Jul;37(6-7):318-324. [Medline: 2791117]
16. Lakhani AD, Avery A, Gordon A, Tait N. Evaluation of a home based health record booklet. Arch Dis Child 1984 Nov;59(11):1076-1081. [Medline: 6508341]
17. Coleman EA, Parry C, Chalmers S, Min SJ. The care transitions intervention: results of a randomized controlled trial. Arch Intern Med 2006 Sep 25;166(17):1822-1828 [FREE Full text] [doi: 10.1001/archinte.166.17.1822] [Medline: 17000937]
18. Parry C, Coleman EA, Smith JD, Frank J, Kramer AM. The care transitions intervention: a patient-centered approach to ensuring effective transfers between sites of geriatric care. Home Health Care Serv Q 2003;22(3):1-17. [Medline: 14629081]
19. Schnipper JL, Gandhi TK, Wald JS, Grant RW, Poon EG, Volk LA, et al. Design and implementation of a web-based patient portal linked to an electronic health record designed to improve medication safety: the Patient Gateway medications module. Inform Prim Care 2008;16(2):147-155. [Medline: 18713531]
20. Nelson SJ, Zeng K, Bodenreider O. MyMedicationList: integrating personal medication records with resources. AMIA Annu Symp Proc 2008:1228. [Medline: 18998987]
21. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA 2002 Oct 16;288(15):1909-1914. [Medline: 12377092]
22. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. JAMA 2002 Oct 9;288(14):1775-1779. [Medline: 12365965]
23. Katz MG, Kripalani S, Weiss BD. Use of pictorial aids in medication instructions: a review of the literature. Am J Health Syst Pharm 2006 Dec 1;63(23):2391-2397. [doi: 10.2146/ajhp060162] [Medline: 17106013]
24. National Program Office, University of Wisconsin School of Nursing. Project HealthDesign.: Robert Wood Johnson Foundation; 2011. URL: http://www.projecthealthdesign.org/ [accessed 2011-03-20] [WebCite Cache ID 5xNin3jqz]
25. Slagle JM, Gordon JS, Harris CE, Davison CL, Culpepper DK, Scott P, et al. MyMediHealth - designing a next generation system for child-centered medication management. J Biomed Inform 2010 Oct;43(5 Suppl):S27-S31. [doi: 10.1016/j.jbi.2010.06.006] [Medline: 20937481]
26. Siek KA, Ross SE, Khan DU, Haverhals LM, Cali SR, Meyers J. Colorado Care Tablet: the design of an interoperable Personal Health Application to help older adults with multimorbidity manage their medications. J Biomed Inform 2010 Oct;43(5 Suppl):S22-S26. [doi: 10.1016/j.jbi.2010.05.007] [Medline: 20937480]
27. Sujansky WV, Faus SA, Stone E, Brennan PF. A method to implement fine-grained access control for personal health records through standard relational database queries. J Biomed Inform 2010 Oct;43(5 Suppl):S46-S50. [doi: 10.1016/j.jbi.2010.08.001] [Medline: 20696276]
28. Wixon D. Evaluating usability methods: why the current literature fails the practitioner. interactions 2003;10(4):28-34. [doi: 10.1145/838830.838870]
29. Kjeldskov J, Skov MB, Stage J. Instant data analysis: conducting usability evaluations in a day. In: Proceedings of the Third Nordic Conference on Human-Computer Interaction. New York: ACM Press; 2004 Presented at: NordiCHI 2004; October 23-27, 2004; Tampere, Finland URL: http://portal.acm.org/citation.cfm?id=1028014&picked=prox&cfid=28925924&cftoken=14962541 [doi: 10.1145/1028014.1028050]
30. Zeng K, Bodenreider O, Kilbourne J, Nelson S. RxNav: a web service for standard drug information. AMIA Annu Symp Proc 2006:1156. [Medline: 17238775]
31. Lister Hill National Center for Biomedical Communications. RxNav: RxNorm API.: National Library of Medicine; 2011. URL: http://mor.nlm.nih.gov/download/rxnav/RxNormAPI.html [accessed 2011-03-22] [WebCite Cache ID 5xNj3oYg6]
32. National Council for Prescription Drug Programs. Structured and Codified Sig Format: Implementation Guide, version 1.0. Scottsdale, AZ: National Council for Prescription Drug Programs; 2006 Jun. URL: http://www.ncpdp.org/pdf/SIG_standard_imp_guide_2006-06.pdf [accessed 2011-03-22] [WebCite Cache ID 5xNjQ5SW6]
33. Moen A, Brennan PF. Health@Home: the work of health information management in the household (HIMH): implications for consumer health informatics (CHI) innovations. J Am Med Inform Assoc 2005;12(6):648-656. [doi: 10.1197/jamia.M1758] [Medline: 16049230]
34. Main DS, Quintela J, Araya-Guerra R, Holcomb S, Pace WD. Exploring patient reactions to pen-tablet computers: a report from CaReNet. Ann Fam Med 2004;2(5):421-424 [FREE Full text] [Medline: 15506574]
35. Charness N, Bosman E, Elliott R. Senior-friendly input devices: is the pen mightier than the mouse? 1995 Aug 15 Presented at: 103rd Annual Convention of the American Psychological Association; August 15, 1995; New York, NY.
36. Connelly KH, Rogers Y, Siek KA, Jones J, Kraus MA, Perkins S. Designing a PDA interface for dialysis patients to monitor diet in their everyday life. In: Salvendy G, editor. HCI 2005 CD - Volumes 1-11. Mahwah, NJ: Lawrence Erlbaum Associates; 2005.

37. PHP Group. PHP Hypertext Processor. 2011. URL: http://www.php.net/ [accessed 2011-03-22] [WebCite Cache ID 5xNkSixn9]
38. Vanderbilt University. eLab. Nashville, TN: Owen Graduate School of Management and Vanderbilt University; 2006. URL: http://elab.vanderbilt.edu/ [accessed 2011-03-22] [WebCite Cache ID 5xNkfe7n5]
39. National Program Office, University of Wisconsin School of Nursing. Project HealthDesign: 2006-2008 Work Products.: Robert Wood Johnson Foundation; 2010. URL: http://www.projecthealthdesign.org/resources/pastprojects-products-open [accessed 2011-03-21] [WebCite Cache ID 5xNri8dKg]
40. Kim MI, Johnson KB. Personal health records: evaluation of functionality and utility. J Am Med Inform Assoc 2002;9(2):171-180. [Medline: 11861632]
41. Kim MI, Johnson KB. Patient entry of information: evaluation of user interfaces. J Med Internet Res 2004 May 14;6(2):e13 [FREE Full text] [doi: 10.2196/jmir.6.2.e13] [Medline: 15249262]
42. Zeng K, Bodenreider O, Nelson SJ. Design and implementation of a personal medication record-MyMedicationList. AMIA Annu Symp Proc 2008:844-848. [Medline: 18998962]
43. Kew S. Text messaging: an innovative method of data collection in medical research. BMC Res Notes 2010;3:342 [FREE Full text] [doi: 10.1186/1756-0500-3-342] [Medline: 21172018]
44. Ollivier L, Romand O, Marimoutou C, Michel R, Pognant C, Todesco A, et al. Use of short message service (SMS) to improve malaria chemoprophylaxis compliance after returning from a malaria endemic area. Malar J 2009;8:236 [FREE Full text] [doi: 10.1186/1475-2875-8-236] [Medline: 19852811]
45. Shet A, Arumugam K, Rodrigues R, Rajagopalan N, Shubha K, Raj T, et al. Designing a mobile phone-based intervention to promote adherence to antiretroviral therapy in South India. AIDS Behav 2010 Jun;14(3):716-720. [doi: 10.1007/s10461-009-9658-3] [Medline: 20054634]
46. Mao Y, Zhang Y, Zhai S. Mobile phone text messaging for pharmaceutical care in a hospital in China. J Telemed Telecare 2008;14(8):410-414. [doi: 10.1258/jtt.2008.080406] [Medline: 19047450]
47. Cocosila M, Coursaris C, Yuan Y. M-healthcare for patient self-management: a case for diabetics. Int J Electron Healthc 2004;1(2):221-241. [Medline: 18048222]

Abbreviations

CCT: Colorado Care Tablet
MMH: MyMediHealth
NDC: National Drug Code
PHR: personal health record
SMS: short message service

Edited by G Eysenbach; submitted 22.03.11; peer-reviewed by B Massoudi; comments to author 18.04.11; revised version received 25.05.11; accepted 26.05.11; published 12.07.11.

Please cite as:
Ross SE, Johnson KB, Siek KA, Gordon JS, Khan DU, Haverhals LM
Two Complementary Personal Medication Management Applications Developed on a Common Platform: Case Report
J Med Internet Res 2011;13(3):e45
URL: http://www.jmir.org/2011/3/e45/
doi:10.2196/jmir.1815
PMID:21749966

Original Paper

Development and Validation of the Online Social Support for Smokers Scale

Amanda L Graham[1,2], PhD; George D Papandonatos[3], PhD; Hakmook Kang[3], PhD; Jose L Moreno[1], MA; David B Abrams[1,2,4], PhD

[1]Schroeder Institute for Tobacco Research and Policy Studies, American Legacy Foundation, Washington, DC, United States
[2]Cancer Control Program, Georgetown University Medical Center/Lombardi Comprehensive Cancer Center, Washington, DC, United States
[3]Center for Statistical Sciences, Brown University, Providence, RI, United States
[4]The Johns Hopkins Bloomberg School of Public Health, Department of Health, Behavior and Society, Baltimore, MD, United States

Corresponding Author:
Amanda L Graham, PhD
Schroeder Institute for Tobacco Research and Policy Studies
American Legacy Foundation
1724 Massachusetts Avenue, NW
Washington, DC, 20036
United States
Phone: 1 202 454 5938
Fax: 1 202 454 5785
Email: agraham@legacyforhealth.org

Abstract

Background: Social networks play an important role in smoking. Provision of social support during cessation is a cornerstone of treatment. Online social networks for cessation are ubiquitous and represent a promising modality for smokers to receive and provide the support necessary for cessation. There are no existing measures specific to online social support for smoking cessation.

Objective: The objective was to develop a measure of social support to be used in online smoking cessation treatment research.

Methods: Initial items for the Online Social Support for Smokers Scale (OS4) were based on existing theory and scales delineated in various taxonomies. Preliminary field analysis (N = 73) was conducted on 23 initial items to optimize the scale. Further development was conducted on a refined 15-item scale in the context of a large randomized trial of Internet and telephone cessation treatment with follow-ups at 3, 6, 12, and 18 months. In all, 1326 participants were randomized to an enhanced Internet arm that included a large online social network; psychometric analyses employed 3-month follow-up data from those reporting use of the enhanced Internet intervention at least once (n = 873). Items were subjected to a factor analysis, and the internal consistency reliability of the scale was examined along with construct and criterion validity. Other measures used in the study included demographics, nicotine dependence, partner support for cessation, general social support, social integration, stress, depression, health status, online community use, Internet use behaviors, intervention satisfaction, and 30-day point prevalence abstinence.

Results: The final 12-item OS4 scale demonstrated high internal consistency reliability (Cronbach alphas .86-.89) across demographic and smoking strata of interest. The OS4 also demonstrated good construct and criterion validity, with the directionality of the observed associations providing support for most a priori hypotheses. Significant Pearson correlations were observed between the OS4 and the Partner Interaction Questionnaire (PIQ) Positive subscale ($\rho = .24$, $P < .001$). As hypothesized, participants with the highest OS4 scores were more likely to have actively participated in the enhanced Internet community and to have high levels of satisfaction with the enhanced Internet intervention. In logistic regression analyses, the OS4 was highly predictive of 30-day point-prevalence abstinence at 6, 12, and 18 months (all P values <.001). The odds of abstinence at 6 months rose by 48% for each standard unit increase in online social support (95% confidence interval [CI] 1.17 - 1.71), dropping only slightly to 37% at 12 and 18 months (95% CI 1.17 - 1.59).

Conclusions: The OS4 is a brief, reliable, and valid instrument for measuring online social support for smoking cessation. Results should be replicated and extended, but this study suggests the OS4 can be used to advance theory, understand mechanisms, and potentially help to improve the tailoring of Internet-based smoking cessation treatments. It can also inspire development of similar measures for other online health-related intervention research.

Trial registration: Clinicaltrials.gov #NCT00282009; http://clinicaltrials.gov/ct2/show/NCT00282009 (Archived by WebCite at http://www.webcitation.org/60XNj3xM6)

(J Med Internet Res 2011;13(3):e69) doi:10.2196/jmir.1801

KEYWORDS

Psychometrics; social support; Internet; smoking cessation

Introduction

It has long been recognized that social support and social connections play an important role in smoking initiation, maintenance, quitting, and relapse. High levels of social support have been linked to better cessation outcomes in correlational and epidemiological studies [1-4] and low levels of support (eg, negative behaviors from a spouse) are a barrier to cessation and maintenance [2,4,5]. Recent analyses by Christakis and colleagues [6] demonstrated that social networks play a powerful role in smoking cessation in that cessation propagates more rapidly among smokers in the proximal social network of a quitter. Given the available evidence, the provision of social support during the process of quitting is one of the cornerstones of evidence-based cessation treatment [7].

Online social networks for smoking cessation have become ubiquitous and, thus, may represent a promising modality for smokers to both receive and provide the kind of support necessary for cessation and relapse prevention. Through Internet-based social networks, smokers have round-the-clock access to thousands of other individuals who are actively quitting smoking, struggling to maintain abstinence, or celebrating various milestones of abstinence. Access in real time to a diverse mix of individuals in all stages of the cessation journey is a unique aspect of online social networks. No other cessation treatment modality provides an ongoing opportunity for current and former smokers to interact and influence each other. In addition, smokers benefit not only from active interactions with other network members, but also from various passive sources of social influence and social support. Smokers can establish personal connections with other network members, or can browse ("lurk") the messages and profiles of others. These kinds of active and passive interactions may influence an individual's motivation to quit, reinforce the undesirability of smoking, assist in buffering cessation-related stressors, enhance coping skills, and provide suggestions for eliminating smoking cues.

To date there have been few published studies of online social networks for cessation. Several studies have described the frequency, intensity, and nature of interactions among online social network members [8-12]. Other studies have examined the association of participation in online communities with cessation outcomes [11,13,14]. These associations appear to be relatively robust, with higher levels of social exchanges (eg, messages, forum posts, and blog posts) and social connectedness (eg, number of buddies and number of people sending messages to and receiving messages from) associated with higher likelihood of abstinence. While these associations are compelling, we know of no studies that have examined whether interactions in an online social network for cessation do, in fact, lead to changes in perceived social support. If observational findings are to be harnessed in interventions that attempt to manipulate social support to improve cessation outcomes, a measure of perceived social support from online social networks is needed, both as a manipulation check and also as a measure of a potentially important mediating mechanism. To our knowledge, there are no existing measures of online social support specific to smoking cessation.

Thus, the purpose of the present study was to develop a brief, reliable, and valid measure of social support for online smoking cessation research that could be used with minimal burden on respondents. Following item generation, we conducted a series of psychometric analyses to examine the performance of individual items and to optimize the scale. We subjected items to a factor analysis and examined the quality of the scale as reflected by the internal consistency reliability coefficient, Cronbach alpha. Finally, we examined the construct and criterion validity of the scale, relying on published studies to guide our hypotheses where possible.

Construct validity is the extent to which a psychometric scale, as operationalized in a particular study, actually measures the theoretical construct that it purports to measure. This requires evidence of similarity with measures known to be related to the underlying construct (convergent validity), and lack of similarity with measures of theoretically distinct constructs (discriminant validity) [15]. With regard to convergent validity, we hypothesized that the scale would be positively correlated with other measures of social support, including smoking-specific social support from a partner or friend [2,16], a general multidimensional measure of support [17], and measures of social integration [18] and of frequency of online communications via social media. With regard to discriminant validity, we hypothesized that our measure of online perceived support would show no association with smoking variables (nicotine dependence, age of first smoking, daily smoking rate, number of quit attempts in the past year, desire to quit, or confidence in quitting), psychosocial variables (stress, current depressive symptoms, and history of anxiety/depression), perceived health status, or duration and frequency of Internet use.

Criterion validity examines the degree to which test scores on a particular psychometric scale correlate well with one or more criteria taken as representative of the construct. This requires evidence of similarity with criteria obtained at approximately the same time (concurrent validity), as well as following test administration (predictive validity) [15]. With regard to concurrent validity, we hypothesized that higher scores on our measure of online social support would distinguish those with higher levels of satisfaction with the website; those who participated in the online community more intensively as indicated by self-report data (amount of perceived "help" given to and received from other community members); and those who participated in the online community more intensively as indicated by automated tracking of online activities (ie, use of any community features, internal messages sent to other members, and buddies in the online community). With regard to predictive validity, we hypothesized that participants with higher scores of online social support 3 months after beginning their participation in an online cessation intervention would be more likely to be abstinent at each of the subsequent follow-up assessments.

Methods

Overview

Development of the Online Social Support for Smokers Scale (OS4) was conducted in the context of a randomized controlled trial (Clinicaltrials.gov #NCT00282009) of Internet cessation treatment alone and in conjunction with proactive telephone counseling [19,20]. We developed the OS4 in order to examine theory-driven hypotheses about the role of online social support as a potential mediating mechanism of treatment outcome. Briefly, the trial randomized 2005 current smokers to one of three treatment conditions: (1) enhanced Internet (n = 651), (2) enhanced Internet plus proactive telephone counseling (n = 675), or (3) a static, information only basic Internet comparison condition (n = 679). Participants were followed at 3, 6, 12, and 18 months postrandomization.

For the enhanced Internet condition, participants were provided free access to QuitNet.com, an interactive, commercial cessation website that provides evidence-based cessation treatment in accordance with national guidelines [7]. Described elsewhere [11,19], QuitNet provides (1) advice to quit; (2) assistance in setting a quit date; (3) assessment of motivation, smoking history, demographics, and nicotine dependence; (4) individually-tailored information; (5) problem solving/skills training content; (6) tailored assistance in using pharmacotherapies approved by the US Food and Drug Administration; and (7) social support within its large online social network [10]. For over 10 years, QuitNet has enrolled individuals into a network of current and former smokers seeking to quit or stay quit and provided multiple mechanisms of social support and social influence. QuitNet's community features allow for multiple forms of social support and social influence. Communication can occur through asynchronous channels, such as private internal email ("Q-Mail") or one-to-many messaging in the threaded forums, as well as synchronous channels such as chat rooms. Users can self-affiliate into clubs that are essentially user-initiated mini-sites complete with a dedicated forum. "Buddy lists" allow individuals to keep track of their friends. Social influence regarding cessation is conveyed through profile pages, journals (similar to a blog), anniversary lists, and testimonials. Users are encouraged to publically share their quit dates, which are set through a "wizard" tool, and users are prompted for updates at each login. QuitNet maintains a complete transactional history of all events, including communications that occur throughout the site. Active events (eg, sending internal email or posting a public message) and passive actions (eg, reading messages or viewing another individual's profile) are logged into a relational database.

Questionnaire Development

Development of the OS4 began with a comprehensive review of the literature to gather existing measures of social support specific to smoking cessation and more broadly related to health behavior change. We also gathered unpublished items being used in ongoing studies from tobacco experts. Measures of social support often distinguish between socially supportive functions that are perceived to be available (*perceived support*) and functions that were recently provided (*received support*) [21]. Given reports that a small percentage of people actively participate in online networks for cessation [11,13,14] and that a much larger number of people browse/lurk in the community, we included items that addressed both perceived and received support in order to account for the possibility that the potential availability of support as needed is as important as support actually received.

Initial items were based on existing scales but were adapted to reflect the specific social context of Internet interventions. Items covered each of the five domains of social support delineated in various taxonomies [22-24]: (1) emotional or esteem support, which refers to the availability of people to talk to about one's problems who can provide indications of caring and acceptance, empathy, reassurance, liking, and respect; (2) instrumental or tangible support, which refers to the perceived availability of material aid or practical support; (3) informational support, which refers to advice or guidance to solve a problem; (4) companionship or belonging support, which refers to perceived social companionship or social integration; and (5) validation or appraisal support, which provides feedback or social comparison about the normative nature of an individual's behaviors or feelings and their relative status in the population. A total of 23 items were generated to provide adequate redundancy within each domain. Items were written at a sixth grade reading level.

Preliminary field testing of the OS4 items was conducted within the QuitNet online social network. Using the internal QuitNet mail system, an invitation to complete an online survey was sent to active community members who had logged into the system at least 10 times. This criterion was selected to ensure that respondents had adequate experience within the community to knowledgeably respond to the relevance or appropriateness of the items. Participants were asked to respond to each item using a 5-point scale where 1 = definitely false, 2 = probably false, 3 = no opinion, 4 = probably true, and 5 = definitely true. Next to each item, participants could also enter comments about the relevance or appropriateness of the item or suggest alternate wording.

A total of 85 people visited the survey link between June 6, 2005, and June 10, 2005; of these, 73 completed the survey. We examined the mean, standard deviation, and range of the original 23 items as well as feedback provided about specific items. A total of 13 items with little variability and/or wording that participants indicated was unclear or tangential to their experience in the community were dropped. Based on participant feedback and expert review, modifications were made to the 10 remaining items to enhance their clarity and maximize their relevance, and 5 new items were added. As shown in Table 1, the scale was composed of 15 items (3 items in each of the 5 domains mentioned previously).

Table 1. Online Social Support Scale for Smokers: Original scale items[a]

Q1. I connected with other people on QuitNet on topics other than smoking.

Q2. I never posted messages on QuitNet.[b]

Q3. I felt comfortable sharing private or personal thoughts with other members of QuitNet in the public forums.

Q4. I felt comfortable sharing private or personal thoughts through Q-Mail to individual members of QuitNet.

Q5. By giving advice to other members of QuitNet, my own efforts in quitting were reinforced.

Q6. Being anonymous made it easier to share personal information with people on QuitNet.

Q7. Using QuitNet helped me cope with cravings.

Q8. I got advice and support on QuitNet that I could not find anywhere else.

Q9. It was comforting to know that I wasn't alone in the struggle to get and stay quit.

Q10. The fact that QuitNet is available whenever I need it, night or day, was important to me.

Q11. I felt supported and encouraged by other QuitNet members.

Q12. Advice and support from people in different stages of quitting was helpful to me.

Q13. I received negative or critical comments from other QuitNet members.[b]

Q14. I received some bad information or advice from someone on QuitNet.[b]

Q15. Being in a different time zone from other members made it difficult to get the support I needed.[b]

[a] Scoring structure: 1 = disagree a lot, 2 = disagree a little, 3 = agree a little, 4 = agree a lot

[b] reverse-scored item

Procedure

At each follow-up assessment, participants randomized to the enhanced Internet and enhanced Internet plus telephone treatment arms were asked how many times they had used the QuitNet website during the follow-up period. Those who had used the website at least once were administered the refined 15-item OS4. The psychometric analyses reported here used data from the 3-month follow-up, since website utilization is highest immediately after registration and tails off within 3 months for the majority of new members. Of the 990 participants who were reached at the 3-month follow-up, 873 participants reported using the QuitNet website at least once and completed the OS4; these participants were used as a validation sample. The remaining 117 participants reported no use of the website during the first 3 months of the study; this sample was used to examine generalizability.

Measures

Demographics and Smoking History

Demographic information collected at baseline included age, gender, education, race, ethnicity, household income, marital status, and employment status. The smoking history questionnaire assessed age of first smoking, daily smoking rate, and number of intentional quit attempts in the past year. Desire and confidence in quitting were each measured on a scale of 1 to 10 where 1 = not at all and 10 = very much.

Nicotine Dependence

Nicotine dependence was measured using the Fagerstrom Test for Nicotine Dependence (FTND) [25], a 6-item measure of dependence considered a standard instrument in the field. Greater scores indicate higher levels of dependence. Internal consistency reliability at 3 months was moderate (Cronbach alpha = .69).

Partner Support for Cessation

The Partner Interaction Questionnaire (PIQ) [2] is the most commonly used measure of spouse/partner support related to cessation. We administered a modified version of the PIQ that measures the receipt of specific behaviors from the person who follows a participant's efforts to quit smoking most closely, not just a spouse/partner [26,27]. The modified version used a 5-point Likert scale to assess how frequently the participant's support person exhibited 3 positive and 3 negative behaviors [16]. Positive items were "express pleasure at your efforts to quit," "congratulate you for your decision to quit smoking," and "express confidence in your ability to quit/remain quit." Negative items were "mention being bothered by smoke," "ask you to quit smoking," and "criticize your smoking." Response options were 0 = never, 1 = almost never, 2 = sometimes, 3 = fairly often, and 4 = very often. For the PIQ scale, internal consistency reliability at 3 months reached .84 for the Positive subscale, .85 for the Negative subscale, and .74 for the difference of the two.

General Social Support

The 12-item version of the Interpersonal Support Evaluation List (ISEL) [17] was used to assess the perceived availability of social resources. The ISEL is composed of three subscales that assess the perceived availability of distinct functions of social support: the Appraisal subscale measures the perceived availability of someone to talk to about one's problems; the Belonging subscale assesses the perceived availability of people with whom to engage in activities; the Tangible subscale measures the perceived availability of instrumental support or material aid. A total measure of perceived support can also be calculated. Respondents indicated their level of agreement with each statement using a 4-point scale: 0 = disagree a lot, 1 = disagree a little, 2 = agree a little, and 3 = agree a lot. At the 3-month follow-up, internal consistency reliability reached .82 for Appraisal, .76 for Belonging, .67 for Tangible, and .87 for the overall scale.

Social Integration

Designed as a measure of social integration, the Social Network Index [18] assesses participation in twelve types of social relationships (eg, spouse, friend, workmate, volunteer). Social network diversity is calculated by assigning one point for each type of relationship (possible score of 12) for which respondents indicate communication at least every 2 weeks. The total number of network members is calculated as the total number of persons with whom they speak at least once every 2 weeks.

Perceived Stress

The 4-item Perceived Stress Scale (PSS) [28] assesses the degree to which participants find their lives to be unpredictable and uncontrollable. Each item is rated on a 5-point Likert scale to indicate how frequently the individual has felt a particular way during the past month. Response options were 0 = never, 1 = almost never, 2 = sometimes, 3 = fairly often, and 4 = very often. Internal consistency reliability was .82 at 3 months.

Depression

Symptoms of current depression were measured using the 10-item Center for Epidemiological Studies Depression Scale (CES-D) [29]. Participants indicated the frequency of occurrence of each symptom during the past week (less than a day, 1-2 days, 3-4 days, and 5-7 days). Internal consistency reliability was .82 at 3 months. Participants also reported past year diagnosis of nervous trouble or depression (yes/no).

Perceived Health Status

Using the item from the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36), participants rated their current health status on a 5-point scale from 1 (excellent) to 5 (poor) [30].

Internet Use Behavior

Participants reported the number of years they had used the Internet and the frequency of Internet use. Participants also rated the frequency of online communications other than email (eg, blogging and use of chat rooms): 1 = never, 2 = less often, 3 = every few weeks, 4 = 1 to 2 days a week, 5 = 3 to 5 days a week, 6 = about once a day, 7 = several times a day.

Intervention Satisfaction

At each follow-up, participants rated their overall satisfaction with the QuitNet website, its perceived helpfulness, and how well the website met their expectations on a scale from 1 to 10 where 1 = not at all and 10 = very much. In addition, participants indicated how much help they had provided to other QuitNet community members and how much help they had received from other QuitNet community members on a 4-point scale where 1 = none, 2 = a little, 3 = some, and 4 = a lot.

Online Community Participation Metrics

At the 3-month follow-up, the following selected metrics of active participation in the QuitNet community were extracted: (1) any use of community features; (2) number of Q-Mails sent to other members; and (3) number of buddies designated. These particular metrics were selected based on their expected association with a measure of perceived support.

Smoking Cessation

At each follow-up assessment, participants self-reported smoking status over the past 30 days which was used as the primary outcome for the parent trial [20].

Data Analysis

Data analysis was conducted in multiple phases. First, to examine generalizability, we compared the validation sample to the 117 participants who had not used the website on a range of baseline demographic, smoking, and psychosocial variables. Two-sample *t* tests were used for continuous variables, Chi-square tests were used for categorical variables, and Poisson regression was used for count data. For categorical variables with small cell frequencies (<5 subjects per cell), significance levels were computed using Fisher's exact test. Sensitivity of *t* test findings to skewness in continuous variables was assessed by Wilcoxon signed-rank tests.

Next, we performed a factor analysis on the OS4 of the interitem covariance matrix using maximum likelihood estimation followed by a varimax rotation. The loadings from the resulting 2-factor solution were used to construct a low-dimensional representation of the interitem covariance matrix known as a biplot [31,32]. By representing each item by a directed arrow, biplots can be used to visually examine individual item characteristics and between-item relationships in cases where the number of items is too large to allow such information to be easily discerned from the interitem covariance matrix itself. Arrow orientation depends on item loadings, with arrows in the first quadrant having positive loadings on both the first and second factors. Arrows pointing in the same direction indicate items that are positively correlated; arrows pointing in the opposite direction indicate negatively correlated items; uncorrelated items have arrows that appear at right angles. Arrow length is related to item variability, with long arrows reflecting highly variable and, hence, more informative items. Long arrows that overlap are indicative of items that may be informative on their own but potentially contribute redundant information. Short arrows are indicative of items with little between-subject variability in the space of the first two factors, the implication being that they can be dropped from the measurement scale with little loss of information about the underlying construct. A biplot suggestive of a single factor solution would be one in which all arrows lie in the positive quadrant after rotation. When the arrows not only point in the same direction but are also approximately equal in length, then all items are informative to a similar degree and a total score should provide a good approximation to the factor score from a single factor solution.

After dropping 3 uninformative items, we repeated the factor analysis, comparing the 1-factor and 2-factor solutions in terms of proportion of variance explained among the remaining 12 items. Once a single-factor solution was established, we used Cronbach alpha to determine whether deletion of partially overlapping items would adversely affect overall scale reliability, with a value of .80 set as the lower acceptable bound [33]. To ensure that use of the OS4 is appropriate for various subgroups of smokers, we examined changes in reliability across population strata defined by gender, race/ethnicity (non-Hispanic white vs other categories of race/ethnicity), marital status, education, and income. Finally, we examined the correlation between the actual factor scores and the total score obtained by simply adding the items loading on a particular factor. When the two scores are highly correlated, there is little information lost when calculating subject-specific measures of online social support for quitting smoking by weighing each scale item equally instead of using the optimal item weights suggested by the factor analysis itself.

Having finalized our choice of weights for the items used to measure online social support, we proceeded to examine construct validity. Convergent validity was assessed using cross-sectional associations between the OS4 and other theory-driven measures of social support. Discriminant validity was assessed using cross-sectional associations between the OS4 and theoretically distinct measures (ie, smoking variables, stress and depression, perceived health status, and Internet use variables). The association between variables of interest and OS4 was measured using Pearson correlations for continuous variables and polyserial correlations for binary variables. Correlations were corrected for attenuation due to measurement error by inflating them by the inverse square root of the reliability coefficients of the respective psychometric scales in our validation sample. In addition, significance levels were corrected for finite sample uncertainty in estimation of the reliability coefficients themselves [34]. No correction was made for the lack of perfect reliability in the OS4 itself, as one must use the OS4 scale as it exists rather than in terms of its unknown “true score” [35].

To assess concurrent criterion validity, we compared the three metrics of intervention satisfaction (overall satisfaction, perceived helpfulness, the degree to which the site met expectations) and metrics of active participation in the community (any community use, sent any Q-Mail, designated any buddies, and gave or received some or a lot of help to or from other members) at 3 months across OS4 quartiles using normal linear regression for continuous variables and logistic regression for binary variables. Statistically significant between-group differences in means or proportions were taken as evidence that the OS4 does indeed have the discriminatory power one would have expected it to demonstrate on such measures. Finally, predictive criterion validity was assessed in terms of the ability of the OS4 at 3 months to predict self-reported 30-day point prevalence abstinence at 6, 12, and 18 months in a logistic regression model that controlled for the effects of treatment assignment (enhanced Internet vs enhanced Internet + phone).

Results

Participants

Detailed characteristics of all trial participants have been reported elsewhere [20]: There were no differences on any of the demographic, smoking, or psychosocial variables across treatment arms. Briefly, among the 873 participants in the validation sample, mean age was 36.5 years (SD 11.1) and 52.2% (456) were female. Most were white (786 or 88%), had completed at least some college (688 or 79%), were employed full-time (605 or 69.3%), and were long-term and frequent Internet users: 81.4% (709) had used the Internet for more than 5 years, and 79.4% (693) used the Internet several times a day. Participants smoked an average of 20 cigarettes per day (SD 9.4), reported a high level of desire to quit (mean 9.04, SD 1.3) and slightly lower confidence in quitting (mean 6.26, SD 2.2), and had made an average of 3.4 (SD 10.2) quit attempts in the past year.

To examine generalizability, we compared the validation sample with the 117 participants who did not use the enhanced Internet treatment during the first 3 months of the study. The validation sample had a higher percentage of women (52.2% or 456/873 vs 41.0% or 48/117, P = .02), reported lower levels of nicotine dependence (mean 5.1, SD 2.3 vs mean 4.5, SD 2.5, P = .02), and was more likely to use the Internet on a daily basis to communicate via blogs or instant messaging (40.8% or 356/873 vs 33.3% or 39/117, P = .009). There were no differences on any of the other variables examined.

Factor Analysis of the Covariance Matrix

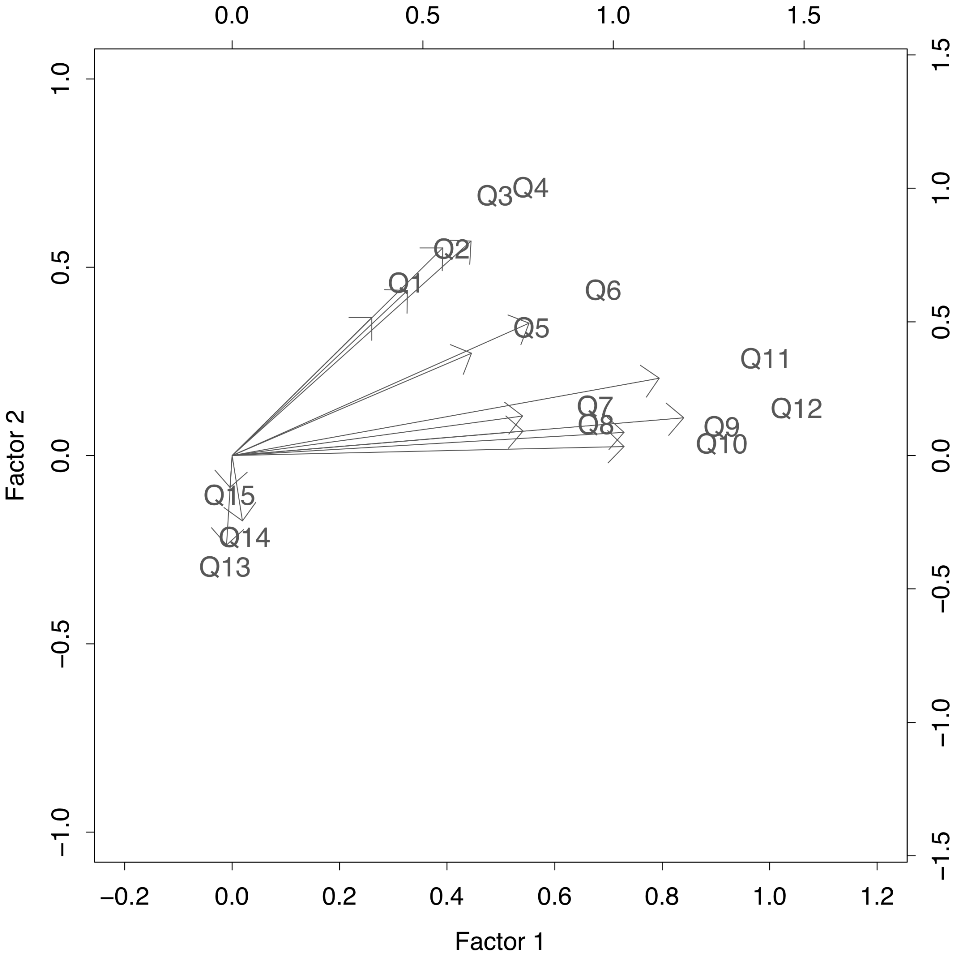
Results from the factor analysis of the items in Table 1 are presented in Figure 1 in biplot form (see Multimedia Appendix 1 for interitem correlation matrix). Factor analysis of the 15 items revealed a homogeneous cluster of 12 items (Q1 through Q12) with strong intraclass correlation (ICC = .36) and high reliability (Cronbach alpha = .87). The 3 remaining items (Q13 through Q15) appeared relatively uninformative. Further examination of their frequency distribution indicated that Q13 and Q14 showed minimal between-subject variation, with 84% and 90% of the sample respectively endorsing the highest (ie, least negative) category. In contrast, Q15 showed larger between-subject variation, but seemed to lie in a dimension other than that described by the first two factors depicted in Figure 1. Therefore, we decided to drop these 3 items from further consideration.

The biplot also suggested that the remaining items could be grouped into 2 highly correlated subsets (Q1 through Q4 and Q7 through Q12), with Q5 and Q6 equidistant from them. Examination of factor loadings from a 2–factor solution of Q1 through Q12 indicated that Q7 through Q12 defined a single factor explaining 23.6% of the variance, while Q1 through Q4 defined a second factor explaining 21.5% of the variance; Q5 and Q6 had approximately equal loadings on each of these factors. However, any increases in intraclass correlation among the items of these two possible online support subscales could not compensate for decreases in subscale-specific reliability due to halving the number of items loading on each (Cronbach alpha < .80 for both). Therefore, we decided to treat all 12 items as belonging to a single construct, with the resulting 1-factor solution explaining 37.3% of the variance versus a combined 45.1% for the 2-factor solution.

Stratification by gender, race/ethnicity (non-Hispanic white vs other), marital status, education, and income showed that reliability remained high across all subgroups of interest (Cronbach alpha .86-.89).

As expected by the length of the arrows in Figure 1, the 1-factor solution was dominated by highly informative items Q11 and Q12. However, the magnitude of the loadings showed only moderate variation across items, ranging from .44 to .78. This led us to consider the possibility of using the total score of Q1 through Q12 as a scalar summary of online social support for quitting smoking. The resulting Pearson correlation (ρ = .99) between the total and factor scores suggested that the two measures were extremely highly correlated. Given the simplicity of calculating the Q1 through Q12 total score, we decided to use it as a proxy of the factor score from a single factor solution.

Figure 1. Biplot of Covariance Matrix.



Construct Validity

Table 2 presents results of convergent validity analyses. After correction for attenuation, Pearson correlations in the moderate range (ρ = .10-.30) according to Cohen’s definition [36] were observed between the OS4 and the PIQ Positive subscale (ρ = .24, *P* < .001), and between the OS4 and the difference between the PIQ Positive and Negative subscales (ρ = .20, *P* < .001). Despite the large sample size (n = 873), which was large enough to guarantee 84% power to detect Pearson correlations with magnitude as low as .10, the correlation between the OS4 and the Negative subscale of the PIQ failed to attain significance. Unexpectedly, small correlations were also detected with the ISEL Total Score (ρ = .06, *P* = .08) and the Appraisal (ρ= .07, *P* = .05) and Belonging (ρ=.08, *P* = .04) subscales; the correlation of the OS4 with the ISEL Tangible subscale was not significant. As expected, correlations between the OS4 and metrics of social network integration (number network members: ρ = .08, *P* = .02; social network diversity: ρ = .16, *P* < .001) as well as with frequency of Internet communications via social media (ρ = .16, *P* < .001) were in the positive direction but fell in the small-to-moderate range.

As hypothesized with regard to discriminant validity, results presented in Table 2 suggest that there was no association of the OS4 with the FTND Total Score, age of becoming a daily smoker, the number quit attempts in the past year, current symptoms of stress and depression, perceived health status, and duration and frequency of Internet use. Small-to-moderate correlations were observed between the OS4 and daily smoking rate at baseline (ρ = -.15, *P* < .001), desire to quit (ρ = .12, *P* = .006), confidence in quitting (ρ = .09, *P* = .04), and a past-year diagnosis of anxiety/depression (polyserial ρ = .12, *P* = .01).

Table 2. Construct validity analyses

	Correlation With OS4 Total Score	Correlation[a] With OS4 Total Score	*P* value[a]
Convergent validity			
Partner Interaction Questionnaire (PIQ)			
Positive subscale	.224	.244	<.001
Negative subscale	-.033	-.035	.36
Positive-Negative subscale	.168	.195	<.001
Interpersonal Support Evaluation List (ISEL)			
Appraisal subscale	.066	.073	.053
Belonging subscale	.069	.079	.04
Tangible subscale	.012	.015	.73
Total Score	.060	.064	.08
Social Network Index (SNI)			
Number of network members	.077		.02
Social network diversity	.159		<.001
Frequency of Internet Communications	.156		<.001
Discriminant validity			
Smoking variables			
Fagerstrom Test for Nicotine Dependence (FTND) total score	-.045	-.054	.32
Age of becoming a daily smoker	.000		.99
Daily smoking rate (at baseline)	-.147		<.001
Number quit attempts past year	-.039		.28
Desire to quit	.122		.006
Confidence in quitting	.092		.04
Psychosocial variables			
Perceived Stress Scale (PSS)	-.020	-.022	.58
Center for Epidemiological Studies Depression Scale (CES-D)	-.009	-.010	.79
Past year diagnosis anxiety/depression (yes/no)	.116		.01
Health status			
Perceived health status	-.049		.17
Internet use			
Duration of Internet use	-.035		.24
Frequency of Internet use	-.042		.32

[a] After correction for attenuation due to measurement error

Criterion Validity

In order to examine concurrent criterion validity, we first standardized the observed OS4 total score (mean 31.44, SD 7.96) to zero mean and unit variance in our overall sample and then calculated its average value within quartiles of the criterion of interest. As a result, observed between-group differences can be directly compared with Cohen's definitions [36] of effect size for continuous outcomes (small = .20, moderate = .50, large = .80), providing a yardstick for the practical significance of the findings.

Table 3 shows that higher ratings on each of the variables measuring satisfaction with the enhanced Internet website were associated with higher total scores on the OS4. Further, pairwise differences in mean OS4 total score across successive quartiles of intervention satisfaction level were all statistically significant (all *P*'s < .004), irrespective of the quartiles of interest and of the particular metric used to measure satisfaction levels. Discriminatory power was strong across the entire range of intervention satisfaction, with large differences in mean OS4 total score at the lower end of intervention satisfaction and moderate differences elsewhere (quartile differences: 1st vs 2nd = .64-.84, 2nd vs 3rd = .27-.44, 3rd vs 4th = .33-.53).

Table 3. Average OS4 score by level of intervention satisfaction

	Quartile of Intervention Satisfaction Metric			
	1st	2nd	3rd	4th
Satisfaction with website	-0.82	-0.18	0.20	0.53
Website met expectations	-0.78	0.10	0.17	0.51
Perceived helpfulness	-0.95	-0.19	0.22	0.75

Similarly, Table 4 shows that active participants in the online community had significantly larger OS4 scores than nonparticipants (all *P*'s < .001) on all of the participation metrics (use of any community features, internal messages sent to other members, and online buddies).

Table 4. Average OS4 score by community participation

	No	Yes	Difference
Used community features	-0.42	0.27	0.69
Sent Q-Mail	-0.24	0.93	1.17
Acquired online buddies	-0.08	1.44	1.52

With regard to predictive validity, logistic regression analyses revealed that 3-month values of the OS4 (mean 31.44, SD 7.96) adjusted for differences between treatment arm (enhanced Internet vs enhanced Internet + phone) were highly predictive of 30-day point-prevalence abstinence at 6, 12, and 18 months (all *P*'s < .001). Among those reached at follow-up, the odds of abstinence at 6 months rose by 48% for each standard unit increase in OS4 total score (adjusted odds ratio [OR] = 1.48, 95% CI 1.17-1.71), dropping only slightly to 37% at 12 and 18 months (adjusted OR = 1.37, 95% CI 1.17-1.59).

Discussion

The OS4 is a specific measure of online social support for smoking cessation, developed using previous theory and measures that have shown promise in understanding the role of social support in smoking cessation and relapse prevention. Developed within the context of a treatment outcome study of Internet and telephone treatment for smoking cessation, the OS4 emerged as a reliable and valid instrument. In a relatively large sample of Internet users, this 12-item scale demonstrated strong intraclass correlations across sociodemographic groups, resulting in high internal consistency reliability (Cronbach alpha .86-.89).

The OS4 demonstrated adequate construct validity. With regard to convergent validity, small-to-moderate correlations were observed with the Positive Support subscale of the PIQ, but not with the Negative Support subscale, a discrepancy that may be due to the fact that the OS4 was designed to capture the positive elements of supportive interactions in an online social network for cessation. A positive correlation with the ISEL Total Score was smaller than expected due to a lack of correlation between the OS4 and the Tangible subscale which contained items largely irrelevant to the experience of individuals interacting in an online social network (eg, If I was stranded 10 miles from home, there is someone I could call who could come and get me; If I were sick, I could easily find someone to help me with my daily chores). Small-to-moderate correlations were also observed with measures of social integration and the frequency of online communications via blogs and chat rooms. The OS4 also showed good discriminant validity in that it was not associated with measures such as nicotine dependence, general Internet use, and health status. Significant relationships of the OS4 with daily smoking rate, desire and confidence in quitting, and a history of nervous trouble/depression were quite small in magnitude but may indicate that smokers who are more motivated or who perceive the need for greater assistance in quitting tend to proactively reach out for support and engage in the community to a greater degree. Future research will need to clarify the nature of these associations.

All hypotheses regarding criterion validity were strongly supported. Concurrent validity was demonstrated by the significantly higher scores on the OS4 observed among subjects who reported higher levels of customer satisfaction as well as those who actively participated in the online community. Importantly for cessation research, the OS4 demonstrated excellent predictive validity with higher scores at 3 months predicting a greater likelihood of abstinence at 6, 12 and 18 months. Indeed, the OS4 may help to provide new insights into the role of social support in the cessation process and effective ways to harness support in intervention research. Despite historically robust associations between social support and better cessation treatment outcomes, numerous attempts to increase social support and enhance treatment effectiveness have been largely ineffective in increasing abstinence. For the most part, these interventions took place in face-to-face treatments and included spouse or partner training and “buddy” interventions

[37-40]. Several explanations have been offered for the lack of effectiveness of these kinds of support interventions. First, it may be difficult to change longstanding interpersonal dynamics through interventions with a spouse or partner. Second, intensive face-to-face treatment programs may provide a sufficient level of support such that additional components provide no added value with respect to social support; the challenge, however, is that less than 5% of smokers are interested in attending face-to-face treatment programs [41], making it critical to identify an appealing and accessible treatment modality that can provide the same type of intensive support. Third, social support may be a stable or "traitlike" construct that is resistant to change within a time-limited intervention; traditional treatment programs are time-limited in nature, typically lasting only 8 to 12 weeks. It may be that the creation or modification of meaningful supportive relationships requires more sustained interaction than traditional modalities can provide. Finally, it may be that the number and/or characteristics of people in an individual's "real-world" (ie, face-to-face) network may not be sufficient to provide the type or frequency of supportive interactions necessary to influence cessation outcomes.

The changing landscape of Internet-based social interactions and the ubiquity of online social networks provide an exciting opportunity to revisit social support mechanisms and interventions. By their nature, online social networks for cessation now make possible the provision and receipt of support in ways that were not feasible, convenient, or practical within face-to-face interventions or social networks of family and friends. The Internet affords continuous and real-time availability of thousands of supportive others in all stages of the quitting process, the rapid spread of information through network ties, and the ability to remain anonymous, among other factors. Christakis and colleagues [6,42] showed recently that smoking cessation and obesity spread more rapidly in the proximal social networks of probands than in unrelated networks, illustrating the importance of network effects in addition to interindividual effects in the social support process. As an assessment instrument specifically designed for exploring the links of perceived online social support to intermediate variables and cessation outcomes among those trying to quit, the OS4 may be helpful in advancing theory and improving the design and effectiveness of online cessation interventions.

Results should be considered in the context of several limitations. First, the measure was derived and validated on the same sample, potentially exaggerating the significance of the findings. Future work is needed to independently validate the factor structure, reliability, and validity of the OS4 in a new and different sample of smokers. Second, the measure was developed specifically within the context of one Web-based smoking cessation intervention, QuitNet.com. Development efforts ensured that items were relevant to the features and functionality of this particular website, and items specifically referenced QuitNet. Other research will need to adapt this measure to the specific Internet resource being evaluated and confirm that items remain relevant.

In summary, the 12-item OS4 is a reliable and valid instrument, developed to advance understanding of the emergent role of online social networks for smoking cessation treatment. To our knowledge, it represents the first psychometric scale developed for this purpose and is a relatively brief instrument that can be included in intervention research where response burden is a concern. The measure can also improve our understanding of basic mechanisms of action, develop and advance theories of behavior change on the Internet, and inform the development of tailored interventions to improve the effectiveness of interventions on cessation outcomes and relapse prevention. Development of an instrument to measure online social network and support activities, such as the OS4, can also inspire similar work in other areas of health promotion, disease prevention, and chronic disease management where social support also plays an important role.

Acknowledgments

This research was funded by grant R01CA104836 from the National Cancer Institute of the National Institutes of Health. Mr. Moreno is now at the University of Missouri – Kansas City in Kansas City, Missouri. Hakmook Kang, PhD, is now in the Department of Biostatistics at Vanderbilt University School of Medicine, Nashville, Tennessee.

Conflicts of Interest

None declared

Multimedia Appendix 1

Inter-item correlation matrix of OS4.

[PDF File (Adobe PDF File), 27KB - jmir_v13i3e69_app1.pdf]

References

1. Cohen S, Williamson GM. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health. Newbury Park, CA: Sage Publications; 1988:31-67.
2. Cohen S, Lichtenstein E. Partner behaviors that support quitting smoking. J Consult Clin Psychol 1990 Jun;58(3):304-309. [Medline: 2365893]

3. Coppotelli HC, Orleans CT. Partner support and other determinants of smoking cessation maintenance among women. J Consult Clin Psychol 1985 Aug;53(4):455-460. [Medline: 4031200]
4. Mermelstein R, Cohen S, Lichtenstein E, Baer JS, Kamarck T. Social support and smoking cessation and maintenance. J Consult Clin Psychol 1986 Aug;54(4):447-453. [Medline: 3745596]
5. Roski J, Schmid LA, Lando HA. Long-term associations of helpful and harmful spousal behaviors with smoking cessation. Addict Behav 1996;21(2):173-185. [Medline: 8730519]
6. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258 [FREE Full text] [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
7. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176. [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
8. Burri M, Baujard V, Etter JF. A qualitative analysis of an internet discussion forum for recent ex-smokers. Nicotine Tob Res 2006 Dec;8 Suppl 1:S13-S19. [Medline: 17491166]
9. Etter JF. Internet-based smoking cessation programs. Int J Med Inform 2006 Jan;75(1):110-116. [doi: 10.1016/j.ijmedinf.2005.06.014] [Medline: 16125450]
10. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289. [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]
11. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216. [doi: 10.1080/14622200500055319] [Medline: 16036277]
12. Vallone DM, Duke JC, Cullen J, McCausland KL, Allen JA. Evaluation of EX: a national mass media smoking cessation campaign. Am J Public Health 2011 Feb;101(2):302-309. [doi: 10.2105/AJPH.2009.190454] [Medline: 21164094]
13. An LC, Schillo BA, Saul JE, Wendling AH, Klatt CM, Berg CJ, et al. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. J Med Internet Res 2008;10(5):e55 [FREE Full text] [doi: 10.2196/jmir.1018] [Medline: 19103587]
14. Zbikowski SM, Hapgood J, Smucker Barnwell S, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. J Med Internet Res 2008;10(5):e41 [FREE Full text] [doi: 10.2196/jmir.999] [Medline: 19017583]
15. DeVellis RF. Scale Development: Theory and Applications. Thousand Oaks, CA: Sage; 2003.
16. Graham AL, Papandonatos GD, Bock BC, Cobb NK, Baskin-Sommers A, Niaura R, et al. Internet- vs. telephone-administered questionnaires in a randomized trial of smoking cessation. Nicotine Tob Res 2006 Dec;8 Suppl 1:S49-S57. [Medline: 17491171]
17. Cohen S, Mermelstein R, Kamarck T, Hoberman H. Measuring the functional components of social support. In: Sarason IG, Sarason BR, editors. Social Support: Theory, Research, and Application. The Hague, the Netherlands: Martinus Nijhoff; 1985:73-94.
18. Cohen S. Social supports and physical health: Symptoms, health behaviors and infectious disease. In: Cummings EM, Greene AL, Karraker KH, editors. Life-span Developmental Psychology: Perspectives on Stress and Coping. Hillsdale, NJ: Erlbaum Associates; 1991:213-234.
19. Graham AL, Bock BC, Cobb NK, Niaura R, Abrams DB. Characteristics of smokers reached and recruited to an internet smoking cessation trial: a case of denominators. Nicotine Tob Res 2006 Dec;8 Suppl 1:S43-S48. [Medline: 17491170]
20. Graham AL, Cobb NK, Papandonatos GD, Moreno JL, Kang H, Tinkelman DG, et al. A randomized trial of Internet and telephone treatment for smoking cessation. Arch Intern Med 2011 Jan 10;171(1):46-53. [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]
21. Wills TA, Shinar O. Measuring perceived and received social support. In: Cohen S, Underwood LG, Gottlieb BH, editors. Social Support Measurement and Intervention: A Guide for Health and Social Scientists. New York, NY: Oxford University Press; 2000:86-135.
22. House JS, Kahn RL. Measures and concepts of social support. In: Cohen S, Syme SL, editors. Social Support and Health. Orlando, FL.: Academic Press; 1985:83-108.
23. Wills TA. Supportive functions of interpersonal relationships. In: Cohen S, Syme SL, editors. Social Support and Health. Orlando, FL: Academic Press; 1985:61-82.
24. Argyle M. Benefits produced by supportive social relationships. In: Veiel HOF, Baumann U, editors. The Meaning and Measurement of Social Support. New York, NY: Hemisphere; 1992:13-32.
25. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
26. McMahon SD, Jason LA. Social support in a worksite smoking intervention. A test of theoretical models. Behav Modif 2000 Apr;24(2):184-201. [Medline: 10804679]
27. Gruder CL, Mermelstein RJ, Kirkendol S, Hedeker D, Wong SC, Schreckengost J, et al. Effects of social support and relapse prevention training as adjuncts to a televised smoking-cessation intervention. J Consult Clin Psychol 1993 Feb;61(1):113-120. [Medline: 8450096]

28. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
29. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
30. 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]
31. Gabriel KR. The biplot graphical display of matrices with application to principal component analysis. Biometrika 1971;58:453-467. [doi: 10.1093/biomet/58.3.453]
32. Gower JD, Hand J. Biplots. London, England: Chapman & Hall; 2006.
33. Nunnally JC. Psychometric Theory. 2nd edition. New York, NY: McGraw-Hill; 1978.
34. Salgado JF. VALCOR: A program for estimating standard error, confidence intervals, and probability of corrected validity. Behavior Research Methods, Instruments, & Computers 1997;29(3):464-467. [doi: 10.3758/BF03200601]
35. Schmidt FL, Hunter JE, Urry VW. Statistical power in criterion-related validated studies. Journal of Applied Psychology 1976;61(4):473-485. [doi: 10.1037/0021-9010.61.4.473]
36. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Erlbaum; 1988.
37. May S, West R, Hajek P, McEwen A, McRobbie H. Randomized controlled trial of a social support ('buddy') intervention for smoking cessation. Patient Educ Couns 2006 Dec;64(1-3):235-241. [doi: 10.1016/j.pec.2006.02.008] [Medline: 16616450]
38. Lichtenstein E, Glasgow RE, Abrams DB. Social support in smoking cessation: In search of effective interventions. Behavior Therapy 1986;17(5):607-619. [doi: 10.1016/S0005-7894(86)80098-3]
39. Lichtenstein E, Glasgow RE. Smoking cessation: what have we learned over the past decade? J Consult Clin Psychol 1992 Aug;60(4):518-527. [Medline: 1506500]
40. Park EW, Schultz JK, Tudiver F, Campbell T, Becker L. Enhancing partner support to improve smoking cessation. Cochrane Database Syst Rev 2004(3):CD002928. [doi: 10.1002/14651858.CD002928.pub2] [Medline: 15266469]
41. Lichtenstein E, Hollis J. Patient referral to a smoking cessation program: who follows through? J Fam Pract 1992 Jun;34(6):739-744. [Medline: 1593248]
42. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007 Jul 26;357(4):370-379 [FREE Full text] [doi: 10.1056/NEJMsa066082] [Medline: 17652652]

Abbreviations

CES-D: Center for Epidemiological Studies Depression Scale
FTND: Fagerstrom Test for Nicotine Dependence
ISEL: Interpersonal Support Evaluation List
OS4: Online Social Support for Smokers Scale
PIQ: Partner Interaction Questionnaire
PSS: Perceived Stress Scale
SNI: Social Network Index

Edited by G Eysenbach; submitted 16.03.11; peer-reviewed by R Mermelstein; comments to author 06.05.11; revised version received 10.05.11; accepted 11.05.11; published 28.09.11.

Please cite as:
Graham AL, Papandonatos GD, Kang H, Moreno JL, Abrams DB
Development and Validation of the Online Social Support for Smokers Scale
J Med Internet Res 2011;13(3):e69
URL: http://www.jmir.org/2011/3/e69/
doi:10.2196/jmir.1801
PMID:21955465

Original Paper

Perceived Threat and Corroboration: Key Factors That Improve a Predictive Model of Trust in Internet-based Health Information and Advice

Peter R Harris[1], PhD; Elizabeth Sillence[2], PhD; Pam Briggs[2], PhD

[1]Department of Psychology, University of Sheffield, Sheffield, United Kingdom

[2]Department of Psychology, School of Life Sciences, Northumbria University, Newcastle upon Tyne, United Kingdom

Corresponding Author:
Elizabeth Sillence, PhD
Department of Psychology
School of Life Sciences
Northumbria University
Northumberland Building
Newcastle upon Tyne, NE1 8ST
United Kingdom
Phone: 44 191 2437247
Fax: 44 191 227 3190
Email: elizabeth.sillence@northumbria.ac.uk

Abstract

Background: How do people decide which sites to use when seeking health advice online? We can assume, from related work in e-commerce, that general design factors known to affect trust in the site are important, but in this paper we also address the impact of factors specific to the health domain.

Objective: The current study aimed to (1) assess the factorial structure of a general measure of Web trust, (2) model how the resultant factors predicted trust in, and readiness to act on, the advice found on health-related websites, and (3) test whether adding variables from social cognition models to capture elements of the response to threatening, online health-risk information enhanced the prediction of these outcomes.

Methods: Participants were asked to recall a site they had used to search for health-related information and to think of that site when answering an online questionnaire. The questionnaire consisted of a general Web trust questionnaire plus items assessing appraisals of the site, including threat appraisals, information checking, and corroboration. It was promoted on the hungersite.com website. The URL was distributed via Yahoo and local print media. We assessed the factorial structure of the measures using principal components analysis and modeled how well they predicted the outcome measures using structural equation modeling (SEM) with EQS software.

Results: We report an analysis of the responses of participants who searched for health advice for themselves (N = 561). Analysis of the general Web trust questionnaire revealed 4 factors: information quality, personalization, impartiality, and credible design. In the final SEM model, information quality and impartiality were direct predictors of trust. However, variables specific to eHealth (perceived threat, coping, and corroboration) added substantially to the ability of the model to predict variance in trust and readiness to act on advice on the site. The final model achieved a satisfactory fit: $\chi^2_5 = 10.8$ ($P = .21$), comparative fit index = .99, root mean square error of approximation = .052. The model accounted for 66% of the variance in trust and 49% of the variance in readiness to act on the advice.

Conclusions: Adding variables specific to eHealth enhanced the ability of a model of trust to predict trust and readiness to act on advice.

(J Med Internet Res 2011;13(3):e51) doi:10.2196/jmir.1821

KEYWORDS

Internet, trust, e-health, threat, fear-appeal, social cognition models.

Introduction

The Internet is an important source of health information and advice. Eight in 10 Internet users have looked online for health information [1] with young people in particular finding it to be a congenial source of health information and advice [2,3]. People use the Internet to help them with a variety of problems related to health, such as seeking reassurance or obtaining information before visiting the physician [4]. They may also turn to it after such visits if, for example, they feel they have not been given enough time or information to make sensible treatment choices [5] or find it difficult to recall elements of their consultations [6,7]. Yet the volume of available material and the unregulated nature of health information on the Internet pose potential problems for users. Indeed, investigations of eHealth material across a range of contexts typically conclude that quality is a problem (e.g., [8,9]). In the face of such variable quality, how do people decide which information to accept when seeking health information and advice?

Trust—"an expectancy held by an individual or group that the word, promise, verbal or written statement of another individual or group can be relied upon" (p. 651) [10]—is widely considered to be pivotal to this process. However, empirical research suggests that the strategies people use when deciding which Internet sites to trust are often suboptimal. For example, e-commerce consumers are influenced by potentially misleading cues, such as the look and feel of the site, trusting sites rated highly in visual appeal and mistrusting sites with poor visual design or with unprofessional errors [11-13]. In the health domain, users are also prone to reject high-quality medical sites because of poor visual appeal [14]; they also typically fail to check website authorship or read disclosure statements, despite suggesting these as important quality markers beforehand [15]. They also make prolific use of general search engines, thereby potentially exposing themselves to large numbers of poor-quality sites [16]. Perceptions of the motives underlying the site also appear to be important and may strongly influence the outcome. For instance, UK participants mistrusted the advice and information on websites openly sponsored by pharmaceutical or other commercial companies [14,17], even though such sites are often recommended by expert reviewers as providing the most accurate health information [18]. Likewise, very early in the selection process participants may reject high-credibility, high-quality sites if they do not appear to be aimed at "people like them" [17]. Users also have a broad view of what constitutes expertise, rating highly expertise displayed by patients and caregivers, as well as medical personnel, which may lead to disparities between the decisions the users make about site quality and those of expert reviewers [19].

To better understand the factors that determine trust, researchers have developed models of the process by which users form trust relationships online [14,16,20,21]. For example, Briggs et al [16] developed a 3-factor model, following analysis of a scale they developed to assess trust in e-commerce, in which trust was determined by 3 key predictors: source credibility (the extent to which the advice came from a knowledgeable, expert source), personalization (the extent to which the respondent felt involved in the advice process and the advice was tailored to them), and predictability (the extent to which the site appeared familiar and the advice met their expectations).

Such models provide a promising starting point for modeling trust in eHealth but, because they are typically either generic (eg, [20,21]) or have been developed to understand trust in another domain, such as e-commerce (eg, [16,22]), they are likely to require augmentation to improve their fit to trust in eHealth. People are not neutral processors of health-risk information [23]: those searching for an explanation for their symptoms naturally prefer one that is benign, and those taking tests to assess their chances of contracting some disease prefer to discover their chances are low [24,25]. In the health domain, that is, people often have strong initial expectations and show preferences for particular sorts of information, and these may shape their search strategies [26] and influence which sites they trust. They may respond differently to sites containing information and advice they find unwelcome or threatening than they respond to sites containing information they find congenial and comforting [23,27]. This distinctive element may not be captured by existing models of trust that have their origins elsewhere, such as in e-commerce. Accordingly, a key aim of this study was to examine whether adding variables from models designed to capture how people respond to health-risk information enhances the capacity of a general model of trust to account for (1) trust in advice presented on health websites and (2) readiness to act on that advice.

Researchers in health psychology have developed models to account for the ways in which people respond to health-risk information, including information the individual may find uncomfortable or threatening (eg, protection motivation theory [PMT] [28,29]; the extended parallel process model (EPPM) [30]) [31]. For example, in PMT and the EPPM, perceived threat—a product of the person's appraisal of the harm that would occur as a result of the hazard (severity) and their personal susceptibility to it (vulnerability)—is a critical predictor of response to health information. The EPPM proposes that, as perceived threat increases, so does the individual's appraisal of the extent to which they can take steps to control the hazard. Where they are persuaded that they can alleviate the hazard, they are motivated to do so [28,30]. Consequently, we assessed whether adding measures of threat, control, and coping affected perceptions of trust. Given the unregulated nature of the Internet, we also examined the possibility that the extent to which people are prepared to trust and accept unpalatable or threatening online information depends on whether it is corroborated. To test this, we included measures of the extent to which participants reported checking the information and finding it consistent with information they obtained elsewhere.

In the current study, we modeled trust and readiness to act on online health advice (our outcome measures). We did this in 2 steps: first, we assessed the factorial structure of a general measure of trust in online eHealth sources and modeled how well the factors predicted the outcomes. We did this by taking a model of online trust developed from data on e-commerce [16], supplementing it with items derived from our qualitative research on trust in eHealth [14,32], and testing it against a data set that covers a range of conditions, diseases, and health-related issues. Second, we assessed whether adding variables designed to capture elements of the response to health-risk information that is uncongenial (such as threat and control) and online (such as information checking and corroboration) enhanced the model's predictive power.

Method

Participants

In total, 1902 participants completed the online questionnaire, which was promoted on the hungersite.com website. Visitors to the website could click on an advertising graphic and were transferred to the online survey. For each click on an advertisement on the hungersite page, a donation of US $0.05 is made to the UN World Food Programme. We had used the hungersite successfully in the past and chose it for this study because of its relatively broad international appeal and its focus on charitable donation. The URL for the questionnaire was also submitted to Yahoo and other search engines. A press release was also put out to local print media (eg, university newsletters and the local newspaper).

Incomplete questionnaires (mostly comprising 1 or 2 initial answers only) were removed (n = 415) and an internal consistency check was applied to the data to eliminate duplicate responders. This involved matching the respondent's stated location with their Internet protocol (IP) address and led to a further 5 replies being removed, leaving a total of 1482 respondents. Of these, 1103 reported having used the Internet for health advice, of whom just over half (561, 51%) reported searching the Internet for health advice for themselves. This group (ie, those reporting searching for advice for themselves) formed the sample for the current paper (other respondents were not directed to the pages asking the threat appraisal questions). Of the sample, 402 (72%) were female and 92 (16%) were male (the remainder did not specify), and age ranged from under 18 to over 64 years and was spread quite evenly within this range (the modal age was 25–35 years old). There were no significant associations between searching for self and searching for other people in gender (n = 832, χ^2_1 = .35, P = .552), age (n = 832, χ^2_5 = 5.76, P = .330), highest education level (n = 832, χ^2_3 = 3.19, P = .364), or country of residence (n = 832, χ^2_4 = 9.31, P = .054; see Table 1).

Table 1. Background characteristics of participants and the health topics searched for

Participant characteristic	Frequency (%)	International Classification of Primary Care (where applicable)
Gender		
Male	92 (16%)	
Female	402 (72%)	
Age (years)		
<18	21 (4%)	
18–24	94 (19%)	
25–35	129 (26%)	
36–44	73 (15%)	
45–54	90 (18%)	
55–64	65 (13%)	
>64	24 (5%)	
Highest education level		
High school	70 (14%)	
College	133 (27%)	
University	151 (30%)	
Postgraduate	142 (29%)	
Country/region of residence		
United States	290 (59%)	
Canada	37 (8%)	
United Kingdom	84 (17%)	
Western Europe	28 (6%)	
Eastern Europe	8 (2%)	
Australasia	17 (4%)	
Central/South America	9 (2%)	
Middle East	4 (<1%)	
Africa	1 (<1%)	
Japan	5 (1%)	
Other	5 (1%)	
Internet use (years)		
1–2	9 (2%)	
3–5	66 (14%)	
6–9	198 (41.2%)	
10–14	169 (35.1%)	
15–19	30 (6%)	
≥20	9 (2%)	
Health topic		
Allergies	19 (3%)	A92
Arthritis	23 (4%)	L88
Alternative health	42 (7%)	
Cancer	17 (3%)	Type of cancer not specified
Children’s health	8 (1%)	Not specified

Participant characteristic	Frequency (%)	International Classification of Primary Care (where applicable)
Diabetes	10 (2%)	T89
Diet	43 (8%)	
Depression	20 (4%)	P76
Fitness	37 (7%)	
Heart disease	19 (3%)	K71
Hypertension	8 (1%)	K86
Men’s health	19 (3%)	
Mental health (excluding depression)	12 (2%)	
Migraine	9 (2%)	N89
Women’s health	98 (17%)	
Sexually transmitted diseases	20 (4%)	
Thyroid problems	17 (3%)	T85; T86
Other	140	
Skin conditions (eczema, psoriasis, rosacea)	9 (2%)	S87; S91; S90
Fibromyalgia	5 (<1%)	L18
Fertility issues, pregnancy, miscarriage	5 (<1%)	W15; W78; W82
Influenza, pneumonia, colds	5 (<1%)	R80; R81; R74
Parkinson disease, multiple sclerosis, Alzheimer disease	4 (<1%)	N87; N86; P70
Anorexia	3 (<1%)	P86
Back pain	3 (<1%)	L02
Cold sores	2 (<1%)	S71
Heartburn	2 (<1%)	D03
Mumps	1 (<1%)	D71
High cholesterol	1 (<1%)	T93

Procedure

On the first page of the online questionnaire, participants were asked whether they had sought advice online about health. Those responding yes were taken to the subsequent screens of questions, including questions about the site they had previously used and their reasons for searching online, as well as demographic information (gender, age, educational attainment, country of residence, and length of Internet use) and the predictor and outcome measures. We report data from those participants answering yes to either of 2 key questions: “Last time you searched online for health advice were you trying to find out whether you might already have a particular disease/condition?” and “Last time you searched online for health advice were you searching for information/advice about the chances of you getting or preventing yourself from getting a particular disease in the future?” Participants were then asked to “Think about any one site that you visited during that search” and to answer the remaining questions with respect to that site—that is, “Thinking about the information or advice on the site please rate your agreement with the following statements.”

Measurements

Once participants had responded to the above statements, the following measures were taken. Except where indicated, responses were given on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*).

General Web Trust Questionnaire

The first measure contained the items designed to assess various aspects of trust in online sources derived from the 18-item trust questionnaire developed by Briggs et al [16] supplemented by 6 items derived from the qualitative research on eHealth conducted by Sillence and colleagues [14,32]. (The full set of items can be found in Table 2.)

Threat appraisal

Consistent with measurement of threat in PMT and the EPPM, threat appraisal was measured by combining 2 items we developed, one to assess susceptibility, “the site said my chances of having/getting the disease/condition were” (1 = *very low* to 5 = *very high*), and one to assess severity, “the site said my consequences of having the disease or condition were” (1 = *not at all severe* to 5 = *extremely severe*).

Reactions to the site

Affective and cognitive reactions to the site were assessed by 8 items. Following the stem “Looking at this site made me feel...” came the variables worried, reassured, at risk, confused, anxious about the risks, optimistic, in control, and able to cope. Responses were given on a 6-point scale with the following labels: 1 = *less*, 2 = *slightly less*, 3 = *no different*, 4 = *slightly more*, and 5 = *more*, plus *not applicable*.

Information checking and corroboration

In each case, two items measured (1) the extent to which participants checked other sources of information in addition to the website (“I checked other websites” and “I checked other sources”; *checked*, $r = .52$, $P < .001$), and (2) how consistent the advice was with these other sources (“I found the advice consistent with other websites” and “I found the advice consistent with other sources”; *corroboration*, $r = .74$, $P < .001$).

Outcome measures

The two outcome measures were *trust*, “I trusted the site,” and *readiness to act on the advice* the participants found on the site, “I intended to act upon the advice.” These were developed for the study.

Results

Participants reported a wide range of diseases and conditions, from cancer, depression, and Alzheimer disease to thyroid problems, allergies, and mumps (see Table 1). On average, participants reported moderate levels of threat (mean 3.63 [SD 1.11]).

To retain the separate identities of the general trust questionnaire and the measure developed here of specific reactions to the site, these measures were factor analyzed separately using principal component analysis with varimax rotation. Analysis of the general Web trust questionnaire revealed 4 factors accounting for 55.6% of the variance (see Table 2). The number of factors was determined by consulting the scree plot and with reference to Kaiser’s eigenvalue > 1 criterion. Factor 1 (alpha = .85), labeled *access to quality information*, brought together items mainly describing ease of use and access to high-quality information. Factor 2 (alpha = .86), labeled *personalization*, brought together items mainly describing the importance of tailored information and the ability to interact with “like-minded people” on the website. Factor 3 (alpha = .74), labeled *perceived impartiality,* brought together items describing the extent to which the advice on the website appeared impartial and objective. Factor 4 (alpha = .70), labeled *credible design*, brought together items describing the extent to which the site had credible design features.

Table 2. Four factors on the general Web trust questionnaire and items with loading on the relevant factor

Item number	Items and factor loading
Access to quality information	
6	The language on the site made it easy to understand (.77)[a]
7	The site helped me understand the issue better (.76)
5	The site was easy to use (.75)
1	The site told me most of what I needed to know (.67)
8	The layout was consistent with other sites (.50)
2	The advice appeared to be prepared by an expert (.49)
12	The advice seemed to be offered in my best interests (.49)
11	The advice came from a knowledgeable source (.48)
Personalization	
19	The site gave me a sense of being part of a community (.81)[a]
10	I was able to contribute to content on the site (.76)[a]
18	I felt involved in the way the site tried to find appropriate advice (.73)
17	It felt like the advice was tailored to me personally (.71)
14	The site contained contributions from like-minded people (.70)
16	There were opportunities to interact with other people on the site (.70)[a]
9	I identified with the site (.49)[a]
22	The reasoning behind the advice was explained to me (.46)
Perceived impartiality	
21	The advice appeared to be impartial and independent. (.78)
20	The site was free from advertisements (.73)[a]
13	The advice seemed objective (ie, no hidden agenda) (.61)
23	The advice seemed credible (.56)
Credible design	
4	The site was owned by a well-known organization (.76)
15	The site featured familiar logos (.72)
3	The site had a professional design (.60)
24	The site had an attractive design (.52)

[a] Item derived from Sillence and colleagues [14,31].

Analysis of the 8 items assessing reactions to the site revealed 2 factors: factor 1, which we labeled *coping* (alpha = .87), comprised the control, coping, and optimism items; factor 2, which we labeled *worry* (alpha = .77), comprised the worried, anxious, and at-risk items. The remaining items (confused and reassured) did not load significantly on either factor, so were entered separately in analyses.

To assess how scores on each of the above factors contributed to the prediction of trust and readiness to act on the advice, mean scores on each factor for each participant were calculated and entered into models as described below. Descriptive statistics for each variable in the final model, together with their intercorrelations, are presented in Table 3.

Table 3. Descriptive statistics and zero-order correlations between the measures[a]

	Information quality	Personalization	Impartiality	Credible design	Coping	Threat	Corroboration	Trust	Mean	SD
Information quality									3.95	0.64
Personalization	.43**								2.72	0.90
Perceived impartiality	.61**	.29**							3.72	0.82
Credible design	.54**	.30**	.28**						3.41	0.84
Coping	.41**	.33**	.30**	.18**					3.73	1.02
Threat	.18**	.09*	.10*	.07	.66**				3.63	1.11
Corroboration	.23**	.08	.17**	.09*	.19**	.57**			3.44	1.76
Outcomes										
Trust	.39**	.17**	.33**	.19**	.27**	.66**	.71**		3.68	1.58
Readiness to act	.24**	.17**	.17**	.08	.29**	.55**	.62**	.63**	3.30	1.88

[a] The range of each scale was the same as the scale points described in the procedure

*$P < .05$;

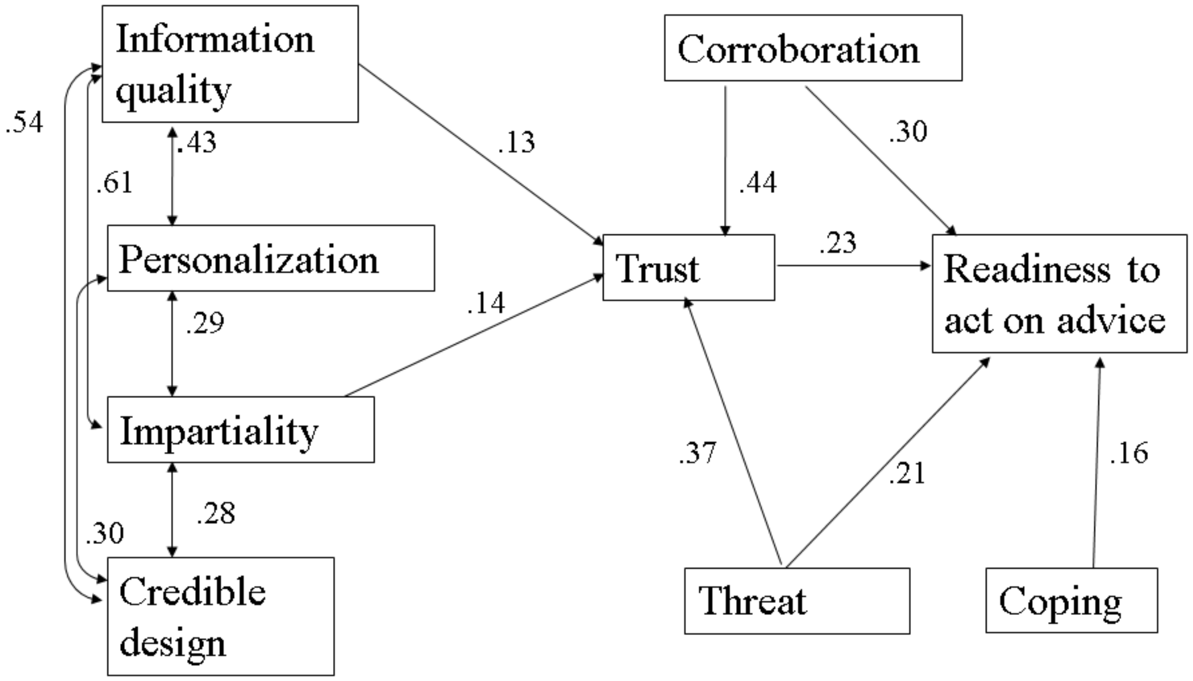
**$P < .01$.

The data were analyzed next using structural equation modeling with EQS (version 6.1; Multivariate Software Inc, Encino, CA, USA). The fit of the models was evaluated using chi-square, the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). Satisfactory fit of the model is obtained when chi-square is nonsignificant, CFI is >.90, and RMSEA is <.08 [33]. Path coefficients and R^2 values were also inspected in evaluating the predictive power of the models.

EQS was first used to test a model that included all paths from the 4 factors of the general Web trust questionnaire to the outcomes, trust and readiness to act on the advice. The extent to which the variables additional to these 4 factors had unmediated effects on the outcomes was then examined by introducing, in turn, a direct path between each predictor variable and each outcome variable. Alternative models were compared using the different fit indices and the extent to which they explained variance in trust and readiness to act (R^2).

The initial model accounted for 15% of the variance in trust and 39% of the variance in readiness to act and had a poor fit ($\chi^2_4 = 11.3$, $P = .02$, CFI = .99, RMSEA = .057); as predicted, adding *threat* improved the fit of the model, as did adding *coping* and *corroboration*. Adding the remaining variables (ie, *worry*, *checked*, *confused*, *reassured*) did not improve fit. The added variables produced no substantive changes in other paths in the model. The fit of the final model was good: chi-square was not significant ($\chi^2_5 = 10.8$, $P = .21$), the other measures of fit indicated a good fit (CFI = .99, RMSEA = .052), and the model accounted for 66% of the variance in trust ($R^2 = .66$) and 49% of the variance in readiness to act on the advice ($R^2 = .49$). The final model, with its significant pathways ($P < .05$), is displayed in Figure 1. Descriptive statistics and intercorrelations for the measures not included in the final model are in Multimedia Appendix 1.

Figure 1. The final model, showing the significant path coefficients (P < .05)



Discussion

The current study aimed to (1) assess the factorial structure of a general measure of Web trust, (2) model how the resultant factors predicted trust in, and readiness to act on, the advice found on health-related websites, and (3) test whether adding variables to capture elements of the response to uncongenial, online health-risk information enhanced the prediction of these outcomes. As predicted, incorporating the latter variables added to the ability of the model to predict variance in both trust and expressed readiness to act on advice. The final model accounted for substantial amounts of variance in both outcome measures. Four factors emerged from analysis of the general Web trust questionnaire. However, a key feature of our final model is that only 2 of these, impartiality and information quality, had direct effects on trust in health-related websites; the effects of the other 2 factors, personalization and credible design, were indirect and mediated through impartiality and information quality. This is consistent with findings from the earlier, qualitative phase of the current research program, which also signaled that, in the health domain, seeking high-quality, independent advice might be critical [17]. It is notable that in previous research in other domains impartiality has emerged as a predictor but not occupied the pivotal role it has here [16]. Indeed, impartiality is not necessarily expected in such domains as e-commerce, although when it is encountered it is valued [13]. The importance of credibility through impartiality here is also consistent with the basis of trust in patient–physician relationships, in which there is an assumption that physicians will act in the best interests of the patient [34].

Clearly, elements of the general trust component of the current model correspond with those in other trust models [21], suggesting that researchers are isolating the core general features of trust across a range of domains, from eHealth through e-commerce to leisure [20,22,35,36]. What differs appears to be the relative importance of certain factors in each domain and at different phases of the trust process [16,17]. However, it is notable that the initial model, which contained only the general Web trust factors, formed a model with relatively poor fit to the data, supporting the decision to search for additional variables to enhance fit to a specific domain (in this case, eHealth).

Both threat and corroboration contributed to the prediction of trust, with direct positive relationships in each case. According to these data, therefore, people are prepared to trust sites that tell them things they can verify elsewhere and sites that tell them things they would prefer not to hear. It is interesting to find that they do not, therefore, appear to let unwelcome news interfere with the process of assessing trust. Research needs next to address moderators of this relationship. For example, do those high in need for closure [37] or low in tolerance for uncertainty [38] show less readiness to trust sites containing high-quality but uncongenial information? As a variable, trust may also have potential to help clarify the different relationships found in previous research between threat and message acceptance, where meta-analyses have, for example, found a linear relationship between fear and intentions [30], but studies have also shown that the most threatened groups often also show the least acceptance (see [23,27] for recent reviews).

The absence of a relationship between coping and trust is also worthy of note. Just as one might have expected people to want to trust sites telling them things about vulnerability and disease severity they found congenial, one might have expected people also to want to trust sites telling them positive things about their ability to control diseases or otherwise boost their sense of efficacy and coping; yet here this was not the case. Indeed, in many ways the current findings illustrate that trust is not simply likeability: people appear to be prepared to trust sites containing information they dislike. In developing trust perceptions, our findings suggest that people may be more prepared to accept information they dislike than basic design features they dislike [32].

Encountering a site that increased one's perceptions of being able to cope, along with threat and corroboration, contributed independently and positively to the individual's readiness to act on the advice the site provided. Indeed, while trust has been shown in this study to predict readiness to respond to advice, and to partially mediate the effects of threat and corroboration in the process, in the model outlined here it is only one of several independent contributors to readiness to act. The benefits from adding threat-related and other variables suggest that our understanding of the process by which trust perceptions are translated into relevant behavior is likely to be enhanced by further integration with relevant social cognition models of behavior, such as the PMT or the EPPM. The model outlined here is also one of the few models of trust to explicitly account for the perceived costs of undertaking a transaction (see [36] for another). Notions of vulnerability, cost, or perceived risk have been argued to be effective in improving the predictive value of trust models in a variety of contexts [39].

The study has a number of limitations. It is not possible to assess the extent to which retrospective biases (such as a schema for rationality) may be contributing to the picture presented. Nevertheless, with such a large sample size and range of diseases and medical conditions (see Table 1), it is clear that this rational model comprises a significant element of the picture people tell the world (and perhaps also themselves) about their responses to health information they encounter on the Internet. Unfortunately, it was not possible to test whether those who started to complete the online questionnaire, but quickly dropped out, differed in key ways from those who persisted. It remains possible, therefore, that the current sample may be unrepresentative. However, the sample tested here (those searching for information for themselves) did not differ from the other group of respondents who also completed the survey (those searching for information for someone else) in terms of the key variables reported in Table 1, which offers some reassurance about their representativeness. It would also be useful to assess in future research of this kind the individual respondent's health status, to establish whether this moderates responses. Likewise, while it is clear from the research reported here that the inclusion of threat-related variables in a model of trust helps predict readiness to act on the advice, it is also known that there is often a gap between intentions and behavior. A complete picture needs to dovetail the processes modeled here with those known to affect the process of translating intentions into behavior [31,40]. What determines impartiality perceptions is also likely to vary between cultures. For example, in societies in which commercial interests are more readily linked to health care provision than in the United Kingdom, the predictors of impartiality may differ. The role of such factors may also be the focus of valuable future research.

The final model supports the hypothesis that people will seek to validate what they find on websites against other sources of information. If they find that information is corroborated elsewhere, this boosts trust in the website. Given the relative novelty of websites as sources of health information, this process can be seen as one of calibrating the credibility of these novel sources against more tried and tested ones. Over time, as people become more experienced and adept at assessing the credibility of Internet sources, we would expect this checking process to become less important.

Finally, for website developers, the current data contain some useful pointers. Four key design factors (information quality, personalization, perceived impartiality, and credible design) have an important (albeit moderated) role to play in influencing trust and the subsequent decision to act on the advice given.

Acknowledgments

The research reported here was supported by ESRC E-Society grant number RES-335-25-0019. We would like to thank Phoenix Mo for her assistance in formatting and editing the paper.

Conflicts of Interest

None declared

Multimedia Appendix 1

Zero-order correlations among the measures not included in the final model together with their descriptive statistics

[PDF file (Adobe PDF File), 78 KB - jmir_v13i3e51_app1.pdf]

References

1. Fox S. Peer-to-peer healthcare. Washington, DC: Pew Internet & American Life Project; 2011 Feb 28. URL: http://www.pewinternet.org/~/media//Files/Reports/2011/Pew_P2PHealthcare_2011.pdf [accessed 2011-07-19] [WebCite Cache ID 60I47XlTi]
2. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
3. Klein JD, Wilson KM. Delivering quality care: adolescents' discussion of health risks with their providers. J Adolesc Health 2002 Mar;30(3):190-195. [Medline: 11869926]
4. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [doi: 10.1016/S0738-3991(03)00116-2] [Medline: 15062905]
5. Carvel J. Lack of information worries NHS patients. The Guardian 2005 Feb 21:10.
6. Kessels RP. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222 [FREE Full text] [Medline: 12724430]
7. Jansen J, Butow PN, van Weert JCM, van Dulmen S. Does age really matter? Recall of information presented to newly referred patients with cancer. J Clin Oncol 2008;28:5450-5457. [Medline: 18936478]
8. Fox S, Rainie L. Pew Internet and American Life Project. Washington, DC: Pew Internet & American Life Project; 2000 Nov 26. The online health care revolution: how the Web helps Americans take better care of themselves URL: http://www.pewinternet.org/~/media//Files/Reports/2000/PIP_Health_Report.pdf.pdf [accessed 2011-07-19] [WebCite Cache ID 60I4MXhx7]
9. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 12020305]
10. Rotter JB. A new scale for the measurement of interpersonal trust. J Pers 1967 Dec;35(4):651-665. [Medline: 4865583]
11. Fogg BJ, Marshall J, Laraki O. What makes web sites credible? A report on a large quantitative study. In: Human Factors & Computing Systems Proceedings, 2001 Mar 31-Apr 5, Seattle. New York, NY: ACM; 2001:61-68.
12. Kim J, Moon JY. Designing towards emotional usability in customer interfaces: trustworthiness of cyber-banking system interfaces. Interacting with Computers 1998;10:1-29. [doi: 10.1016/S0953-5438(97)00037-4]
13. Standford J, Tauber E, Fogg BJ, Marable L. Consumer Reports WebWatch. Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites URL: http://www.consumerwebwatch.org/pdfs/expert-vs-online-consumers.pdf [accessed 2011-07-19] [WebCite Cache ID 60I4PuRuL]
14. Sillence E, Briggs P, Harris P, Fishwick L. A framework for understanding trust factors in web-based health advice. Int J Hum Comput Stud 2006;64:697-713. [doi: 10.1016/j.ijhcs.2006.02.007]
15. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
16. Briggs P, Burford B, De Angeli A, Lynch P. Trust in online advice. Soc Sci Comput Rev 2002;20:321-332. [doi: 10.1177/089443930202000309]
17. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
18. Reed M, Anderson C. Evaluation of patient information Internet web sites about menopause and hormone replacement therapy. Maturitas 2002 Oct 25;43(2):135-154. [Medline: 12385862]
19. Sillence E, Briggs P, Herxheimer A. Personal experiences matter: what patients think about hypertension information online. He@lth Infom Internet 2004;42:3-5.
20. Corritore CL, Kracher B, Wiedenbeck S. On-line trust: concepts, evolving themes, a model. Int J Hum Comput Stud 2003;58:737-758. [doi: 10.1016/S1071-5819(03)00041-7]
21. Wang YD, Emurain HH. An overview of online trust: concepts, elements and implications. Comput Hum Behav 2005;21:105-125. [doi: 10.1016/j.chb.2003.11.008]
22. Egger FN. Affective design of e-commerce user interfaces: how to maximise perceived trustworthiness. In: Helander M, Hkalid HM, Tham MP, editors. Proceedings of CAHD2001. London, UK: Asean Academic Press; 2001 Presented at: Affective Human Factors Design; June 27-29; Singapaore.
23. Good A, Abraham C. Measuring defensive responses to threatening messages: a meta-analysis of measures. Health Psychol Rev 2007;1:208-229. [doi: 10.1080/17437190802280889]
24. Renner B. Biased reasoning: adaptive responses to health risk feedback. Pers Soc Psychol Bull 2004 Mar;30(3):384-396. [doi: 10.1177/0146167203261296] [Medline: 15030627]
25. Ditto PH, Lopez DF. Motivated skepticism: use of differential decision criteria for preferred and nonpreferred conclusions. J Personal Soc Psychol 1992;63:568-584. [doi: 10.1037//0022-3514.63.4.568]
26. Joinson AN, Banyard P. Psychological aspects of information seeking on the Internet. Aslib Proc 2002;54:95-102. [doi: 10.1108/00012530210435220]
27. Harris PR, Epton T. The impact of self-affirmation on health-related cognition and health behaviour: issues and prospects. Soc Personal Psychol Compass 2010;4:439-454. [doi: 10.1111/j.1751-9004.2010.00270.x]

28. Rogers RWCacioppo BL, Petty LL. Cognitive and physiological processes in fear appeals and attitude change: a revised theory of protection motivation. In: Social Psychophysiology: A Sourcebook. New York, NY: Guilford Press; 1983:153-176.
29. Rogers RW, Prentice-Dunn SGochman DS. Protection motivation theory. In: Handbook of Health Behavior Research. New York, NY: Plenum Press; 1997:113-132.
30. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Commun Monogr 1992;59:329-349. [doi: 10.1080/03637759209376276]
31. Sheeran P, Milne S, Webb TL, Gollwitzer PM. Implementation intentions and health behaviour. In: Conner M, Norman P, editors. Predicting Health Behaviour: Research and Practice with Social Cognition Models, 2nd ed. Maidenhead, UK: Open University Press; 2005:276-323.
32. Sillence E, Briggs P, Fishwick L, Harris P. Trust and mistrust of online health sites. In: Chi 2004: Connect: Conference Proceedings: April 24-29, Vienna, Austria: Conference on Human Factors in Computing Systems: Vienn (Chi Letters). New York, NY: ACM; 2004:663-670.
33. Marsh H, Balla J, Hau KT. An evaluation of incremental fit indices: a clarification of mathematical and empirical properties. In: Marcoulides GA, Schumacker RE, editors. Advanced Structural Equation Modeling: Issues and Techniques. Mahwah, NJ: L Erlbaum Associates; 1996.
34. Hall MA, Zheng B, Dugan E, Camacho F, Kidd KE, Mishra A, et al. Measuring patients' trust in their primary care providers. Med Care Res Rev 2002 Sep;59(3):293-318. [Medline: 12205830]
35. Egger FN. “Trust me, I’m an online vendor”: towards a model of trust for e-commerce system design. In: CHI Extended Abstracts on Human Factors in Computing Systems. New York, NY: ACM; 2000.
36. Lee J, Kim J, Moon JY. What makes Internet users visit cyber stores again? Key design factors for customer loyalty. In: CHI 2000, The Future is Here: CHI 2000 Conference proceedings: Conference on Human Factors in Computing Systems. New York, NY: ACM; 2000.
37. Kruglanski AW, Webster DM, Klem A. Motivated resistance and openness to persuasion in the presence or absence of prior information. J Pers Soc Psychol 1993 Nov;65(5):861-876 [FREE Full text] [Medline: 8246114]
38. O'Neill SC, DeMarco T, Peshkin BN, Rogers S, Rispoli J, Brown K, et al. Tolerance for uncertainty and perceived risk among women receiving uninformative BRCA1/2 test results. Am J Med Genet C Semin Med Genet 2006 Nov 15;142C(4):251-259. [doi: 10.1002/ajmg.c.30104] [Medline: 17024668]
39. Patrick AS, Briggs P, Marsh S. Designing systems that people will trust. In: Cranor L, Garfinkel S, editors. Security and Usability: Designing Secure Systems That People Can Use. Sebastopol, CA: O’Reilly & Associates; 2005:99-127.
40. Schwarzer R. Self-efficacy in the adoption and maintenance of health behaviors: theoretical approaches and a new model. In: Schwarzer R, editor. Self-Efficacy: Thought Control Of Action. London, UK: Taylor & Francis; 1992:217-243.

Abbreviations

CFI: comparative fit index
EPPM: extended parallel process model
IP: Internet protocol
PMT: protection motivation theory
RMSEA: root mean square error of approximation

Edited by G Eysenbach; submitted 30.03.11; peer-reviewed by J Weert, D Zulman; comments to author 20.04.11; revised version received 15.06.11; accepted 23.05.11; published 27.07.11.

Please cite as:
Harris PR, Sillence E, Briggs P
Perceived Threat and Corroboration: Key Factors That Improve a Predictive Model of Trust in Internet-based Health Information and Advice
J Med Internet Res 2011;13(3):e51
URL: http://www.jmir.org/2011/3/e51/
doi:10.2196/jmir.1821
PMID:21795237

Original Paper

A Pilot Study of Motivational Interviewing Training in a Virtual World

Suzanne Mitchell[1], MD, MS; Robin Heyden[2], BS; Neil Heyden[3], BA; Paul Schroy[4], MD, MS; Stephen Andrew[5], LICSW; Ekaterina Sadikova[1]; John Wiecha[1], MD, MPH

[1]Boston University School of Medicine, Department of Family Medicine, Boston, MA, United States
[2]Education Consultant, Wellesley, MA, United States
[3]Educational Consultant, Wellesley, MA, United States
[4]Boston University School of Medicine, Department of Gastroenterology, Boston, MA, United States
[5]Health Education and Training Institute, Portland, ME, United States

Corresponding Author:
Suzanne Mitchell, MD, MS
Boston University School of Medicine
Department of Family Medicine
Dowling 5 South
771 Albany Street
Boston, MA, 02118
United States
Phone: 1 617 414 4465
Fax: 1 617 414 3345
Email: suzanne.mitchell@bmc.org

Abstract

Background: Motivational interviewing (MI) is an evidence-based, patient-centered counseling strategy proven to support patients seeking health behavior change. Yet the time and travel commitment for MI training is often a barrier to the adoption of MI by health care professionals. Virtual worlds such as Second Life (SL) are rapidly becoming part of the educational technology landscape and offer not only the potential to improve access to MI training but also to deepen the MI training experience through the use of immersive online environments. Despite SL's potential for medical education applications, little work is published studying its use for this purpose and still less is known of educational outcomes for physician training in MI using a virtual-world platform.

Objective: Our aims were to (1) explore the feasibility, acceptability, and effectiveness of a virtual-world platform for delivering MI training designed for physicians and (2) pilot test instructional designs using SL for MI training.

Methods: We designed and pilot tested an MI training program in the SL virtual world. We trained and enrolled 13 primary care physicians in a two-session, interactive program in SL on the use of MI for counseling patients about colorectal cancer screening. We measured self-reported changes in confidence and clinical practice patterns for counseling on colorectal cancer screening, and acceptability of the virtual-world learning environment and the MI instructional design. Effectiveness of the MI training was assessed by coding and scoring tape-recorded interviews with a blinded mock patient conducted pre- and post-training.

Results: A total of 13 physicians completed the training. Acceptability ratings for the MI training ranged from 4.1 to 4.7 on a 5-point scale. The SL learning environment was also highly rated, with 77% (n = 10) of the doctors reporting SL to be an effective educational medium. Learners' confidence and clinical practice patterns for colorectal cancer screening improved after training. Pre- to post-training mean confidence scores for the ability to elicit and address barriers to colorectal cancer screening (4.5 to 6.2, $P = .004$) and knowledge of decision-making psychology (4.5 to 5.7, $P = .02$) and behavior change psychology (4.9 to 6.2, $P = .02$) increased significantly. Global MI skills scores increased significantly and component scores for the MI skills also increased, with statistically significant improvements in 4 of the 5 component skills: empathy (3.12 to 3.85, $P = .001$), autonomy (3.07 to 3.85, $P < .001$), collaboration (2.88 to 3.46, $P = .02$), and evocative response (2.80 to 3.61, $P = .008$).

Conclusions: The results of this pilot study suggest that virtual worlds offer the potential for a new medical education pedagogy that will enhance learning outcomes for patient-centered communication skills training.

(J Med Internet Res 2011;13(3):e77) doi:10.2196/jmir.1825

KEYWORDS

Medical education; continuing medical education; computer-assisted instruction; computer-aided instruction; distance education; computer simulation; patient simulation; motivational interviewing; virtual world

Introduction

Patient-centered clinical skills, supported by versatile use of technology, enhance the clinical effectiveness of health care professionals. A cornerstone of effective health care is patient-centered communication and health behavior counseling, and the most widely studied approach is motivational interviewing (MI). Yet the intensive time and resource commitments for standard MI training often present barriers to the adoption of MI by health care professionals. Recently, however, the use of novel educational technologies such as virtual worlds has been showing great promise for overcoming access barriers to practical professional development training, such as is required for learning MI, and deepening the MI learning experience.

The popularity and scope of Web-based medical education curricula and continuing medical education (CME) programs has increased dramatically in recent years, but virtual-world venues are still not mainstreamed for these purposes [1-3]. Virtual worlds are rapidly becoming part of the general educational landscape and have untapped potential for meeting the unique and practical demands of medical education. Second Life (SL) is one of the best known virtual worlds, with over 300 colleges and universities staging regular events, seminars, and workshops in SL [4]. Most of these are education and awareness locations featuring kiosks and visual displays, health videos, slideshows, clinical skills simulation exercises, and Web links [5-11]. The advantages of a virtual-world venue such as SL for medical education include not only the convenience of a virtual-world venue that can be accessed from any physical location with an available computer, but also the immersive and 3-dimensionally realistic environment that creates a potentially richer learning experience than standard Web-based training courses.

What is SL? First launched in 2003, SL is an example of an immersive, 3-dimensional environment that supports a high level of social networking and interaction with information. Individuals can enter SL free of charge as avatars. In the SL virtual world, vast opportunities exist for student interaction, intense engagement, scripted immersive experiences, simulations, role-playing, and constructivist learning. The anonymity afforded by the avatar appears to lead to less inhibition and greater interaction. In addition, the greater sense of “presence” in a virtual world positively influences group process and cohesiveness, as well as engagement and attention [12]. We chose SL for our pilot study because it is the most widely used virtual-world platform, there is no access charge, and we had recent success conducting a CME event in SL on diabetes care management [13]. At any given time, 30,000 to 60,000 users are logged into SL. As of January 2009, there were 15 million registered SL accounts [14].

Little is known about the educational effectiveness of health professional training in SL or other virtual worlds, aside from our own published report in 2010 [13]. We published findings in the *Journal of Medical Internet Research* of our pilot study of a CME course conducted in SL on diabetes treatment. In our prior study, we focused on training physicians in a disease management skill. In the current study, we report findings from a new study whereby we explored the use of a virtual environment for training clinicians in MI counseling for colorectal cancer screening.

The purpose of this project was to explore the potential use of the SL virtual-world platform for training physicians in MI. The objectives of this pilot study were to (1) explore the feasibility, acceptability, and effectiveness of a virtual-world platform for delivering MI training designed for physicians, (2) pilot test instructional designs using SL for MI training, and (3) measure participant learning outcomes and feedback.

Methods

Study Sample and Recruitment

We targeted primary care residents and practicing physicians, although being a primary care practitioner was not required for participation. Participants were recruited from two family medicine listservs and from prior participants in our online courses. To be eligible for the pilot study, participants needed to communicate in English, be available on the dates of the synchronous SL sessions, and have computer equipment and capacity to support the SL platform. Given the anticipated time commitment, including training and pre- and post-training surveys, participants were offered an honorarium for completion of all activities. Since this was a research study and not an accredited CME program, no CME credits were provided. All participants took part in the experimental condition. Approval was obtained from Boston University School of Medicine institutional review board.

Instructional and Study Design

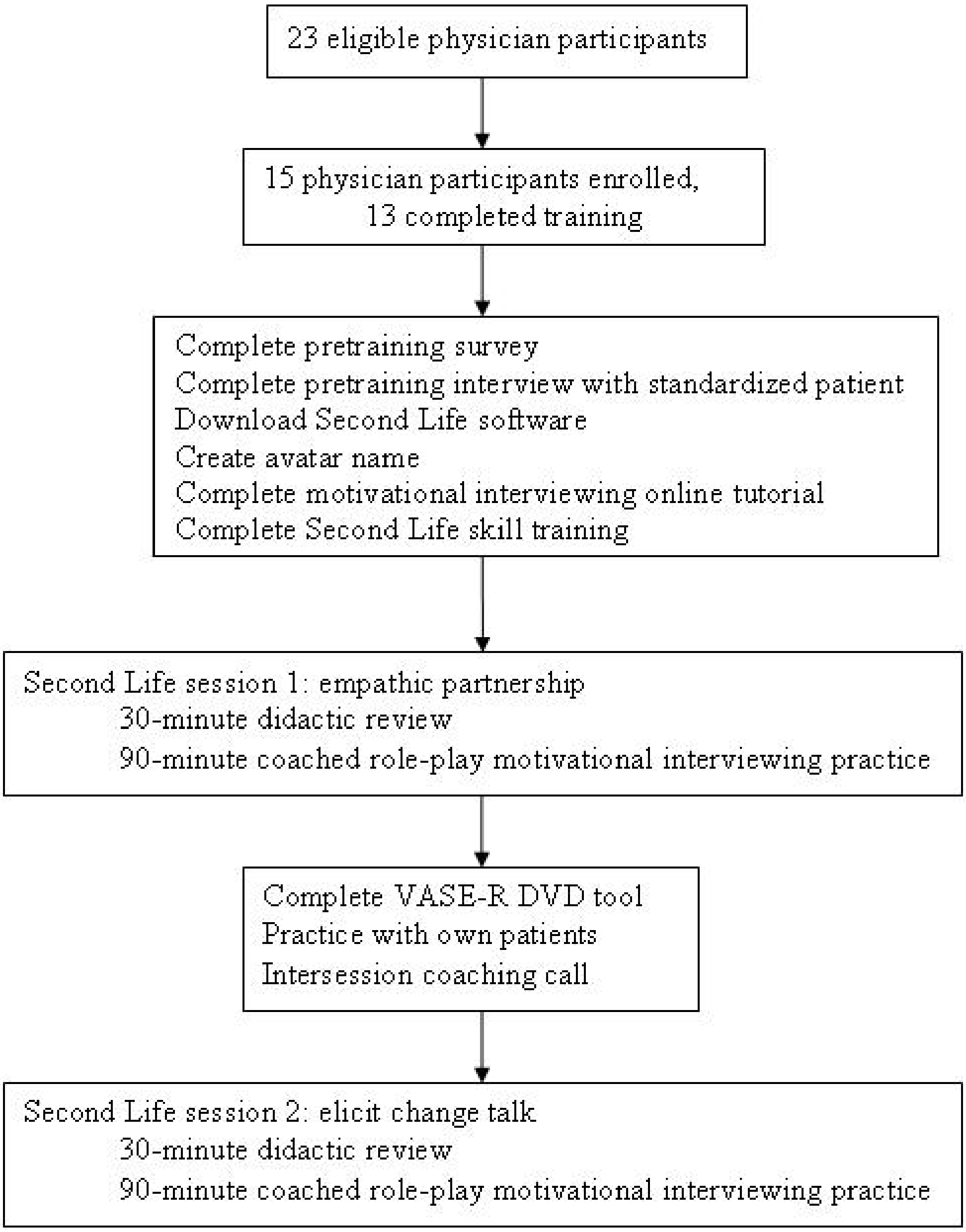
The investigators commissioned an expert MI trainer to provide the session content advice and coach role-play sessions. The instructional emphasis was to understand the conceptual framework underpinning MI and to learn the MI health counseling approach through practice with feedback and coaching, based on available evidence on effective MI training approaches [15]. We focused on two specific MI skills, *developing empathic partnership* and *eliciting change talk*, because of the evidence correlating behavior change with these counseling behaviors [16]. Empathic partnership represents the global spirit of MI and is the process whereby the clinician establishes understanding, cultivates trust with the patient, and creates a safe haven wherein the patient explores his or her ambivalence about undertaking behavior change actions. Eliciting change talk involves the process of encouraging patients to discuss and describe a desire and vision for the specific behavior change. Change talk is considered the precursor step toward taking action [17].

We applied mixed educational approaches to train primary care physicians in MI based on evidence of best training practices (see Figure 1) [15]. In preparation for each SL event, all study participants completed a 1-hour online tutorial in MI to gain a fundamental understanding of the MI philosophy and underlying conceptual model. For the in-world events we modified two existing MI slide decks typically used for live trainings, to support two 30-minute reviews of didactic information focused on *developing empathic partnership* and *eliciting change talk*. We modified the presentation to allow opportunities for interaction between the trainer and the learners using local text chat and other activity afforded by the virtual world such as a simulated visual element to leverage the unique capabilities of SL. In the second SL session, we also added a model MI interview after the didactic presentation, prior to role-play exercises. The interview was conducted by the MI trainer with one of the standardized patients to help demonstrate MI skills in a health counseling session. We developed two standardized patient scripts for use interchangeably by the standardized patients for mock interviews. The scripts focused on a discussion of colorectal cancer screening with a patient who is ambivalent about undergoing screening. The study participants were asked to engage the standardized patient in a discussion about colorectal cancer screening using the principles of MI.

Following the didactic review and model interview, we engaged in practical skills development. To enhance the adoption of new communication skills, we employed a practical skill-building approach using role-play with standardized patients. To make this feasible, we split the learners into two small groups of 6–7 persons each and teleported them to two separate platforms, each with an MI coach and a standardized patient. For the remaining 75–90 minutes, each participating physician took a turn interviewing and counseling the standardized patient on colorectal cancer screening and received immediate feedback from the MI coach and other study participants. Learners also engaged in active observation of peer participants' role-play experiences, feedback, and coaching, providing text chat feedback and commentary for one another.

Between sessions, we encouraged participants to practice the MI skills with their own patients. We also asked participants to complete the video assessment of simulated encounters-revised (VASE-R) DVD-based video assessment tool, developed by MI researchers at the University of Washington. The 18-item instrument includes three video-based vignettes, in which actors portray patients, with each vignette followed by questions that prompt learners to write responses that are then scored against MI standards [18]. The participants completed online answer sheets, which were scored by an independent and blinded MI expert, and we furnished results to participants as feedback. In addition, each participant received an individual coaching session by telephone with an expert MI coach. These coaching sessions were approximately 15 minutes in length and involved positive reinforcement, trouble-shooting of problem situations, and clarification of concepts if requested by the learner. In total, the training was designed to be completed in 8–10 hours of learners' time.

Figure 1. Recruitment and instructional design (VASE-R = video assessment of simulated encounters-revised).



The Virtual-World Venue

We started with an existing Boston University School of Medicine SL build (or "sim") constructed on a private, virtual island owned and developed by Boston University School of Medicine for an earlier joint project between the School and the World Health Organization [19]. Ownership of the island allowed the developers to control access and thus provide security and privacy for the attending physicians. If an SL venue is not private, there is a risk of random avatars potentially disrupting an event. Further details on the virtual-world venue are reported elsewhere [13] (see Figure 2).

The participating physicians were on their own computers in their homes or offices across the United States while the investigators and the speaker conducted the seminar in the same physical location. This allowed easy communication between the study staff during the event and convenient technical support for the speaker during the event. As insurance against sound problems (one of the more often encountered SL technical issues), Skype names were collected from all participants in advance so that a Skype conference call could be placed between the speaker's location and any participants experiencing sound problems [20]. During the two in-world events we had two technical support people available to help solve issues as they arose. Typical technical problems experienced by the participants included sound problems (cannot hear, cannot speak, hearing an echo, forgot to bring headphones), entering SL from a different computer and encountering firewall access problems, and trouble adjusting their camera view to read the slides. The seminar speaker was coached in one-on-one sessions conducted by phone. We conducted two dress rehearsals prior to the live SL training.

Figure 2. The virtual-world venue for motivational interviewing in Second Life training.

Training in SL

Following eligibility screening and enrollment, each participant was given information on how to access the online SL tutorials. The tutorials were constructed based on an essential skills checklist required for participation in the planned activities. Required SL skills included moving the avatar (walking, flying, or sitting), using camera controls for viewing, communicating using text chat, talking, and instant messaging, and troubleshooting technical problems. Participants were instructed to work their way through the online tutorials on their own [21]. If they had questions or experienced problems, they could email for help or telephone during announced "office hours." Prior to the first in-world event, each study participant took part in a short, scheduled skills assessment in order to determine their readiness for the training. Any problems or questions were addressed at that time.

Outcome Measures

We assessed feasibility and acceptability of recruitment, the instructional design, and the virtual-world learning environment using a Web-based self-report questionnaire. We measured limited effectiveness of the MI training by first comparing changes in pre- and post-training responses to a 16-item questionnaire exploring learners' self-reported confidence and clinical practice related to colorectal cancer screening. We measured MI skills proficiency by comparing the pre- and post-training scores from coded audiotaped mock interviews with a standardized patient. The scores were determined by a blinded, independent coder using the Motivational Interviewing Treatment Integrity, version 3.0 coding tool [22]. A difference of 1 on a 5-point scale between pre- and post-training scores was considered a clinically significant change in MI skills proficiency and was used for statistical testing as the relevant effect size [15]. According to this coding instrument, a global clinician score of 3.5 was used as the threshold score for proficiency and 4.0 indicated MI competency, out of a total possible score of 5. We calculated the mean scores for the core components and compared the mean global scores using the Wilcoxon signed rank test.

Results

Recruitment and Characteristics of Study Sample

Initially, 23 physicians expressed interest in the program. Following screening for eligibility, 15 of the original 23 participants enrolled in the study and 13 completed the training. After the first session, 2 enrolled participants dropped out due to conflicts with the workshop dates. The final sample comprised 2 male physicians and 11 female physicians, residing in 5 different states (Pennsylvania, Kansas, Massachusetts, Illinois, and New York). All 13 participants specialized in family medicine. Of the 13 participants, 10 had an average of 9 years of postresidency clinical practice experience (3 of the 13 are still enrolled in a residency program) and 5 are in academic practice. The majority of the participants (10/13, 77%) indicated they work in an environment with more than 10 other doctors, while 2 said they work with 4–6 doctors and 1 worked alone. Participants reported spending between 0 and 10 hours dedicated to the study training, with 4 participants spending more than 10 hours, 3 spending 6–10 hours, 3 spending 1–5 hours, and 3 spending 0 hours.

About a third (5/13, 38%) of the participants considered themselves “expert Internet users,” 62% (8/13) were steady users, and the majority (12/13, 92%) reported being “very comfortable” or “moderately comfortable” with using the computer. Only 2 participants said that they had experience with any virtual worlds other than SL. The sample was relatively evenly split between Mac (7/13, 54%) and PC users (6/13, 46%).

Of the 13 participating physicians, 3 had previous SL experience. The participants found SL to be an effective method of learning and reported they would like to take other courses in SL. They were mostly neutral or unsure about whether SL is a superior learning technique to face-to-face learning. The majority reported they would recommend SL to a colleague, strongly agreed that SL is effective, and exhibited interest in taking more courses with SL. Interestingly, 10 of the 13 participants reported that engaging in role-play was easier in SL than in face-to-face role-play practice.

Motivational Interviewing Course Acceptability Outcomes

Acceptability of the MI course content and training was positive overall. The participants rated the MI online tutorial with an average score of 4 (SD 0.9), or “very good,” out of 5,and the lecture on the *empathic partnership* at 4.1(SD 0.86). The *eliciting change talk* lecture was rated similarly, with a mean score of 4.4 (SD 0.63). The participants rated their experiences with mock patients in role-play interviews as well as the performance of their MI coaches. The first mock interview practice session, focusing on *developing empathic partnership*, received an average score of 4.7 (SD 0.63) out of 5 and the second, focusing on *eliciting change talk*, received a mean score of 4.6 (SD 0.87). Participants’ averaged scores for the VASE-R DVD learning tool was 4.1 (SD 1.12).

Educational Outcomes

Colorectal cancer screening practices and MI skills proficiency changes were measured using survey results and scores from participants’ coded mock interviews of colorectal cancer screening counseling with standardized patients (see Table 1 and Table 2). We examined changes in colorectal cancer screening practice experiences and MI skills proficiency changes. Table 1 shows the results comparing the participants’ pre- and post-training 12-item questionnaire on colorectal cancer screening practice experiences. Following the MI training in SL, the participants reported a significant increase in their knowledge and skills related to colorectal cancer screening counseling. A greater portion (9/13, 69% compared with 5/13, 39%) reported recommending routine colorectal cancer screening to patients over 50 years of age, and 77% (10/13) (compared with 5/13, 38% before training) became “very comfortable” talking to patients about colorectal cancer screenings. In 10 of the 12 survey items, the scores increased with mean differences that are equal to or exceed 1 point on a 7-point Likert scale, which, for the sake of detecting trends in these data, is significant. The participants’ reported ability to address the patients’ barriers to colorectal cancer screening increased significantly—on average, by 1.8 points (see Table 1).

Table 1. Participants’ self-reported mean change in knowledge and confidence scores in colorectal cancer screening counseling practices (N = 13 for each item)

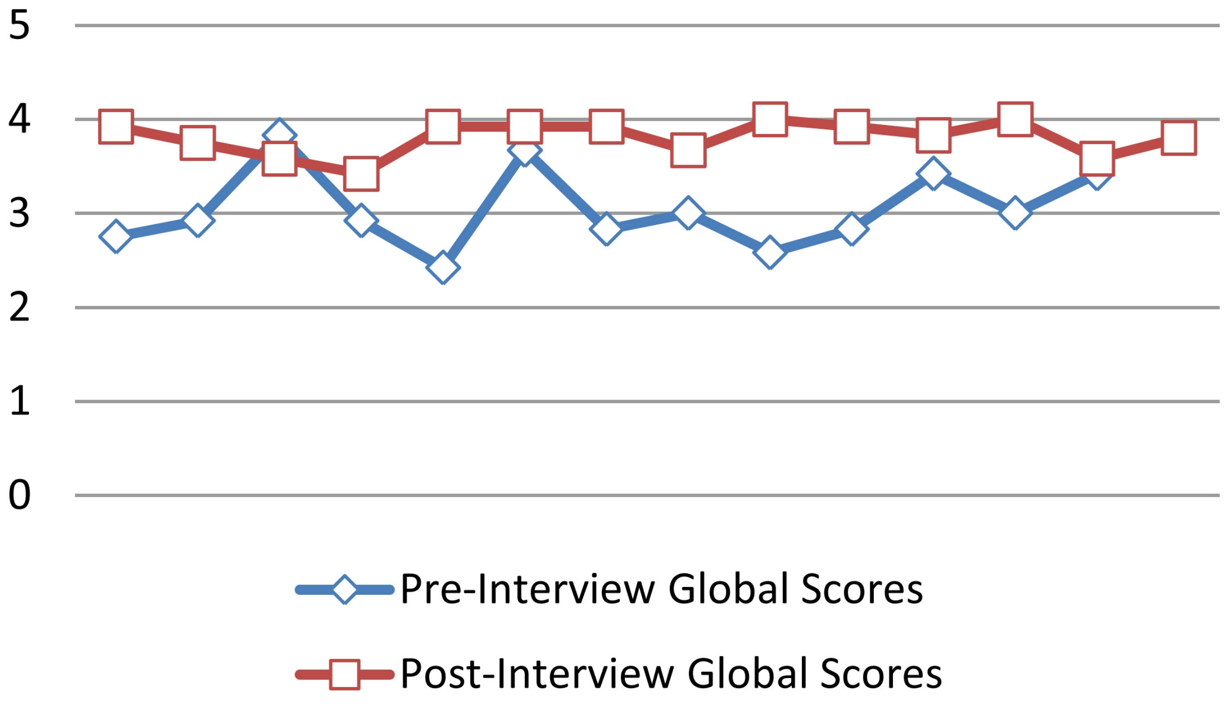
	Precourse mean	Postcourse mean	Difference	*P* value
Routine recommendation	5.69	6.23	0.54	.4
Comfortable discussing screening	5.75	6.46	1.08	.03
Effective help coping with concerns	5.15	6.15	1.00	.04
Successful motivation	4.15	5.77	1.62	.004
Effective help overcoming resistance	3.92	5.46	1.54	.007
Knowledge of psychology of decision	4.46	5.67	1.33	.02
Knowledge of health behavior change stages	4.85	6.23	1.38	.02
Communication to promote colorectal cancer screening	4.23	5.77	1.54	.01
Ability to elicit barriers	4.54	6.15	1.61	.004
Ability to address barriers	4.15	6.0	1.85	.004
Detection of ambivalence	4.42	5.85	1.42	.03
Screening follow-up	2.92	4.62	1.69	.007

Table 2. Mean pre- to post-training change in global scores for proficiency in motivational interviewing (N = 13 for each component)

	Mean pretraining score	Mean post-training score	Mean change	*P* value
Global	3.05	3.80	0.75	.001
Empathy	3.12	3.85	0.73	.001
Autonomy	3.07	3.85	0.78	<.001
Direction	3.50	3.50	0.00	–
Collaboration	2.88	3.46	0.58	.02
Evocative response	2.80	3.61	0.81	.008

In terms of measured MI proficiency, we found significant changes in MI skills based on comparison of audiotaped interview scores before and after MI training in SL (see Table 2 and Figure 3). Before training, only 15% (2/13) of participants received a global clinician score at or above the proficiency level of 3.5 out of a possible 5 points, with a score of 4 indicating MI competency. After the MI training in SL, 12 of the 13 participants scored at or above the MI proficiency level, and 2 participants received scores in the range of MI competency. Group mean scores for the global MI rating and each component score are listed in Table 2. Mean scores for 4 of the 5 component skills increased significantly following MI training (see Table 2): empathy (0.73, *P* = .001), autonomy (0.78, *P* < .001), collaboration (0.58, *P* = .02), and evocative response (0.81, *P* = .008). The evocative response score represents the elements of *eliciting change talk*. Mean global rating scores also increased significantly (0.75, *P* = .001).

Figure 3. Change in global motivational interviewing scores after training.



Discussion

Patient-centered communication and counseling skills, such as MI, are a cornerstone of clinically effective health care. Our results indicate that the virtual world holds tremendous promise as a venue for training physicians in these essential skills and improving access to such training for health care professionals. We demonstrated high acceptability and feasibility for conducting training in MI using the SL virtual-world platform. Further, despite a small pilot study sample size, we were able to demonstrate a significant and positive effect on educational outcomes for developing proficiency in global clinician ratings in use of MI and in colorectal cancer screening practices and counseling among 13 primary care providers.

This pilot study supports adapting traditional face-to-face training to virtual-world environments. We were successful in recruiting physicians to participate in this training experience, and each learner rated the experience favorably. Learners acquired the basic navigational skills for SL using a self-directed tutorial with the additional support of a technical skills assessment process that ensured sufficient proficiency in SL before the live learning aspect of the study period began. By using a virtual world, we were able to engage physicians from 6 different states simultaneously in this program. We did find it necessary to have all study staff in one physical location, but technical support personnel were located remotely in 2 different east coast states.

We also achieved excellent results in our limited measure of educational effectiveness in teaching MI in the virtual world. Our study protocol involved a total of 8–10 hours of active learning whereby learners were trained in the philosophy and conceptual model of MI and engaged in role-play for skills development, compared with the traditional 2- to 3-day in-person commitment typically required for essential MI training. We were able to demonstrate a statistically significant improvement in proficiency scores of recorded counseling interviews with a mock patient. While only 2 participants' scores demonstrate full competency in the global score rating for the MI method overall, we were able to achieve proficiency for 12 of the 13 participants (92%). This suggests that our 8- to 10-hour practical training approach in a virtual world is as effective as a 2- to 3-day in-person training in MI [15]. Finally, our learners reported significant improvements in confidence in colorectal cancer screening counseling after the MI training in SL.

Our training is focused on practical skill-building with a learning-by-doing educational approach. We developed this curriculum based on evidence from the Evaluating Methods for Motivational Enhancement Education clinical trial by Miller et al, indicating the significant impact of feedback and coaching on trainee proficiency in MI [15]. Our findings support this earlier work showing that practice of the MI method combined with coaching and feedback is the most effective approach to clinician skills acquisition. We introduced feedback at two key points in our training, during the SL sessions and between sessions with feedback for participants' responses on the VASE-R DVD tool. In addition, we provided opportunities for both individualized and group feedback and coaching. This method allowed participants to learn to identify key MI behaviors by observing others and to learn from observing feedback and coaching provided to peer learners, thus optimizing SL participation as an active learning opportunity. Our participants rated this approach highly, indicating the substantial benefit of learning from observing others engaging with the standardized patient. This approach is an adaptation from other MI training methods where feedback and coaching are often done individually.

Our study involved only physicians. It is clear, however, that other health care professionals and nonclinical staff could also benefit from virtual-world training opportunities. Our experience indicates that virtual worlds are amenable to communication skills training beyond health behavior counseling such as discussion about end-of-life and advance directives, communication about medical errors, patient-centered communication for nonclinical staff, and cultural competency training. Some of the unique aspects of a virtual-world training that are conducive to communication skills training are the anonymity of the avatar that facilitates role-play, the ability to simulate a deep, immersive, and realistic experience, and the enhancement of technology skills acquired through technology-based learning environments [23].

While our pilot study demonstrates the enormous potential of virtual-world environments for clinical communication skills training, there are several important limitations. Our study did not correlate the achieved MI skills proficiency with clinical outcomes. Thus, we do not know from our findings whether this level of MI training and achieved proficiency is sufficient to make clinically significant improvements in patient outcomes. Further, it is not evident that this training method is conducive to dissemination on a larger scale. The need for small-group learning and individualized coaching between sessions is certainly a potential limitation to cost effectiveness; however, if learners acquire skills proficiency more quickly with this instructional design, our training approach may prove cost effectiveness. Our next effort will involve a fully powered trial testing the effectiveness of this training to increase colorectal cancer screening rates.

We did experience several SL technical problems during our study period. The problems with using a virtual world like SL for education and training lie mostly in the realm of technical and security issues. The software requires a download and has significant system requirements (processing power, up-to-date video card, and a fast broadband Internet connection), the learning curve for navigation and interaction is steep, and the possibilities for technical problems and failures during the actual event are numerous. Many corporate or university firewalls do not allow access to public virtual worlds like SL.

Despite these limitations, 13 participants successfully completed our study. We learned that it is possible for this participant population to master sufficient virtual-world navigation skills without one-to-one intervention. Recognizing the significant support required for one-to-one training from our prior study, in this study we used a self-directed tutorial to support learners in acquiring SL navigation skills. By doing so, we dramatically reduced the time required for SL skills support, providing only a final skills verification session for each participant. Ultimately, all participants successfully learned the essential skills. On the other hand, the course facilitator was challenged with a high degree of multitasking while teaching. For example, the SL facilitator not only conducted a didactic learning session but also was simultaneously required to monitor the local text chat and respond to questions and comments in real time. In the role-play segment of our training, MI coaches were required to simultaneously navigate the avatar, listen to learners conduct interviews with the mock patient, monitor comments from the local chat of observing learners, and then provide feedback to the interviewer. Managing all of these tasks concurrently occasionally affected the momentum of the training but was not disruptive.

Conclusion

Overall, the SL virtual-world training in MI appears to have yielded significant increases in the confidence levels of doctors with respect to their abilities to recommend and guide their patients through the colorectal screening procedure and in their proficiency using MI strategies. These results warrant further research in this area, including future work to determine whether this educational approach can be expanded to a larger group of learners in a cost-effective manner and whether the achieved MI proficiency levels correlate with improved patient outcomes.

Acknowledgments

The authors would like to thank Jennalee Redmond and Elizabeth Dorland for their technical support on this project and Danielle Shiley, our research assistant, for supporting the administration of the research reported here.

Conflicts of Interest

None declared

References

1. HibiscusH. CNN iReport. 2008 Nov 17. AMIA Panel Discussion on Virtual Worlds Takes Place in RL and SL URL: http://ireport.cnn.com/docs/DOC-149486 [accessed 2011-06-07] [WebCite Cache ID 5zFrOFwwZ]
2. Kues J, Davis D, Colburn L, Fordis M, Silver I, Umuhoza O, Association of American Medical Colleges. Society for Academic Continuing Medical Education. 2009. Academic CME in North America: The 2008 AAMC/SACME Harrison Survey URL: http://www.sacme.org/Resources/Documents/SurveyResults/2008_survey_report.pdf [accessed 2011-08-18] [WebCite Cache ID 612EIsC61]
3. Van Harrison R, Davis DA, Fordis M, Kues JR, Manchul L, Mazmanian PE, et al. Society for Academic Continuing Medical Education. 2007 Jan. Survey for 2006: Descriptive Results URL: http://www.sacme.org/site/sacme/assets/pdf/BS_results_2006.pdf [accessed 2011-06-07] [WebCite Cache ID 5zFrhfdVV]
4. EDUCAUSE Learning Initiative. EDUCAUSE. 2011. 7 Things You Should Know About Second Life URL: http://www.educause.edu/ELI/7ThingsYouShouldKnowAboutSecon/163004 [accessed 2011-06-07] [WebCite Cache ID 5zFs0lAZJ]
5. Boulos MN, Hetherington L, Wheeler S. Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. Health Info Libr J 2007 Dec;24(4):233-245. [doi: 10.1111/j.1471-1842.2007.00733.x] [Medline: 18005298]
6. Gorini A, Gaggioli A, Vigna C, Riva G. A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008;10(3):e21 [FREE Full text] [doi: 10.2196/jmir.1029] [Medline: 18678557]
7. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [doi: 10.2196/jmir.1051] [Medline: 18762473]
8. Woodford P. Natl Rev Med. Montreal, Quebec CA: Parkhurst Publishing; 2007 Mar 30. Medicine's not-so-secret Second Life: public health education thrives in so-called "virtual worlds". URL: http://www.nationalreviewofmedicine.com/issue/2007/03_30/4_advances_medicine_6.html [accessed 2011-08-25] [WebCite Cache ID 61CoDRU0x]
9. Toro-Troconis M, Kamel Boulos MN. Musings on the state of '3-D Virtual Worlds for Health and Healthcare' in 2009. J Virtual Worlds Res 2009;2(2):3-15 [FREE Full text]
10. Lafsky M. Discover Magazine. Can training in Second Life teach doctors to save real lives? 2009 Jul 16 URL: http://discovermagazine.com/2009/jul-aug/15-can-medical-students-learn-to-save-real-lives-in-second-life/article_view?b_start:int=1&-C=
11. LeRoy Heinrichs W, Youngblood P, Harter PM, Dev P. Simulation for team training and assessment: case studies of online training with virtual worlds. World J Surg 2008 Feb;32(2):161-170. [doi: 10.1007/s00268-007-9354-2] [Medline: 18188640]
12. Taylor TL. Living digitally: embodiment in virtual worlds. In: Schroeder R, editor. The Social Life of Avatars. London: Springer; 2002:40-62.
13. Wiecha J, Heyden R, Sternthal E, Merialdi M. Learning in a virtual world: experience with using second life for medical education. J Med Internet Res 2010;12(1):e1 [FREE Full text] [doi: 10.2196/jmir.1337] [Medline: 20097652]
14. WebProNews Staff. WebProNews. 2009 Mar 16. Second Life Still Alive and Kicking URL: http://www.webpronews.com/second-life-still-alive-and-kicking-2009-03 [accessed 2011-06-07] [WebCite Cache ID 5zFtbh8pv]
15. Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. J Consult Clin Psychol 2004 Dec;72(6):1050-1062. [doi: 10.1037/0022-006X.72.6.1050] [Medline: 15612851]
16. Amrhein PC, Miller WR, Yahne CE, Palmer M, Fulcher L. Client commitment language during motivational interviewing predicts drug use outcomes. J Consult Clin Psychol 2003 Oct;71(5):862-878. [doi: 10.1037/0022-006X.71.5.862] [Medline: 14516235]
17. Miller WR, Rollnick S, Conforti K. Motivational Interviewing, Second Edition: Preparing People for Change. New York, NY: The Guilford Press; 2002.
18. Rosengren DB, Hartzler B, Baer JS, Wells EA, Dunn CW. The video assessment of simulated encounters-revised (VASE-R): reliability and validity of a revised measure of motivational interviewing skills. Drug Alcohol Depend 2008 Sep 1;97(1-2):130-138. [doi: 10.1016/j.drugalcdep.2008.03.018] [Medline: 18499356]
19. Linden Research I. SLurl. 2011. Second Life URL: http://slurl.com/secondlife/Teaching%209/12/169/22 [accessed 2011-06-07] [WebCite Cache ID 5zFtlIjzS]

20. Skype Limited. Skype.com. 2011. Skype Homepage URL: http://www.skype.com/intl/en-us/home [accessed 2011-06-07] [WebCite Cache ID 5zFtrUkfN]
21. Tangient LLC. Wikispaces. 2011. Getting Started With Second Life URL: http://gettingstartedwithsl.wikispaces.com/ [accessed 2011-06-07] [WebCite Cache ID 5zFu137tp]
22. Center on Alcoholism, Substance Abuse, and Addictions. The University of New Mexico. 2011 Aug 17. Coding Instruments URL: http://casaa.unm.edu/codinginst.html [accessed 2011-06-07] [WebCite Cache ID 5zFu6nUDU]
23. Dede C. Immersive interfaces for engagement and learning. Science 2009 Jan 2;323(5910):66-69 [FREE Full text] [doi: 10.1126/science.1167311] [Medline: 19119219]

Abbreviations

CME: continuing medical education
MI: motivational interviewing
SL: Second Life
VASE-R: video assessment of simulated encounters-revised

Edited by G Eysenbach; submitted 12.04.11; peer-reviewed by D Mohr, P Mills, JM Suelves, G Riva; comments to author 29.04.11; revised version received 12.06.11; accepted 14.06.11; published 26.09.11.

Please cite as:
Mitchell S, Heyden R, Heyden N, Schroy P, Andrew S, Sadikova E, Wiecha J
A Pilot Study of Motivational Interviewing Training in a Virtual World
J Med Internet Res 2011;13(3):e77
URL: http://www.jmir.org/2011/3/e77/
doi:10.2196/jmir.1825
PMID:21946183

Original Paper

Internet Use and eHealth Literacy of Low-Income Parents Whose Children Have Special Health Care Needs

Caprice Knapp[1], PhD; Vanessa Madden[1], BSC; Hua Wang[1], MS; Phyllis Sloyer[2], RN PhD; Elizabeth Shenkman[1], PhD

[1] University of Florida, Department of Health Outcomes and Policy, Gainesville, FL, United States
[2] Florida Department of Health, Children's Medical Services Network, Tallahassee, FL, United States

Corresponding Author:
Caprice Knapp, PhD
University of Florida
Department of Health Outcomes and Policy
1329 SW 16th Street
Room 5130
Gainesville, FL, 32610
United States
Phone: 1 352 265 7220
Fax: 1 352 265 7221
Email: cak@ichp.ufl.edu

Abstract

Background: The Internet has revolutionized the way in which many Americans search for health care information. Unfortunately, being able to use the Internet for this purpose is predicated on having access to the Internet and being able to understand and comprehend online health information. This is especially important for parents of children with special health care needs who are forced to make many medical decisions throughout the lives of their children. Yet, no information is available about this vulnerable group.

Objective: For parents of children with special health care needs we sought to (1) describe their Internet access and use, (2) determine which child and household factors were associated with Internet use, (3) describe eHealth literacy of Internet users, and (4) determine which child and household factors were associated with greater eHealth literacy.

Methods: This was a cross-sectional telephone survey of 2371 parents whose children with special health care needs were enrolled in Florida’s Medicaid and State Children’s Health Insurance Plan (SCHIP) programs (4072 parents were approached). To be enrolled in the program, families must have incomes that are less than or equal to 200% of the federal poverty level. The eHealth Literacy Scale (eHEALS) was used to measure eHealth literacy. Descriptive and multivariate analyses were conducted to address the study objectives.

Results: The survey response rate was 58.2%. Participating parents were mainly female (2154/2371, 91%), white non-Hispanic (915/2371, 39%), English speaking (1827/2371, 77%), high school graduates (721/2371, 30%), married (1252/2371, 53%), and living in a two-parent household (1212/2371, 51%). Additionally, 82% of parents (1945/2371) in the sample reported that they used the Internet, and 49% of those parents used it daily (1158/2371). Almost three-quarters of Internet users had access to the Internet at home while about one-half had access at work. Parents who were African American, non-English speaking, older, and not college graduates were less likely to use the Internet than their referent groups ($P < .001$). About 74% of Internet users (1448/1945) reported that they knew how to find health information for their children. However, only about one-half (1030/1945) reported that they can tell high quality from low quality resources online or that they feel confident in using information accessed online to make health decisions. Multivariate regression results consistently showed that being a non-English speaker, having less than a high school education, and being older were all significantly associated with lower eHealth literacy (all $P < .001$).

Conclusion: Low-income parents of children with special health care needs have access to and use the Internet as a source of information about their children's health. However, some parents are unable to distinguish between high and low quality information and are not confident in using the Internet. This information is timely because as the pressure to use the Internet to empower consumers and exchange information increases, issues related to access and disparities must be better understood.

(J Med Internet Res 2011;13(3):e75) doi:10.2196/jmir.1697

KEYWORDS

Children; Internet; Medicaid

Introduction

There is no doubt that the World Wide Web has significantly impacted the world since the mid-1990s. Information that was once available to only those with time, money, and knowledge is now available at the click of a button to those who have access to the Internet. The Pew Research Center's Pew Internet and American Life Project is perhaps the most comprehensive series of research on how access to, and use of, the Internet has evolved since the early 2000s. The Pew Research Center's first report in 2000 focused on how women used the Internet. Findings from the report noted that about 26 million Americans had used the Internet to keep in touch with a relative that they previously had not previously been in touch with [1]. More recent studies have contemplated the “digital divide." This divide refers to the differential between those who access the Internet and those who do not [2]. The divide has been documented in the United States as well other European countries [3,4]. Several of the Pew Research Center's reports suggest that the digital divide is narrowing over time as the number of mobile devices increases and the broadband population becomes more diverse, especially within the African American community [5]. Lorence, however, has suggested that the divide persists with the emergence of a “digitally underserved” group [6,7].

Likewise, the impact of the World Wide Web on health care has also been radical. Consumers have flocked to the Internet to search for information on diagnoses, treatment regimes, and prognoses. Findings from the Pew Research Center's 2008 nationally representative telephone survey suggest that 61% of Americans use the Internet to find health information, and 60% say that information they found online has impacted a health care decision they made [8]. These results have been corroborated by peer-reviewed studies. McInnes et al found that 29% of veterans had searched for health information online [9]. Health information searching was directly associated with higher levels of education, living in an urban area, and decreased health status. Lea and colleagues studied patients with head and neck cancer who received care at a comprehensive cancer center [10]. Using the computer was associated with increased educational attainment and income but younger age. Walsh conducted a survey with 1784 cancer patients and also found that use of the computer increased with education and income but decreased with age [11].

In pediatrics, there is limited information on Internet use or eHealth literacy, which is defined as the ability to “locate, evaluate, integrate, and apply information gleaned from electronic platforms” [12]. Mackert et al used quantitative methods to study eHealth literacy. They conducted focus groups with low literacy and culturally diverse parents and found that some avoided .edu or .gov websites because such websites are viewed as too complex and that some parents had expressed a lack of trust in government websites [13]. Quantitative methods have also been used. Kind et al surveyed 260 African American parents and found that greater Internet use and access were associated with higher educational attainment and income [14].

Other studies have documented parents’ Internet information seeking activities related to specific diagnoses including genetic counseling, hearing loss, and late effects of cancer treatment [15-17]. Recently, a few studies have emerged which used the eHealth Literacy Scale (eHEALS) to assess eHealth literacy in pediatrics [12]. Knapp and colleagues used this scale to measure eHealth literacy of parents whose children were enrolled in a pediatric palliative care program in Florida. Survey data from 129 parents showed that not having graduated from high school was associated with lower eHealth literacy and using the Internet as the primary information source (as opposed to their child's physician) was associated with higher eHealth literacy [18]. Brown and Dickson used the eHEALS to assess eHealth literacy of health care students in a master's program. They found that the students had high eHealth literacy but lacked the confidence in using information found online to make health care decisions [19]. Finally, the eHEALS was used in a 2010 intervention study to determine the intervention's effectiveness on HIV-positive participants’ ability to access information online. The intervention was found to be associated with positive and sustaining increases in eHealth literacy [20].

Although not the focus of these few studies, it is especially important to investigate the online information seeking behaviors and eHealth literacy for the approximately 13.9% of parents in the United States whose children have special health care needs [21]. Children with special health care needs are defined as “those who have or are at increased risk for a chronic physical, developmental, behavioral, or emotional condition and who also require health and related services of a type or amount beyond that required by children generally” [21]. These children face a number of medical decisions throughout their lives, forcing their parents to routinely seek out and compile large amounts of medical information on their behalf. Oftentimes parents use the Internet to locate, process, and use information. A recent statement by the American Association of Pediatrics (AAP) urged policymakers to take into account the needs of all stakeholders when implementing online and electronic resources, including families [22]. AAP suggested that information technology education and training should be available for patients and families so that they can be involved in decision making and become empowered.

To our knowledge, no studies have sought to describe Internet use and eHealth literacy of parents of children with special health care needs. Our study addresses these gaps in knowledge. We aimed to (1) describe Internet access and use patterns among low-income parents whose children have special health care needs, (2) describe the factors associated with Internet use, (3) describe the eHealth literacy of the parental Internet users, and (4) determine which factors are associated with greater eHealth literacy. We hypothesized that there would be education and age disparities associated with Internet use and eHealth literacy.

Methods

Sample

Study participants were parents of children with special health care needs aged 1 to 21 years enrolled in Florida's Children's Medical Services Network (CMSN) program, the state's Title V program. Title V of the Social Security Act allows for states to receive block grants to improve maternal and child health. CMSN has a network of primary and specialty care providers and each child enrolled in CMSN is assigned a care coordinator. All children with special health care needs enrolled in the program are eligible for Medicaid or the State Children's Health Insurance Plan (SCHIP) and their families have incomes less than or equal to 200% of the federal poverty level. Children must be certified by a physician as having a special health care need.

A random sample of parents whose children were currently enrolled in CMSN were each sent a letter explaining that someone may call them to participate in the study. Telephone surveys were conducted in English and Spanish beginning in July 2009 and ending in October 2009 using the Windows based computer assisted telephone interviewing system, WinCATI (Sawtooth Technologies, Northbrook, IL). Overall, 2371 surveys were completed. The University of Florida's Institutional Review Board approved this study.

Outcome Measures

The two aims of this study were to determine parents' use of and access to the Internet, as well as the eHealth literacy of Internet users. We asked parents if they had ever used the Internet, and if so, what was their frequency of use (daily, weekly, or less often than weekly). Parents were asked where they accessed the Internet (work, home, or mobile device). To determine the eHealth literacy of Internet users, the eHEALS scale was used. The eHEALS scale measures the "ability to locate, evaluate, integrate, and apply information gained from electronic platforms"[12]. There are eight items on the eHEALS scale that measure consumers' perceived information technology skills. The response categories determine the level of agreement (agree, undecided, disagree) with the eight statements about online health information. Psychometric testing on the eHEALS has revealed high internal consistency (Cronbach alpha 0.88) [12].

Factors

The final study aims were to explore what parent, child, and household factors are associated with Internet use and greater eHealth literacy. Several questions were included on the survey to explore these factors including parent's age, parent's race/ethnicity, parent's gender, parental language spoken at home, parent's marital status, parent's educational attainment, type of household (single or two parent), child's age, and child's health status. Children's health status was gauged by asking parents to rate their children's health status as excellent, very good, good, fair, or poor.

Analyses

Descriptive analyses were conducted to describe Internet use, how users accessed the Internet, and responses to the eight items from the eHEALS scale. Multivariate analyses were conducted to explore which child and household factors were associated with Internet use and greater eHealth literacy. A multivariate logistic regression was conducted to determine factors associated with Internet use. In this regression, the dependent variable was a binary indicator equal to one if parents responded that they used the Internet and zero otherwise. Finally, eight multivariate ordinal logistic regressions were performed using the responses to each of the eight statements in the eHEALS scale. Ordinal logistic was chosen because the response categories for each of the eight statements are ordinal (agree, undecided, disagree). STATA version 10.0 (StataCorp LP, College Station, TX) was used to perform the analyses [23].

Results

Sample Characteristics

Survey response rate was 58.2% (4072 had valid contact information and were approached, and, of those, 2371 completed the survey). More than one-half (1212/2371, 51%) of parents lived in a two-parent household; 1252 out of 2371 (53%) were married; 1827 out of 2371 (77%) primarily spoke English; 915 out of 2371 (39%) were white non-Hispanic; 721 out of 2371 (30%) had a high school diploma; and 2154 out of 2371 (91%) of the respondents were female (Table 1). Of all participating parents, 21% (491/2371) reported that their children were in fair or poor health, 914 out of 2,371 (39%), good health, and 943 out of 2371 (40%), excellent or very good health. Parental mean age was 40.5 (SD 10.2 years) and mean age of their children was 10.5 (SD 4.9 years).

Table 1. Summary statistics

Variable	n	% of 2371
Parent's gender		
Female	2154	90.8%
Male	217	9.2%
Missing	0	0.0%
Parent's race/ethnicity		
White non-Hispanic	915	38.6%
Hispanic	688	29.0%
African American non-Hispanic	623	26.3%
Other	125	5.3%
Missing	20	0.8%
Parental language spoken at home		
English	1827	77.1%
Non-English	541	22.8%
Missing	3	0.1%
Parent's educational attainment		
Less than high school	458	19.3%
High school graduate	721	30.4%
Some college	569	24.0%
College graduate	597	25.2%
Missing	26	1.1%
Parents marital status		
Not married	1096	46.2%
Married	1252	52.8%
Missing	23	1.0%
Type of household		
Single parent	1127	47.5%
Two-parent	1212	51.1%
Missing	32	1.3%
Child's health		
Excellent/very good	943	39.8%
Good	914	38.5%
Fair/poor	491	20.7%
Missing	23	1.0%

Internet Use

Overall, 82% of all parents (1945/2371) reported that they used the Internet and 426 out of 2371 (18%) of parents reported that they never used the Internet (Table 2). Of the Internet users, about one-half accessed the Internet or email on a daily basis. Most parents had access to the Internet from home (1681/2371, 71%), 48% had access from work (1143/2371), and 43% of all parents had access from both home and at work (1015/2371). Additionally, 624 out of 2371 (26%) parents had accessed the Internet or email from mobile devices.

Table 2. Internet use and access

Variable Use	n	% of 2371
Frequency of Internet or email use		
Daily	1158	48.8%
Weekly	488	20.6%
Less often than weekly	299	12.6%
Never	426	18.0%
Location		
Internet access from home	1681	70.9%
Internet access from work	1143	48.2%
Internet access from home and work	1015	42.8%
Ever used cell phone or BlackBerry to access Internet or email		
Yes	624	26.3%
No	1747	73.7%

Multivariate Analysis: Internet Use

A logistic regression was performed where the dependent variable was equal to one to indicate Internet use and zero otherwise (Table 3). Parents of African American race, non-English speaking parents, older parents, and parents with less than a college education were less likely to use the Internet. Parents who were married and had a child with excellent or very good health were more likely to use the Internet.

Table 3. Multivariate logistic regression

Independent variables[a]	(Dependent Variable: Internet Use) Coefficient Estimate	95% Upper Confidence Interval	95% Lower Confidence Interval	*P* value
Parent's gender				
Male	0.81	0.53	1.26	.36
Parent's race/ethnicity				
Hispanic	0.78	0.50	1.20	.25
African American non-Hispanic	0.46	0.33	0.64	< .001
Other	1.25	0.55	2.83	.60
Parental language spoken at home				
Non-English	0.42	0.28	0.62	< .001
Parent's educational attainment				
Less than high school	0.06	0.04	0.09	< .001
High school graduate	0.16	0.10	0.25	< .001
Some college	0.32	0.20	0.52	< .001
Parent's marital status				
Married	1.44	0.98	2.13	.07
Type of household				
Two-parent household	1.09	0.74	1.62	.66
Child's health				
Excellent/very good	1.44	1.04	1.99	.03
Good	1.50	1.10	2.04	.011
Age				
Parent's age (years)	0.94	0.93	0.95	< .001
Child's age (years)	1.01	0.98	1.04	.48

[a] Referent groups: female, white non-Hispanic, English speaking, college graduate, not married, two-parent household, fair/poor health

Responses to eHealth Literacy Scale Items

Table 4 shows the response frequencies for each eHEALS item for those 1945 parents who said they used the Internet. Parents who never used the Internet were not asked the eHEALS items.

Table 4. Response frequencies to eHEALS items

eHEALS item	Agree n	Agree % of 1945	Undecided n	Undecided % of 1945	Disagree n	Disagree % of 1945
1. I know what health resources are available on the Internet.	1179	60.6%	486	25.0%	280	14.4%
2. I know where to find helpful health resources on the Internet.	1273	65.4%	371	19.1%	301	15.5%
3. I know how to find helpful health resources on the Internet.	1358	69.8%	309	15.9%	278	14.3%
4. I know how to use the Internet to answer my questions about my child's health.	1409	72.4%	297	15.3%	239	12.3%
5. I know how to use the health information I find on the Internet to help my child.	1448	74.4%	286	14.7%	211	10.8%
6. I have the skills I need to evaluate the health resources I find on the Internet.	1364	70.1%	325	16.7%	256	13.2%
7. I can tell high quality health resources from low quality health resources on the internet.	1030	53.0%	501	25.8%	414	21.3%
8. I feel confident in using information from the Internet to make health decisions.	1024	52.6%	466	24.0%	455	23.4%

Response categories are grouped into agree (including strongly agree and agree), undecided, or disagree (including strongly disagree and disagree). Missing values were recoded as undecided.

The statement “I know how to use the health information I find on the Internet to help my child” had the highest level of agreement (1448/1945, 74%). The two statements that parents had the highest level of disagreement with were related to confidence in using information received from the Internet to make health decisions (455/1945 or 23.4% disagreed with eHEALS item 8) and ability to distinguish between high and low quality information (414/1945 or 21.3% disagreed with eHeals item 7).

Although not presented in the table, bivariate analyses were conducted to determine if there were significant differences between parents who were confident in using health information versus those who were not confident and parents who were and were not able to distinguish the quality of health information. In regard to confidence in using health information, significant differences ($P < .05$) were realized. More confident parents were English speaking parents, younger parents, parents of younger children, and parents whose children had excellent to very good health versus the respective referent groups. In regard to distinguishing between high and low quality health information, significant differences ($P < .05$) were realized. Parents better able to make the distinction had higher levels of education, were younger, were parents of younger children, and had children with excellent to very good health versus the respective referent groups.

Multivariate Analysis: eHealth Literacy

Performed were eight ordinal logistic regressions where the dependent variable represented the levels of agreement with each eHEALS statement (Tables 5). Observations where the dependent variable was missing were dropped. It is important to use this model since the response categories of agree, undecided, and disagree have an ordered nature. Results from all the eight regressions indicate that parental language, parental lower educational attainment, and older parental age were all consistently and significantly associated with lower levels of agreement with the eight eHEALS statements. For example, results from the statement “I know what resources are available on the Internet” imply that non-English speaking parents were about 35% less likely to be in a higher agreement category versus their English speaking peers. Results from the last two regressions are especially important to note given the low percentage of parents who agreed with these statement. For the statement “I can tell high quality health resources from low quality health resources on the Internet,” parents of another race, parents with less than a college degree, parents living in a two-parent household, and older parents were all significantly less likely to be in higher agreement. It is interesting that this is the only statement where non-English speaking was not significantly associated with higher agreement. For the statement “I feel confident in using information from the Internet to make health decisions,” Hispanic parents, non-English speaking parents, parents with less than a high school degree, parents who were married, parents living in a two-parent household, parents having a child with excellent/good health, and parents who were older were all significantly less likely to be in higher agreement.

Table 5. Multivariate ordered logit regressions (eHEALs items 1 through 4)

Independent variables[a]	Dependent Variables eHEALS Item 1 *Coefficient Estimate*	 eHEALS Item 1 *P* Value	eHEALS Item 2 *Coefficient Estimate*	eHEALS Item 2 *P* Value	eHEALS Item 3 *Coefficient Estimate*	eHEALS Item 3 *P* Value	eHEALS Item 4 *Coefficient Estimate*	eHEALS Item 4 *P* Value
Parent's gender								
Male	1.123	.49	1.107	.56	1.069	.72	0.912	.62
Parent's race/ethnicity								
Hispanic	0.922	.59	1.008	.96	0.882	.45	1.063	.72
African American	0.689	.003	0.627	<.001	0.663	.003	0.863	.30
Other	0.659	.05	0.819	.39	0.873	.58	1.114	.67
Parental language spoken at home								
Non-English	0.652	.005	0.571	<.001	0.530	<.001	0.452	<.001
Parent's educational attainment								
Less than high school	0.564	<.001	0.601	.001	0.576	.001	0.492	<.001
High school graduate	0.774	.04	0.792	.07	0.711	.01	0.710	.02
Some college	0.869	.28	0.924	.56	0.867	.32	0.742	.04
Parent's marital status								
Married	0.777	.12	0.993	.97	1.036	.84	0.838	.33
Household type								
Two-parent household	0.702	.03	0.848	.32	0.862	.39	0.847	.36
Child's health								
Excellent/very good	1.229	.11	1.187	.21	1.143	.35	1.185	.24
Good	1.217	.13	1.022	.87	1.051	.72	1.061	.68
Age								
Parent's age (years)	0.986	.02	0.981	.001	0.986	.03	0.978	<.001
Child's age (years)	0.989	.34	0.988	.31	0.976	.049	0.984	.21

[a] Referent groups: female, white non-Hispanic, English speaking, college graduate, not married, two-parent household, fair/poor health

Table 6. Multivariate ordered logit regressions (eHEALS items 5 through 8)

	Dependent Variables							
	eHEALS Item 5		eHEALS Item 6		eHEALS Item 7		eHEALS Item 8	
Independent variables[a]	Coefficient Estimate	*P* Value	Coefficient Estimate	*P* Value	Coefficient Estimate	*P* Value	Coefficient Estimate	*P* Value
Parent's gender								
Male	1.009	.96	0.761	.12	1.047	.77	1.258	.16
Parent's race/ethnicity								
Hispanic	1.089	.63	1.067	.70	1.083	.57	1.342	.04
African American	0.949	.72	0.903	.47	0.991	.94	0.980	.87
Other	1.073	.77	0.863	.54	0.616	.02	0.816	.33
Parental language spoken at home								
Non-English	0.542	<.001	0.418	<.001	1.097	.54	0.616	.001
Parent's educational attainment								
Less than high school	0.527	<.001	0.432	<.001	0.497	<.001	0.654	.003
High school graduate	0.770	.07	0.630	.001	0.576	<.001	0.863	.20
Some college	0.888	.43	0.758	.06	0.652	.001	0.837	.14
Parent's marital status								
Married	0.761	.14	0.783	.18	0.760	.08	0.614	.002
Household type								
Two-parent household	0.761	.14	0.716	.07	0.704	.02	0.656	.008
Child's health								
Excellent/very good	1.326	.06	1.624	.001	1.245	.08	1.398	.008
Good	1.078	.60	1.257	.10	1.091	.48	1.192	.16
Age								
Parent's age (years)	0.979	.001	0.969	<.001	0.985	.009	0.978	<.001
Child's age (years)	0.981	.13	0.987	.28	0.980	.06	0.992	.47

[a] Referent groups: female, white non-Hispanic, English speaking, college graduate, not married, two parent household, fair/poor health

Discussion

To our knowledge, our study is the first to focus on Internet use and eHealth literacy of parents of children with special health care needs. This study is unique in the population surveyed, the number of completed surveys, the ability to test for disparities, and the focus on children with special health care needs. Our study suggests that most parents have access to the Internet and use it on a daily basis, and most know how to find health information for their child. However, parents are concerned that they are unable to distinguish between high and low quality information online. Our findings expand the extant knowledge in the following ways.

First, our findings allow us to comment on access to the Internet for low-income parents of children with special health care needs. Parents of children with special health care needs are likely to have greater impetus to seek out health information compared with parents whose children do not have special health care needs. As a result, it is important that the parents' information needs be met, and the Internet may serve as a convenient, low-cost repository of information for these parents if they have access. Our results show that 1945 of 2371 (82%) parents in our sample use the Internet and most (1681/2371, 71%) have access to the Internet at home. Compared with the Pew Research Center's Internet and American Life Project study, which found that 57% of adults with household incomes lower than $30,000 use the Internet, parents in our study seemed to have greater Internet use [8]. Results from our study are more aligned with the Pew Research Center's findings that approximately 74% of the general population has gone online to access the Internet, World Wide Web, or to receive email [8]. Interestingly, 624 out of 2371 (26%) parents in our sample have used mobile devices to access the Internet. Findings from the Pew Research Center's report Mobile Access 2010 showed that 39% of the adult population are *motivated* mobile users and that African Americans and non-English speakers are among the highest users of the mobile Web [24]. Future studies should determine if the lower mobile use trends demonstrated in our sample are due to differences in income.

Second, our findings provide new insights into the factors associated with parental Internet use. Our findings corroborate

Kind et al [14] who also found that higher levels of education were associated with greater Internet use for African American parents. Our additional findings are novel to the literature on parental Internet use, although they do corroborate existing information on Internet use in other subgroups. For example, we found that African American parents of children with special health care needs were less likely to have used the Internet. Although new evidence in the literature for children with special health care needs, this has been corroborated in other studies of the digital divide. Our findings concerning African American parents may be explained in part by broadband access trends. There has been a distinct gap in broadband access between African Americans and whites, although this gap has slightly narrowed in the past few years (from a 19 point gap in 2009 to an 11 point gap in 2010) [5]. Our results that older parents, less educated, and non-English speaking parents are less likely to use the Internet present opportunities for interventions. Salovey et al [25] described the creation of a community-based computer center designed to improve computer literacy of Latino and African American parents whose children were enrolled in the Head Start program. The study suggests that parental knowledge was improved. Perhaps an intervention similar to this could be used with the population in our study, parents whose children with special health care needs enrolled in Florida's Medicaid and SCHIP programs, if these parents are interested in using the Internet, but do not have the skills to do so. Increasing Internet use for this vulnerable group is critical as many national initiatives push for the implementation of email communication between parents and providers, the adoption of electronic personal health records, and online education tools [26,27].

Third, our findings contribute to the literature on eHealth literacy. To our knowledge, our study is the first to assess the eHealth literacy of parents of children with special health care needs, and we used eHEALS, a validated survey instrument. Other studies that have explored eHealth literacy have relied on qualitative methods or nonvalidated questionnaires to understand eHealth literacy; therefore, it is difficult to compare our findings with those of other studies. However, our findings that parents may not being able to decipher the quality of online information and lack of confidence in using information to make decisions can lead to positive and negative effects on the delivery of care. For example, lack of confidence and inability to distinguish quality information may prompt parents to follow up with nurses and physicians during the health care encounter, leading to new dialogue and improved shared decision making between parents and providers. However, this may also lead to increases in competing demands on providers' time since it would not be possible to discuss the breadth of information that parents have discovered during a single encounter. Results from the Pew Research Center indicate that 53% of adults said their most recent health search led them to seek out a second opinion or to ask their physician new questions [8]. Of course, providers routinely give educational materials to families that include links to recommended websites. Yet, for low-income families, additional interventions may be necessary such as question and answer sessions with a case manager or care coordinator to build eHealth skills and confidence.

Fourth, results from the multivariate analysis show that there is a significant, negative association between most of the eight components of eHealth literacy and not speaking English, lower educational attainment, and being an older parent. Although these findings may not seem particularly surprising, the results of the other factors were surprising. For example, Hispanic parents are about 34% more likely than their white non-Hispanic counterparts to report higher agreement with feeling confident in using information found online. Prior evidence has noted no difference between Hispanic white and non-Hispanic white parents' trust in physicians, and our results imply that perhaps this trust translates to online health sources whereby Hispanic parents are just as confident in what they find [28]. African American parents in our sample were significantly less likely to locate information on the Internet, but they reported no less agreement in their ability to use or distinguish the quality of information compared with white parents of children with special health care needs. Perhaps African American parents are proficient at seeking out additional information when they do not understand what they find on the Internet, or perhaps there is incongruence in what they believe they understand and what they actually do. Finally, our results suggest that parents whose children with special health care needs had excellent to good health had significantly higher agreement in feeling confident about using the information they found online versus parents whose children with special health care needs were in poor health (the referent group). Future research should focus on how interventions can be targeted to improve the confidence of this particularly vulnerable group of parents whose children with special health care needs are in poor health. Given that caregiver burden is directly associated with severity of a child's illness, this burden may be preventing these parents from having the time to develop eHealth literacy [29]. Ironically, parents of more severely ill children may require more information due to the complexities of their children's illnesses, and lower eHealth literacy skills may inhibit their ability to make decisions.

Several study limitations merit attention. First, although we had more than 2300 completed surveys, the response rate for the survey was 58%. While this response rate is consistent with prior surveys conducted with this population, it is possible that inherent differences between responders and nonresponders exist [30,31]. Second, our sample consists of parents whose children are enrolled in publicly funded health insurance programs. By definition, all children enrolled in the program are members of families with low incomes. Lower socioeconomic status is associated with lower health literacy, but we could not identify any studies that have investigated the effect of socioeconomic status on eHealth literacy. Both health literacy and eHealth literacy are important for utilizing Web-based applications. Our findings, which may be less generalizable to broader socioeconomic groups, show that even in low socioeconomic groups, Internet use is high (1945/2371, 82%) and that several questions on the eHEALS had high levels of agreement. Research should be conducted to determine the relationship between health literacy and eHealth literacy. Third, our study focused on parents of children with special health care needs. However, we acknowledge that parents of children without special health care needs may have different eHealth literacy and Internet use patterns. Fourth, eHealth literacy in this study was self-reported. Future research should develop a method to compare self-assessment with expert assessment in order to better interpret the results. Fifth, our study does not consider information presented to parents in multiple formats and multiple languages. Finally, we did not assess Internet use and eHealth literacy of the children and adolescents, who should also be engaged in the process of seeking health information.

Despite these limitations, our findings contribute to the dearth of evidence in the pediatric literature on parents' Internet use and eHealth literacy. The ability to measure the eHealth literacy of parents whose children have special health care needs in an easy, valid manner as well as comment on the factors that are associated with greater eHealth literacy highlight some of the opportunities and challenges that the pediatric community faces if it wants to design Web-based applications to improve the health outcomes of children and their families. Organizations that provide information to parents on the Internet should show parents different websites to help them distinguish between high and low quality information, provide parents with information on what to look for on websites (such as citations from scientific studies), and should provide parents with information about online education and training opportunities in their area as they are available. Given the national push for health information technology adoption, understanding these issues in pediatric care is important. Future research should focus on systematically developing and testing interventions that could raise eHealth literacy and ultimately increase family empowerment.

Acknowledgments

This work was funded by contract COQQA between the University of Florida and the Florida Children's Medical Services Network Division (principal investigators Shenkman and Knapp). Publication of this paper was funded in part by the University of Florida Open-Access Publishing Fund.

Conflicts of Interest

None declared

References

1. Raine L, Fox S, Horrigan JB, Lenhart A, Spooner T. Tracking online life: How women use the Internet to cultivate relationships with family and friends. Washington, DC: The Pew Center Internet & American Life Project; 2000 May 10. URL: http://www.pewinternet.org/~/media//Files/Reports/2000/Report1.pdf.pdf [accessed 2010-11-24] [WebCite Cache ID 5uU3ikMQA]
2. Irving L, Levy K. Falling Through the Net: Defining the Digital Divide. Washington, DC: National Telecommunications and Information Administration, US Department of Commerce; 1999 Jul 08. URL: http://www.ntia.doc.gov/ntiahome/fttn99/contents.html [accessed 2010-11-24] [WebCite Cache ID 5uTzUk84i]
3. Wangberg S, Andreassen H, Kummervold P, Wynn R, Sørensen T. Use of the internet for health purposes: trends in Norway 2000-2010. Scand J Caring Sci 2009 Dec;23(4):691-696. [doi: 10.1111/j.1471-6712.2008.00662.x] [Medline: 19473317]
4. Renahy E, Parizot I, Chauvin P. [Internet and health: lessons learned from the French national survey on households living conditions in 2005]. Rev Epidemiol Sante Publique 2008 Apr;56(2):117-125. [doi: 10.1016/j.respe.2008.01.009] [Medline: 18472374]
5. Smith A. Home broadband 2010. Washington, DC: Pew Internet & American Life Project; 2010 Aug 11. URL: http://www.pewinternet.org/~/media//Files/Reports/2010/Home%20broadband%202010.pdf [accessed 2010-11-24] [WebCite Cache ID 5uU3uk8T0]
6. Lorence D, Park H. Group disparities and health information: a study of online access for the underserved. Health Informatics J 2008 Mar;14(1):29-38. [doi: 10.1177/1460458207086332] [Medline: 18258673]
7. Lorence D, Park H. Study of education disparities and health information seeking behavior. Cyberpsychol Behav 2007 Feb;10(1):149-151. [doi: 10.1089/cpb.2006.9977] [Medline: 17305464]
8. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2009 Jun. URL: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2010-11-24] [WebCite Cache ID 5uU3xLI3e]
9. McInnes DK, Gifford AL, Kazis LE, Wagner TH. Disparities in health-related internet use by US veterans: results from a national survey. Inform Prim Care 2010;18(1):59-68. [Medline: 20429979]
10. Lea J, Lockwood G, Ringash J. Survey of computer use for health topics by patients with head and neck cancer. Head Neck 2005 Jan;27(1):8-14. [doi: 10.1002/hed.20114] [Medline: 15551302]
11. Walsh MC, Trentham-Dietz A, Schroepfer TA, Reding DJ, Campbell B, Foote ML, et al. Cancer information sources used by patients to inform and influence treatment decisions. J Health Commun 2010 Jun;15(4):445-463. [doi: 10.1080/10810731003753109] [Medline: 20574881]
12. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

13. Mackert M, Kahlor L, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: addressing the obesity epidemic. Telemed J E Health 2009 Sep;15(7):672-677. [doi: 10.1089/tmj.2009.0012] [Medline: 19694596]
14. Kind T, Huang ZJ, Farr D, Pomerantz KL. Internet and computer access and use for health information in an underserved community. Ambul Pediatr 2005;5(2):117-121. [doi: 10.1367/A04-107R.1] [Medline: 15780014]
15. Roche MI, Skinner D. How parents search, interpret, and evaluate genetic information obtained from the internet. J Genet Couns 2009 Apr;18(2):119-129. [doi: 10.1007/s10897-008-9198-4] [Medline: 18937062]
16. Knijnenburg SL, Kremer LC, van den Bos C, Braam KI, Jaspers MW. Health information needs of childhood cancer survivors and their family. Pediatr Blood Cancer 2010 Jan;54(1):123-127. [doi: 10.1002/pbc.22207] [Medline: 19743299]
17. Porter A, Edirippulige S. Parents of deaf children seeking hearing loss-related information on the internet: the Australian experience. J Deaf Stud Deaf Educ 2007;12(4):518-529 [FREE Full text] [doi: 10.1093/deafed/enm009] [Medline: 17456507]
18. Knapp C, Madden V, Marcu M, Wang H, Curtis C, Sloyer P, et al. Information seeking behaviors of parents whose children have life-threatening illnesses. Pediatr Blood Cancer 2011 May;56(5):805-811. [doi: 10.1002/pbc.22674] [Medline: 21370415]
19. Brown CA, Dickson R. Healthcare students' e-literacy skills. J Allied Health 2010;39(3):179-184. [Medline: 21174023]
20. Robinson C, Graham J. Perceived Internet health literacy of HIV-positive people through the provision of a computer and Internet health education intervention. Health Info Libr J 2010 Dec;27(4):295-303. [doi: 10.1111/j.1471-1842.2010.00898.x] [Medline: 21050372]
21. McPherson M, Arango P, Fox H, Lauver C, McManus M, Newacheck PW, et al. A new definition of children with special health care needs. Pediatrics 1998 Jul;102(1 Pt 1):137-140. [Medline: 9714637]
22. Shiffman RN, Spooner SA, Kwiatkowski K, Brennan PF. Information technology for children's health and health care: report on the Information Technology in Children's Health Care Expert Meeting, September 21-22, 2000. J Am Med Inform Assoc 2001;8(6):546-551 [FREE Full text] [Medline: 11687562]
23. Statistical Analysis Software for Professionals, Release 9 [computer software]. Cary, NC: SAS Institute Inc URL: http://www.stata.com/error/404.html [accessed 2010-11-24] [WebCite Cache ID 5uTzDBNbN]
24. Smith A. Mobile access 2010. Washington, DC: Pew Internet & American Life Project; 2010 Jul 07. URL: http://www.pewinternet.org/~/media//Files/Reports/2010/PIP_Mobile_Access_2010.pdf [accessed 2010-11-24] [WebCite Cache ID 5uU40MEaO]
25. Salovey P, Williams-Piehota P, Mowad L, Moret ME, Edlund D, Andersen J. Bridging the digital divide by increasing computer and cancer literacy: community technology centers for head-start parents and families. J Health Commun 2009;14(3):228-245. [doi: 10.1080/10810730902805804] [Medline: 19440907]
26. Health Care Financing and Organization. Health Information Technology. Washington, DC: Health Care Financing & Organization; 2009 Apr. URL: http://www.hcfo.org/publications/health-information-technology [accessed 2010-11-24] [WebCite Cache ID 5uTzFRJ8b]
27. Centers for Medicare and Medicaid Services. Medicare and Medicaid Health Information Technology: Title IV of the American Recovery and Reinvestment Act. Washington, DC: US Department of Health & Human Services; 2009. URL: http://www.cms.hhs.gov/apps/media/press/factsheet.asp?Counter=3466&intNumPerPage=10&checkDate=&checkKey=&srchType=1&numDays=3500&srchOpt=0&srchData=&keywordType=All&chkNewsType=6&intPage=&showAll=&pYear=& [WebCite Cache ID 5uTzGdfEp]
28. Stepanikova I, Mollborn S, Cook KS, Thom DH, Kramer RM. Patients' race, ethnicity, language, and trust in a physician. J Health Soc Behav 2006 Dec;47(4):390-405. [Medline: 17240927]
29. Raina P, O'Donnell M, Rosenbaum P, Brehaut J, Walter SD, Russell D, et al. The health and well-being of caregivers of children with cerebral palsy. Pediatrics 2005 Jun;115(6):e626-e636 [FREE Full text] [doi: 10.1542/peds.2004-1689] [Medline: 15930188]
30. Knapp C, Huang IC, Madden V, Vadaparampil S, Quinn G, Shenkman E. An evaluation of two decision-making scales for children with life-limiting illnesses. Palliat Med 2009 Sep;23(6):518-525. [doi: 10.1177/0269216309104892] [Medline: 19346274]
31. Huang IC, Shenkman EA, Leite W, Knapp CA, Thompson LA, Revicki DA. Agreement was not found in adolescents' quality of life rated by parents and adolescents. J Clin Epidemiol 2009 Mar;62(3):337-346. [doi: 10.1016/j.jclinepi.2008.06.012] [Medline: 18834712]

Edited by G Eysenbach; submitted 29.11.10; peer-reviewed by D Charbonneau, L LaRue; comments to author 20.04.11; accepted 14.06.11; published 29.09.11.

Please cite as:
Knapp C, Madden V, Wang H, Sloyer P, Shenkman E
Internet Use and eHealth Literacy of Low-Income Parents Whose Children Have Special Health Care Needs
J Med Internet Res 2011;13(3):e75
URL: http://www.jmir.org/2011/3/e75/
doi:10.2196/jmir.1697
PMID:21960017

Original Paper

A Comparison of Physician Pre-Adoption and Adoption Views on Electronic Health Records in Canadian Medical Practices

Norm Archer[1], PhD; Mihail Cocosila[2], PhD

[1]DeGroote School of Business, McMaster University, Hamilton, ON, Canada
[2]Athabasca University, Athabasca, AB, Canada

Corresponding Author:
Norm Archer, PhD
DeGroote School of Business
McMaster University
1280 Main St. West
Hamilton, ON, L8S 4M4
Canada
Phone: 1 905 525 9140 ext 23944
Fax: 1 905 528 0556
Email: archer@mcmaster.ca

Abstract

Background: There is a major campaign involving large expenditures of public money to increase the adoption rate of electronic health record (EHR) systems in Canada. To maximize the chances of success in this effort, physician views on EHRs must be addressed, since user perceptions are key to successful implementation of technology innovations.

Objective: We propose a theoretical model comprising behavioral factors either favoring or against EHR adoption and use in Canadian medical practices, from the physicians' point of view. EHR perceptions of physicians already using EHR systems are compared with those not using one, through the lens of this model.

Methods: We conducted an online cross-sectional survey in both English and French among medical practitioners across Canada. Data were collected both from physicians using EHRs and those not using EHRs, and analyzed with structural equation modeling (SEM) techniques.

Results: We collected 119 responses from EHR users and 100 from nonusers, resulting in 2 valid samples of 102 and 83 participants, respectively. The theoretical adoption model explained 55.8% of the variance in behavioral intention to continue using EHRs for physicians already using them, and 66.8% of the variance in nonuser intention to adopt such systems. Perception of ease of use was found to be the strongest motivator for EHR users (total effect .525), while perceptions of usefulness and of ease of use were the key determinants for nonusers (total effect .538 and .519, respectively) to adopt the system. Users see perceived overall risk associated with EHR adoption as a major obstacle (total effect –.371), while nonusers perceive risk only as a weak indirect demotivator. Of the 13 paths of the SEM model, 5 showed significant differences between the 2 samples (at the .05 level): general doubts about using the system ($P = .02$), the necessity for the system to be relevant for their job ($P < .001$), and the necessity for the system to be useful ($P = .049$) are more important for EHR nonusers than for users, while perceptions of overall obstacles to adoption ($P = .03$) and system ease of use ($P = .042$) count more for EHR users than for nonusers.

Conclusions: Relatively few differences in perceptions about EHR system adoption and use exist between physicians already using such systems and those not yet using the systems. To maximize the chances of success for new EHR implementations from a behavioral point of view, general doubts about the rationale for such systems must be mitigated through improving design, stressing how EHRs are relevant to physician jobs, and providing substantiating evidence that EHRs are easier to use and more effective than nonusers might expect.

(J Med Internet Res 2011;13(3):e57) doi:10.2196/jmir.1726

KEYWORDS

Electronic health record; information technology; medical practice; Canada

Introduction

Context

Less than 25% of Canadian doctors use electronic health record (EHR) systems, less than in many other countries and ranking last among industrialized nations [1]. As a result, although almost all physicians use some form of computer support for scheduling and billing purposes, patient clinical records are mostly on paper, and are scattered and often inaccessible in doctors' offices, clinics, test centers, labs, and hospitals.

An EHR (a term often used interchangeably with EMR, or electronic medical record) is a repository of information in computer-processable form that is employed by a physician to record and access information regarding the health of a patient. EHRs are becoming an essential artifact of current health care: to properly care for their patients, health care practitioners must have timely and accurate access to all relevant medical records. Absence of EHR systems can cause significant delays, duplication of effort, and even inaccuracy in diagnosing problems [2-4].

Despite the obvious advantages of EHR systems, their adoption rate has been slow in Canadian medical practices. Insufficient adoption of EHR systems is a highly complex problem that has not been addressed adequately in a comprehensive manner. Many interdependent factors influence adoption, and these must be considered simultaneously [5]. Above all, user perceptions are key in determining the success of an information technology (IT) deployment in any context [6].

Several studies regarding adoption of EHR systems in Canada and elsewhere have been done, but most of this work addressed in isolation a few relevant motivations for, or obstacles to, adoption [7-10]. It is important for such research to address the major issues comprehensively, or health care policies may be adopted that encourage the use of EHRs for the wrong reasons, leaving physicians with problems rather than benefits. The end result may be abandonment or underutilization of such systems. It is therefore important to have a broader understanding of physician views regarding salient factors for EHR adoption, and to compare how and why these views differ between adopters and nonadopters.

Theoretical Background and Model

Investigating user adoption of new IT applications in various fields, including health care, is an established topic of information systems research. Various theories have been validated that attempt to better explicate user reasons to adopt new IT applications [11]. However, applying these models to medical practice situations is a challenging task because these practices involve a blend of autonomous physician activities with team interaction, thus requiring theories of both individual and organizational IT adoption. Therefore, one approach to examining medical practitioner perceptions of EHRs is to consider 2 sets of factors: (1) elements of individual use that stem from theories of individual decision making (eg, Technology Acceptance Model (TAM) [12] or Unified Theory of Acceptance and Use of Technology [11]), and (2) elements of organizational use (eg, the organizational version of TAM, TAM2 [13], or the Theory of Reasoned Action [14]).

Both of these categories of theories investigate primary factors such as behavioral intent to adopt and use an IT application, including performance expectancy (perceived usefulness) and effort expectancy (ease of use).

Of the many other (secondary) factors that may influence positively the intention to adopt EHRs, we consider 3 as more important: personal IT innovativeness (willingness of an individual to try new IT applications [15]), job relevance (functionalities provided for physician needs [16]), and social influence (influence of colleagues and significant others on adopting a system [17,18]).

A recent trend in information systems research is to also consider factors of resistance to a new IT implementation [19]. This category is of particular interest for investigating IT deployment in health care [20], as this field has high social sensitivity while lagging behind other industry sectors (eg, banking or tourism) in IT deployment. Factors having a negative effect on intention to adopt IT can be explained through the concept of perceived risk borrowed from consumer behavior, where consumers express doubts about purchasing due to anxieties such as wasting money and time, and privacy considerations [21]. Medical practitioners who consider adopting EHRs would tend to perceive these risks as being associated with themselves, even though the decision is in an organizational context, similarly to consumers considering the adoption of innovations [22].

Obstacles to EHR adoption are captured through the use of perceived overall risk, which expresses perceived negative consequences of EHR use. Recent information systems research has shown that perceived risk has several facets [21,23]. These depend on the context of the activity users would perceive as risky, but their aggregated influence should generate the same overall result. In the context of considering the adoption of EHRs, 3 risk facets are likely to be important antecedents of overall risk: (1) perceived performance risk (fear that the system will fail to perform as expected), (2) perceived legal and privacy risk (fear of legal and privacy problems from EHR use), and (3) perceived psychological risk (anxiety and stress about EHR implementation).

Building on previous consumer behavior and information systems research [19,24], we can find perceived overall risk to affect negatively both the perception of usefulness (hence performance expectancy) and the intention to use EHR systems. On the other hand, risk perception is alleviated when users perceive the system as easy to use [21].

Based on the above theoretical reasoning and the results of previous research, the following theoretical model (Figure 1) and hypotheses are proposed:

Hypothesis 1a) perceived performance risk will relate positively to perceived overall risk.

Hypothesis 1b) perceived psychological risk will relate positively to perceived overall risk.

Hypothesis 1c) perceived legal and privacy risk will relate positively to perceived overall risk.

Hypothesis 2) perceived overall risk will relate negatively to performance expectancy.

Hypothesis 3) perceived overall risk will relate negatively to behavioral intention to adopt.

Hypothesis 4) job relevance will relate positively to performance expectancy.

Hypothesis 5) social influence will relate positively to performance expectancy.

Hypothesis 6) personal IT innovativeness will relate positively to performance expectancy.

Hypothesis 7) personal IT innovativeness will relate positively to effort expectancy.

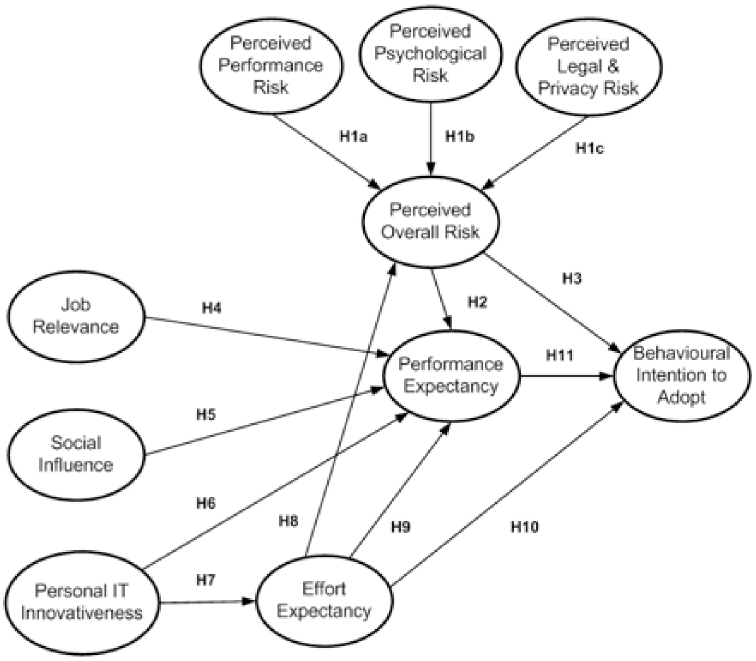
Hypothesis 8) effort expectancy will relate negatively to perceived overall risk.

Hypothesis 9) effort expectancy will relate positively to performance expectancy.

Hypothesis 10) effort expectancy will relate positively to behavioral intention to adopt.

Hypothesis 11) performance expectancy will relate positively to behavioral intention to adopt.

Figure 1. Theoretical model of electronic health record adoption in Canadian medical practices showing hypothesized (H) relationships



Purpose

This theoretical model captures the key behavioral factors that may favor or disfavor EHR adoption and use in Canadian medical practices, from the physicians' point of view. Perceptions of physicians who have adopted EHRs can be systematically compared with those who have not, through the model's lens.

Methods

Participant Recruitment and Data Collection

Data were collected through a cross-sectional online survey comprising an instrument based on the theoretical model hypotheses, together with relevant demographic questions. The items of all the constructs in the theoretical model were measured with 7-point Likert scales having as anchors *strongly disagree* and *strongly agree*. Relevant constructs were adapted from previous research in information systems, health care, and consumer behavior [11,14,20,21,23].

We designed 2 versions of the survey: one for physicians using EHR systems, and the other one for physicians not using such systems. The only difference between the 2 versions was that items capturing various perceptions for physicians using EHRs were formulated in the present tense (eg, "I find EHR useful"), while the questions addressed to the other category of physicians were phrased in the conditional present (eg, "I would find EHR useful").

The survey was pretested by 4 PhD students in a Canadian university and, after improvement, pilot-tested with 3 practicing physicians. The research was approved by the Research Ethics Board of McMaster University.

Participants targeted for data collection were physicians working in medical practices in Canada. Because these health professionals have a busy schedule and are difficult to recruit for research, participant recruitment and survey administration were outsourced to a commercial firm having a panel of almost 67,000 preregistered physicians. The survey was conducted across Canada in both English and French, and participants were compensated for completing the survey. The total sample was of 220 participants with a balanced distribution between physicians working in clinics already using EHRs and physicians from practices not using EHRs. Additionally, an attempt was made to balance participants between those working in small (1 or 2 physicians) and larger clinics, and between general practitioners and specialists.

Theoretical Model Analyses

Data were analyzed through descriptive statistics and structural equation modeling (SEM) techniques. A popular SEM approach, partial least squares (PLS), was used due to its suitability for predominantly exploratory research using complex models [25]. In addition, PLS requires no assumptions about sample data distributions [26] and works well with relatively small samples [27]. PLS analysis was done with SmartPLS software (release 2.0 [beta]; SmartPLS, Hamburg, Germany) in 2 stages: evaluation of the measurement model, followed by that of the structural model [28]. Demographic factors collected in the study were also tested as possible control variables in each of the 2 samples.

Analysis of Differences between EHR Users and Nonusers

Outcomes of model analyses were compared through differences in path coefficients between the 2 models. The degree of difference was assessed with the t statistic with $N1+N2-2$ degrees of freedom [29,30]:

$t = (Path1 – Path 2) / [Spooled * sqrt (1 / N1 + 1 / N2)]$

where *Path1*, *Path2* are the corresponding path coefficients in the 2 models and *N1*, *N2* are the respective sample sizes.

Spooled is the pooled estimator for the variance, calculated as:

$Spooled = sqrt \{[square\ of(N1 – 1)/(N1 + N2 – 2)] * square\ of\ SE1 +[square\ of\ (N2 – 1) / (N1 + N2 – 2)] *square\ of\ SE2\}$

where *SE1*, *SE2* are the standard errors of the respective model path coefficients.

The above approach is suitable when the variance of the 2 samples that are being compared is approximately the same, or when sample sizes are relatively large.

Results

Characteristics of Participants

We collected 219 completed questionnaires: 119 EHR users and 100 nonusers. Eliminating questionnaires with more than 10% missing data left 102 and 83 valid responses, respectively. Missing data were replaced by the multiple imputation approach through a predictive mean matching procedure with 10 imputations [31]. The resulting corrected samples were the basis for subsequent statistical analyses.

Table 1 shows the demographic characteristics of the 2 samples used in the research.

Table 1. Participant characteristics

	EHR users	EHR nonusers
Sample size	102	83
Practice size		
Small	47 (46%)	47 (57%)
Large	55 (54%)	36 (43%)
Respondent practice type		
General practice	50 (49%)	44 (53%)
Specialists	52 (51%)	39 (47%)
Average medical experience (years)	18.5	20.7
Average number of physicians per practice	6.9	3.8
Work schedule		
Full-time	99 (97%)	81 (98%)
Part-time	3 (3%)	2 (2%)
Gender		
Female	27 (26%)	15 (18%)
Male	75 (73%)	68 (82%)

Theoretical Model Analyses

A first evaluation of the measurement model indicated the necessity to drop 7 of 35 items from the EHR user model and 5 of 35 items from the EHR nonuser model because of unsatisfactory reliability and construct validity values. After rerunning SmartPLS for the remaining items, all constructs for both models displayed Cronbach alpha (composite reliability) values above .7, average variance extracted above .5, and item loadings above .7. Thus, the measurement model had acceptable reliability and convergent validity [32-34]. A visual inspection of the matrix of loadings and cross-loadings produced by SmartPLS indicated that the loadings of measurement items on their assigned constructs was larger than cross-loadings on other constructs, leading to the conclusion that constructs of both models had sufficient discriminant validity [28]. Overall, the statistical analysis of data for both models indicated adequate reliability and construct validity, leading to structural analysis of the model, the second step of PLS analysis.

Evaluation of the structural model involved running SmartPLS with a bootstrap of 200 resamples. Results of path coefficients, their significance levels, and hypotheses outcomes, together with the values of the coefficient of determination, are shown in Figure 2 and Table 2 for the EHR users model and in Figure 3 and Table 3 for the EHR nonusers model.

Figure 2. Structural evaluation of the electronic health record user model

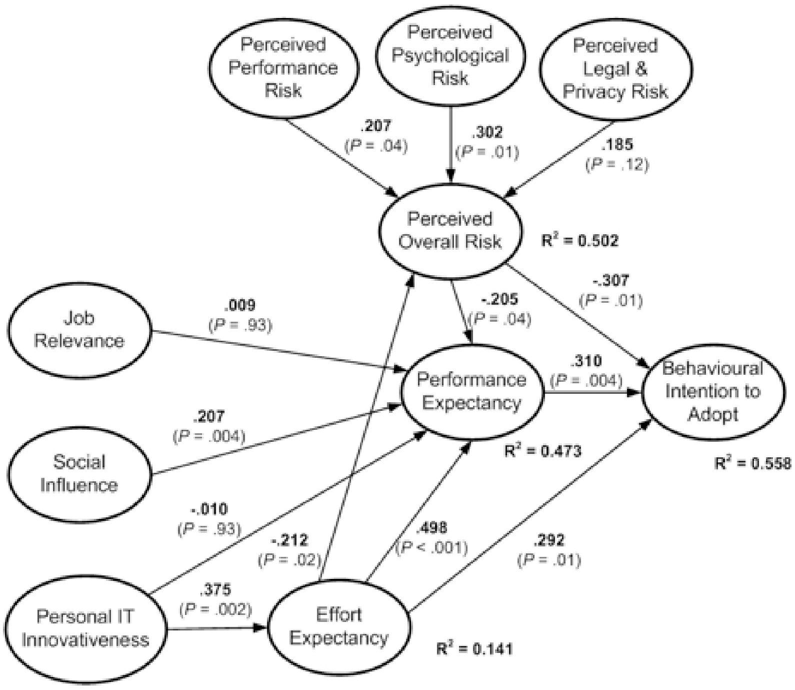


Table 2. Hypothesis test results for the EHR user model

Hypothesis	Path	Coefficient	*P* value	Outcome
H1a	Performance risk → overall risk	.207	.04	Supported
H1b	Psychological risk → overall risk	.302	.01	Supported
H1c	Legal and privacy risk → overall risk	.185	.12	Rejected
H2	Overall risk → performance expectancy	–.205	.04	Supported
H3	Overall risk → behavioral intention	–.307	.01	Supported
H4	Job relevance → performance expectancy	.009	.93	Rejected
H5	Social influence → performance expectancy	.207	.004	Supported
H6	Personal IT[a] innovativeness → performance expectancy	–.010	.93	Rejected
H7	Personal IT[a] innovativeness → effort expectancy	.375	.002	Supported
H8	Effort expectancy → overall risk	–.212	.02	Supported
H9	Effort expectancy → performance expectancy	.498	<.001	Supported
H10	Effort expectancy → behavioral intention	.292	.01	Supported
H11	Performance expectancy → behavioral intention	.310	.004	Supported

[a] Information technology.

Figure 3. Structural evaluation of the electronic health record nonuser model

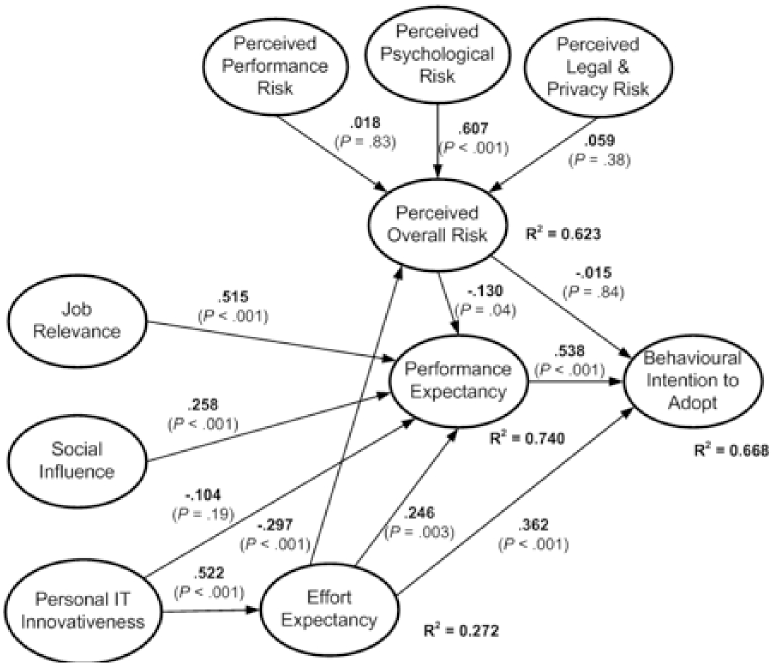


Table 3. Hypothesis test results for the EHR nonuser model

Hypothesis	Path	Coefficient	*P* value	Outcome
H1a	Performance risk → overall risk	.018	.83	Rejected
H1b	Psychological risk → overall risk	.607	<.001	Supported
H1c	Legal and privacy risk → overall risk	.059	.38	Rejected
H2	Overall risk → performance expectancy	–.130	.04	Supported
H3	Overall risk → behavioral intention	–.015	.84	Rejected
H4	Job relevance → performance expectancy	.515	<.001	Supported
H5	Social influence → performance expectancy	.258	<.001	Supported
H6	Personal IT[a] innovativeness → performance expectancy	–.104	.19	Rejected
H7	Personal IT[a] innovativeness → effort expectancy	.522	<.001	Supported
H8	Effort expectancy → overall risk	–.297	<.001	Supported
H9	Effort expectancy → performance expectancy	.246	.003	Supported
H10	Effort expectancy→ behavioral intention	.362	<.001	Supported
H11	Performance expectancy → behavioral intention	.538	<.001	Supported

[a] Information technology.

SmartPLS results also provided the total effects of the factors in the theoretical model on behavioral intention to use EHRs, for the 2 categories of participants (see Table 4).

Number of physicians per medical practice, years of medical experience, and gender of the respondents were tested as potential control variables. Path coefficients and significance levels for the paths from these constructs to the endogenous latent variables of both theoretical models were calculated by running PLS analysis. A fourth potential control variable (work schedule) did not produce any results in the PLS analysis due to its heavy bias: over 97% of the physicians in both samples reported working full time. Of all the control variables tested, only the number of physicians per practice in the EHR nonusers model was found to be significant, with a path coefficient of .106 ($P = .04$) to effort expectancy. The variance explained by this variable increased from $R^2 = 0.272$ to $R^2 = 0.283$.

Table 4. Total effects and their significance levels on behavioral intention

Antecedent construct	EHR users	EHR nonusers
Perceived performance risk	–.077 (P = .048)	–.002 (P = .88)
Perceived psychological risk	–.112 (P = .07)	–.052 (P = .28)
Perceived legal and privacy risk	–.068 (P = .17)	–.005 (P = .61)
Perceived overall risk	–.371 (P < .001)	–.085 (P = .27)
Job relevance	.003 (P = .94)	.277 (P < .001)
Social influence	.064 (P = .06)	.139 (P < .001)
Personal IT[a] innovativeness	.194 (P = .03)	.215 (P = .005)
Effort expectancy	.525 (P < .001)	.519 (P < .001)
Performance expectancy	.310 (P = .004)	.538 (P < .001)

[a] Information technology.

Analysis of Differences Between EHR Users and Nonusers

The standard deviations of the corresponding path coefficients of the 2 models are of similar orders of magnitude in most cases. In addition, sample sizes were relatively large, at more than twice the minimum sample required for PLS [34]. Accordingly, the t statistic formula described above was used to calculate the degree of difference between the path coefficients of the 2 models. The results are presented in Table 5.

Table 5. Statistical analysis of differences between EHR users and nonusers

Path	EHR users path coefficient	EHR nonusers path coefficient	t Value of difference	P value
Performance risk → overall risk	.207	.018	1.444	.07
Psychological risk → overall risk	.302	.607	2.034	.02
Legal and privacy risk → overall risk	.185	.059	0.874	.19
Overall risk → performance expectancy	–.205	–.130	0.618	.27
Overall risk → behavioral intention	–.307	–.015	1.964	.03
Job relevance →> performance expectancy	.009	.515	3.600	<.001
Social influence → performance expectancy	.207	.258	0.544	.29
Personal IT[a] innovativeness → performance expectancy	–.010	–.104	0.685	.25
Personal IT[a] innovativeness → effort expectancy	.375	.522	0.982	.16
Effort expectancy → overall risk	–.212	–.297	0.713	.24
Effort expectancy → performance expectancy	.498	.246	1.733	.042
Effort expectancy → behavioral intention	.292	.362	0.495	.31
Performance expectancy → behavioral intention	.310	.538	1.659	.049

[a] Information technology.

Discussion

Principal Findings and Comparison with Prior Work

The objectives of this study were to identify the main factors influencing the adoption of EHRs in Canada from the medical practitioners' perspective, and to compare the views of physicians already using such systems with those not using them. The most important findings in general are that high performance expectancy and little effort expectancy about EHRs are significant positive adoption factors, while perceived overall risk is a deterrent. Few differences in the perceptions of EHR systems were noticed between users and nonusers: these concern mostly the adoption roles of effort expectancy and perceived overall risk. The following focuses first on the results from existing users, then on results from nonusers, and then on a comparison of differences between the 2 population samples.

EHR User Model

As expected, according to a large body of information systems literature, usefulness perception captured through performance expectancy is a key direct reason influencing physicians already using EHRs to continue using them [6] (Figure 2). Of the hypothesized antecedents of performance expectancy, only social influence proved to have a significant influence. Thus, in a sensitive field like health care, positive references from colleagues and collaborators about EHRs is a factor influencing practitioners to see the usefulness of such systems. This finding is consistent with previous organizational information systems studies that include subjective norm in technology adoption models [13].

Effort expectancy, indicating a perception of ease of use, is an important determinant of both behavioral intention to adopt and performance expectancy, in agreement with information system studies [6]. Moreover, effort expectancy has by far the strongest total effect on behavioral intention (coefficient = .525, $P < .001$), due to its direct path as well as its indirect influence through performance expectancy (Figure 2 and Table 4): ease of use amplifies the perception of usefulness and, hence, intention to continue using the system. Ease of use is augmented by physician personal IT innovativeness, which is concordant with the literature [15,35]: physicians interested in IT innovations have little difficulty in using EHRs, thus reinforcing the perception of usefulness and intention to use. Personal IT innovativeness has a significant overall effect ($P = .03$) on behavioral intention to use EHRs (Table 4).

With a negative effect of –.371 ($P < .001$), perceived overall risk is the main obstacle to intention to continue using EHRs, for physicians already using them (Figure 2 and Table 4). This is consistent with consumer behavior research where perceived risk associated with a product negatively influences intention to purchase it [24]. Of the 3 types of risk considered meaningful in this study, only perceived performance risk ($P = 0.4$) and, especially, perceived psychological risk ($P = .01$) had a significant influence on the overall risk perceived by EHR users. This shows the key deterrent roles of fears about the system not working as expected and general doubts about the role played by EHRs in medical practices. These findings are concordant with previous consumer behavior and information systems studies [19,21]. As shown in Figure 2, a low expectancy of effort eases the perception of risk (coefficient = –.212, $P = .02$). This demonstrates the twofold role played by effort expectancy in the model (a positive direct and indirect influence on one side and a risk-reducing effect on the other side). Consequently, one way to promote more effective use of EHR systems among physicians already using them is to focus on improving ease of use.

Overall, the theoretical model explaining the adoption of EHR by physicians using these systems had a moderately high R^2 (variance explained) values for all the endogenous constructs (between 0.141 for effort expectancy and 0.558 for behavioral intention) and a high proportion of significant paths: 10 out of 13 (Table 2). According to the literature on PLS methodology, the proposed model could be termed appropriate [36].

EHR Nonuser Adoption Model

Similarly to the model for physicians using EHRs, performance expectancy and effort expectancy are strong and significant explanations of behavioral intention to adopt for nonusers (Figure 3). Results in Table 4 show their total effect is about the same level (.538, $P < .001$ and .519, $P < .001$, respectively), demonstrating that it is important for EHRs to be perceived as both useful and easy to use by physicians who have not yet been exposed to this technology. While influences from the social environment are a significant element strengthening perceptions of usefulness, physician IT innovativeness remains a strong antecedent of perceptions of ease of use (Table 3). In addition, nonusers feel that it is important for EHRs to be relevant in their jobs when they decide on adoption (Figure 3 and Table 4).

Perceived overall risk remains a negative factor in the adoption equation for EHR nonusers, but with a reduced influence compared with results from the user model. There is a direct negative effect on the usefulness perception ($P = .04$) but no significant direct or total effect on the intention to adopt such systems (Figure 3 and Table 4). Perceived psychological risk is the only antecedent of overall risk that is significant in the model, but its influence is strong (coefficient = .607, $P < .001$) (Table 3). A possible explanation is that nonusers do not have a deep understanding of the potential obstacles associated with the use of such systems. Nonetheless, anxiety about introducing an EHR system is an important obstacle that needs to be mitigated if adoption is to be encouraged.

Of the 4 endogenous constructs of the nonuser model, 3 displayed moderately high R^2 values (between 0.272 and 0.740) and 9 of the 13 paths hypothesized were significant. Hence, the model explains relatively well the nonuser perceptions of EHR adoption. Therefore, the theoretical model appears to be appropriate for nonusers [36], for reasons similar to the suitability demonstrated for the user model.

Control Variable Tests

The control variables tested (number of physicians per medical practice, years of medical experience, work schedule, and gender of the respondents) did not produce any effect, with one exception: the positive influence of the number of physicians on effort expectancy for EHR nonusers. The moderately strong and significant effect may indicate that the larger the clinic, the lower the effort perceived in adopting and using EHRs. This effect is likely due to the availability of more technologically knowledgeable support staff in larger clinics.

Differences in Perceptions Between EHR Users and Nonusers

The models depicted in Figures 2 show differences in factor influences between EHR users and nonusers. However, results in Table 5 indicate significant differences in only 5 paths out of 13. Thus, perceived psychological risk is a strong and significant antecedent of perceived overall risk in both models, but its influence is much higher for nonusers ($P = .02$). Therefore, decreasing perceived adoption obstacles and easing the acceptance of EHRs among physicians not yet using them would help to address general doubts about the rationale of using EHRs. However, perceived overall risk is a significantly stronger obstacle for users than for nonusers ($P = .03$), indicating that user concerns must be addressed during implementation to avoid negative views in later stages.

As shown in Table 5, job relevance is a much stronger antecedent of performance expectancy for EHR nonusers than for users ($P < .001$). This indicates that, in the process of developing and implementing EHRs for practices not yet using them, it is important that these applications address the most important tasks as seen by physicians.

Effort expectancy has a more significant influence on performance expectancy for EHR users than for nonusers (*P* = .042, in Table 5). This indicates that the ease of use perceived by physicians currently using EHRs is higher than for nonusers. Therefore, the design and implementation process of new EHRs for medical practices should include efforts to mitigate concerns about the difficulty of using the systems.

Performance expectancy is an important antecedent of intention to use in both models. Its influence appears to be stronger for nonusers but only at the limit of significance (*P* = .049).

Limitations

This study has certain limitations. Although study participants were recruited from a large sampling frame, they self-selected. Further, because of feasibility constraints, the resulting sample size was relatively small for such a complex investigation. Although the valid samples of 102 EHR users and 83 EHR nonusers were more than twice as large as the minimum PLS requirements for a reliable statistical analysis [34,37], samples were not particularly homogeneous from a medical practice point of view. This was demonstrated in part by comments received from participants that, although difficult to classify, ranged all the way from frustration to satisfaction for users, and from concerns about productivity to positive anticipation from nonusers.

Although the study showed little effect from participant characteristics (years of medical experience, gender, full time vs part time), there was some differentiation based on clinic size (number of physicians employed), and future research should use larger samples that can discriminate more between such characteristics. Larger samples might also differentiate between physicians involved versus those not involved in decisions on EHR system adoption. For example, various risk perceptions might depend on the degree of respondent involvement in EHR adoption decisions.

This study did not attempt to differentiate between various levels of EHR experience, frequency of use, and complexity of tasks performed with the system by EHR users. Thus, physicians in some clinics might use only basic functionalities of EHRs (eg, tracking patient prescription records) that would perform well but have less impact on work quality and productivity, while physicians in other practices may use more complex functions, better suited to supporting their tasks, but encountering different types of user problems.

For nonusers, a useful future differentiation would be between the medical practices that have not seriously considered adopting EHRs yet and those practices that are seriously considering the use of EHRs within a certain time horizon. Perceptions of certain risks may differ between these 2 categories, depending on physician familiarity with related issues.

Conclusions and Practical Implications

Comparing EHR perceptions of physicians already using the systems with those not yet using them through a rigorous theoretical approach has helped to understand certain key behavioral aspects that must be addressed in the deployment of these systems. To maximize the chances of success of new EHR implementations, it is necessary to focus on mitigating doubts about the rationale for such systems, to stress how EHRs are relevant for physician jobs, and to demonstrate that they are easier to use than nonusers might expect. This cannot be done without involving the end users in the design and evolution of EHRs, so they present highly usable and easy-to-learn interfaces and support decision-support capabilities that are needed by physicians. Further research is expected to deepen the findings from this study. Understanding physician perceptions of EHRs is a critical issue that must be addressed before better systems can be designed and adopted by the medical community.

Acknowledgments

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada, and the Canadian Institutes of Health Research. Neither had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Social Sciences and Humanities Research Council of Canada or the Canadian Institutes of Health Research.

Conflicts of Interest

None declared

Authors' Contributions

These authors contributed equally.

References

1. Davis K, Doty MM, Shea K, Stremikis K. Health information technology and physician perceptions of quality of care and satisfaction. Health Policy 2009 May;90(2-3):239-246. [doi: 10.1016/j.healthpol.2008.10.002] [Medline: 19038472]
2. Wright A, McGlinchey EA, Poon EG, Jenter CA, Bates DW, Simon SR. Ability to generate patient registries among practices with and without electronic health records. J Med Internet Res 2009;11(3):e31 [FREE Full text] [doi: 10.2196/jmir.1166] [Medline: 19674961]

3. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin CT. Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 2005;7(2):e13 [FREE Full text] [doi: 10.2196/jmir.7.2.e13] [Medline: 15914460]
4. Kazemi A, Fors UG, Tofighi S, Tessma MK, Ellenius J. Physician order entry or nurse order entry? Comparison of two implementation strategies for a computerized order entry system aimed at reducing dosing medication errors. J Med Internet Res 2010;12(1):e5 [FREE Full text] [doi: 10.2196/jmir.1284] [Medline: 20185400]
5. Gunter TD, Terry NP. The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. J Med Internet Res 2005;7(1):e3 [FREE Full text] [doi: 10.2196/jmir.7.1.e3] [Medline: 15829475]
6. Venkatesh V, Speier C, Morris MG. User acceptance enablers in individual decision making about technology: toward an integrated model. Decis Sci 2002;33(2):297-316.
7. Jensen TB, Aanestad M. Hospitality and hostility in hospitals: a case study of an EPR adoption among surgeons. Eur J Inf Syst 2007;16:672-680. [doi: 10.1057/palgrave.ejis.3000713]
8. Klein R. An empirical examination of patient-physician portal acceptance. Eur J Inf Syst 2007;16:751-760. [doi: 10.1057/palgrave.ejis.3000719]
9. Koru G, El Emam K, Neisa A, Umarji M. A survey of quality assurance practices in biomedical open source software projects. J Med Internet Res 2007;9(2):e8 [FREE Full text] [doi: 10.2196/jmir.9.2.e8] [Medline: 17513286]
10. Leonard KJ, Sittig DF. Improving information technology adoption and implementation through the identification of appropriate benefits: creating IMPROVE-IT. J Med Internet Res 2007;9(2):e9 [FREE Full text] [doi: 10.2196/jmir.9.2.e9] [Medline: 17513287]
11. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478.
12. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340.
13. Venkatesh V, Davis F. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
14. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley; 1975.
15. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: toward an integrative view. Inform Manag 2006;43:350-363. [doi: 10.1016/j.im.2005.08.006]
16. DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care: a national survey of physicians. N Engl J Med 2008 Jul 3;359(1):50-60 [FREE Full text] [doi: 10.1056/NEJMsa0802005] [Medline: 18565855]
17. Moore GC, Benbasat I. Development of an instrument to measure the perceptions of adopting an information technology innovation. Inf Syst Res 1991;2(3):192-222. [doi: 10.1287/isre.2.3.192]
18. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35(8):982-1002. [doi: 10.1287/mnsc.35.8.982]
19. Cocosila M, Archer N, Yuan Y. Early investigation of new information technology acceptance: a perceived risk-motivation model. Commun Assoc Inf Syst 2009;25(30) [FREE Full text]
20. Lapointe L, Rivard S. Getting physicians to accept new information technology: insights from case studies. CMAJ 2006 May 23;174(11):1573-1578 [FREE Full text] [doi: 10.1503/cmaj.050281] [Medline: 16717265]
21. Featherman MS, Pavlou P. Predicting e-services adoption: a perceived risk facets perspective. Int J Hum Comput Stud 2003;59(4):451-474. [doi: 10.1016/S1071-5819(03)00111-3]
22. Wu JH, Wang SC. What drives mobile commerce? An empirical evaluation of the revised technology acceptance model. Inf Manag 2005;42(5):719-729. [doi: 10.1016/j.im.2004.07.001]
23. Lim N. Consumers' perceived risk: Sources versus consequences. Electron Commer Res Appl 2003;2:216-228.
24. Laroche M, McDougall GHG, Bergeron J, Yang Z. Exploring how intangibility affects perceived risk. J Serv Res 2004;6(4):373-389.
25. Bontis N, Crossan MM, Hulland J. Managing an organizational learning system by aligning stocks and flows. J Manag Stud 2002;39(4):437-469. [doi: 10.1111/1467-6486.t01-1-00299]
26. Jöreskog KG, Wold HOA. The ML and PLS techniques for modeling with latent variables. In: Jöreskog KG, Wold HOA, editors. Systems Under Indirect Observation. Volume 139. New York, NY: North-Holland Publishing Company; 1982:263-270.
27. Chin WW, Newsted PRHoyle RH. Structural equation modeling analysis with small samples using partial least squares. In: Statistical Strategies for Small Sample Research. Thousand Oaks, CA: Sage Publications; 1999:307-337.
28. Gefen D, Straub D. A practical guide to factorial validity using PLS-Graph: tutorial and annotated example. Commun Assoc Inf Syst 2005;16:91-109.
29. Ahuja MK, Thatcher, JB. Moving beyond intentions and toward the theory of trying: effects of work environment and gender on post-adoption information technology use. MIS Q 2005;29(3):427-459.

30. Chin WW. University of Houston, CT Bauer College of Business. 2004. Frequently Asked Questions: Partial Least Squares & PLS-Graph URL: http://disc-nt.cba.uh.edu/chin/plsfaq/plsfaq.htm [accessed 2010-11-27] [WebCite Cache ID 5uZMSOxrz]
31. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York, NY: J Wiley & Sons; 1987.
32. Bontis N. National intellectual capital index: a United Nations initiative for the Arab region. J Intellect Capital 2004;5(1):13-39. [doi: 10.1108/14691930410512905]
33. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Marketing Res 1981;18(1):39-50.
34. Jarvenpaa SL, Shaw TR, Sandy DS. Toward contextualized theories of trust: the role of trust in global virtual teams. Inf Syst Res 2004 Sep;15(3):250-267. [doi: 10.1287/isre.1040.0028]
35. Rogers EM. Diffusion of Innovations, 5th edition. New York, NY: Free Press; 2003.
36. Bontis N, Keow WCC, Richardson S. Intellectual capital and business performance in Malaysian industries. J Intellect Capital 2000;1(1):85-100. [doi: 10.1108/14691930010324188]
37. Chin WW. The partial least squares approach for structural equation modeling. In: Marcoulides A, editor. Modern Methods for Business Research. Mahwah, NJ: Lawrence Erlbaum; 1998:295-336.

Abbreviations

EHR: electronic health record
IT: information technology
PLS: partial least squares
SEM: structural equation modeling
TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 04.01.11; peer-reviewed by A Kazemi; comments to author 02.02.11; accepted 14.06.11; published 12.08.11.

Please cite as:
Archer N, Cocosila M
A Comparison of Physician Pre-Adoption and Adoption Views on Electronic Health Records in Canadian Medical Practices
J Med Internet Res 2011;13(3):e57
URL: http://www.jmir.org/2011/3/e57/
doi:10.2196/jmir.1726
PMID:21840835

Review

Quality of Online Pharmacies and Websites Selling Prescription Drugs: A Systematic Review

Grazia Orizio[1], MD, PhD; Anna Merla[1], MD; Peter J Schulz[2], PhD; Umberto Gelatti[1], MD

[1]Section of Hygiene, Epidemiology and Public Health, Department of Experimental and Applied Medicine, University of Brescia, Brescia, Italy
[2]Institute of Communication and Health, University of Lugano, Lugano, Switzerland

Corresponding Author:
Grazia Orizio, MD, PhD
Section of Hygiene, Epidemiology and Public Health
Department of Experimental and Applied Medicine
University of Brescia
Viale Europa 11
Brescia, 25123
Italy
Phone: 39 030 3717697
Fax: 39 030 3717688
Email: graziamail@gmail.com

Abstract

Background: Online pharmacies are companies that sell pharmaceutical preparations, including prescription-only drugs, on the Internet. Very little is known about this phenomenon because many online pharmacies operate from remote countries, where legal bases and business practices are largely inaccessible to international research.

Objective: The aim of the study was to perform an up-to-date and comprehensive review of the scientific literature focusing on the broader picture of online pharmacies by scanning several scientific and institutional databases, with no publication time limits.

Methods: We searched 4 electronic databases up to January 2011 and the gray literature on the Internet using the Google search engine and its tool Google Scholar. We also investigated the official websites of institutional agencies (World Health Organization, and US and European centers for disease control and drug regulation authorities). We focused specifically on online pharmacies offering prescription-only drugs. We decided to analyze and report only articles with original data, in order to review all the available data regarding online pharmacies and their usage.

Results: We selected 193 relevant articles: 76 articles with original data, and 117 articles without original data (editorials, regulation articles, or the like) including 5 reviews. The articles with original data cover samples of online pharmacies in 47 cases, online drug purchases in 13, consumer characteristics in 15, and case reports on adverse effects of online drugs in 12. The studies show that random samples with no specific limits to prescription requirements found that at least some websites sold drugs without a prescription and that an online questionnaire was a frequent tool to replace prescription. Data about geographical characteristics show that this information can be concealed in many websites. The analysis of drug offer showed that online a consumer can get virtually everything. Regarding quality of drugs, researchers very often found inappropriate packaging and labeling, whereas the chemical composition usually was not as expected in a minority of the studies' samples. Regarding consumers, the majority of studies found that not more than 6% of the samples had bought drugs online.

Conclusions: Online pharmacies are an important phenomenon that is continuing to spread, despite partial regulation, due to intrinsic difficulties linked to the impalpable and evanescent nature of the Web and its global dimension. To enhance the benefits and minimize the risks of online pharmacies, a 2-level approach could be adopted. The first level should focus on policy, with laws regulating the phenomenon at an international level. The second level needs to focus on the individual. This approach should aim to increase health literacy, required for making appropriate health choices, recognizing risks and making the most of the multitude of opportunities offered by the world of medicine 2.0.

(J Med Internet Res 2011;13(3):e74) doi:10.2196/jmir.1795

KEYWORDS

Internet; pharmaceutical preparations; public health; review; online pharmacies

Introduction

> *Neo, which pill would you choose? You take the blue pill—the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill—you stay in Wonderland and I show you how deep the rabbit-hole goes[1].*

Although their choice is much less metaphysical than the question posed to Neo in the science fiction movie *The Matrix*, but still important in terms of health care delivery, consumers nowadays can make another decision: they can choose a pill sold at their local pharmacy, as they have always done, or they can choose one from the Web, by purchasing it from a "cyberpharmacy" or "online pharmacy" [2]. An online pharmacy is a company that sells pharmaceutical preparations, including prescription-only drugs, via online ordering and mail delivery, although—as the evidence will show—very few of them behave like a proper "pharmacy" and many of them are not licensed.

The online sale of drugs started in the late 1990s and has expanded so much that the US Food and Drug Administration (FDA) has implemented an entire section on its website dedicated to "Buying medicines over the Internet" [3,4]. The World Health Organization (WHO) is faced with this issue in the context of counterfeit medicines, which "pose a public health risk" [5-7]. Another aspect that is important to consider is that the Internet can facilitate access and thus support abuse of prescription drugs [8].

It is very difficult to estimate the number of online pharmacies and people buying online, the volume of drugs traded, and the revenue and profits generated by such a hidden business. Moreover, the geographical distribution of the phenomenon seems to be very heterogeneous. With regard to the number of online pharmacies, MarkMonitor in a 2009 press release claimed to have found nearly 3000 websites selling prescription medicines, while a 2010 review by the US National Association of Boards of Pharmacy (NABP) investigated the characteristics of 5859 Internet outlets selling prescription medications [9,10]. Regarding access to drugs online, the 2006 Online Health Search, a US survey by the Pew Internet & American Life Project, showed that "prescription or over-the-counter drugs" was the fifth most widely searched health topic on the Internet [11]. Another US telephone survey concluded that 4% of Americans had purchased prescription drugs on the Internet [12]. Very few estimates regarding the revenue and profits of this phenomenon are available, there being great variability in methods and numbers [13-15]. Besides economic aspects, there is a legal issue and jurisdictional consequences: cases of law enforcement acts and legal prosecutions have been reported in the literature [16-18].

This new market has undoubted advantages for patients: access to drugs for the disabled or housebound, access 24 hours a day, a virtually unlimited number of products available, relative privacy, which may encourage patients to ask questions about embarrassing issues, and more affordable prices [19-21]. But direct access to health services, especially drugs, poses a hazard to consumers because it is difficult to determine whether drugs purchased online are counterfeit, unapproved, or illegal [22]. Besides, the inappropriate use of medicines, the limited or nonexistent opportunity for advice (which blurs the line between willful abuse and unknowing misuse), and the risk of increased antibiotic resistance arising from their misuse have also been suggested as negative consequences of online purchase of medication [23,24]. What is more, the chance to circumvent prescription boundaries can be a potential disruptor at several levels, at both an individual and a public health level. At an individual level, this phenomenon can influence the doctor–patient relationship [25,26]. At a public health level, since each country has a unique system, access to drugs from abroad can disrupt the delicate equilibrium that leads to a certain drug price on the basis of taxation, copayment, reimbursement, and negotiation with industry [27].

Previous Reviews

To our knowledge, 2 reviews about online pharmacies in general and 3 others on specific aspects related to online pharmacies are available. With regard to general reviews, Fung et al [20] searched material published between 1997 and 2002 using 3 scientific databases. They identified 139 articles, although they found that "many of the articles reported about a specific legal case involving an online pharmacy." Although this review is wide-ranging, it is important to note that it dealt with papers issued 9 or more years ago (2002), a considerable length of time in such a dynamic world as that of the Internet and e-commerce. The most recent review, issued in 2009, is by Nielsen and Barratt [28]. They reviewed the literature on prescription drug misuse through the Internet, focusing on online supply, online monitoring of drug use trends, and electronic prescription monitoring [28]. The part relevant to our review is that on Internet supply; although the work is valuable, the article selection method was not described in detail, making it impossible to establish which databases were screened, with which keywords, and the time of publication. Since the review was submitted for publication on February 5, 2008, we can deduce it explored articles published up to 2007, as confirmed by looking at the references. Besides, the review focused only partially on online supply (referring to only 14 papers in connection with this topic).

Other reviews tackled specific issues but were unable to give an overall picture of the phenomenon. A recent review examined counterfeit phosphodiesterase type 5 inhibitors (PDE5Is) for the treatment of erectile dysfunction, similar to the 2000 review of sildenafil and the Internet [29,30]. The review by Baert and De Spiegeleer gives an overview of the different quality attributes that can be evaluated to gain a complete understanding of the quality of the pharmaceutical product traded on the Internet, as well as the current analytical techniques that serve this objective [31].

Objectives

This review sought to investigate the available evidence on the phenomenon of online pharmacies. We report data on the 3 main areas on which the literature focuses: the characteristics of the websites, the quality of pharmaceutical products purchased online, and the number of consumers and their characteristics.

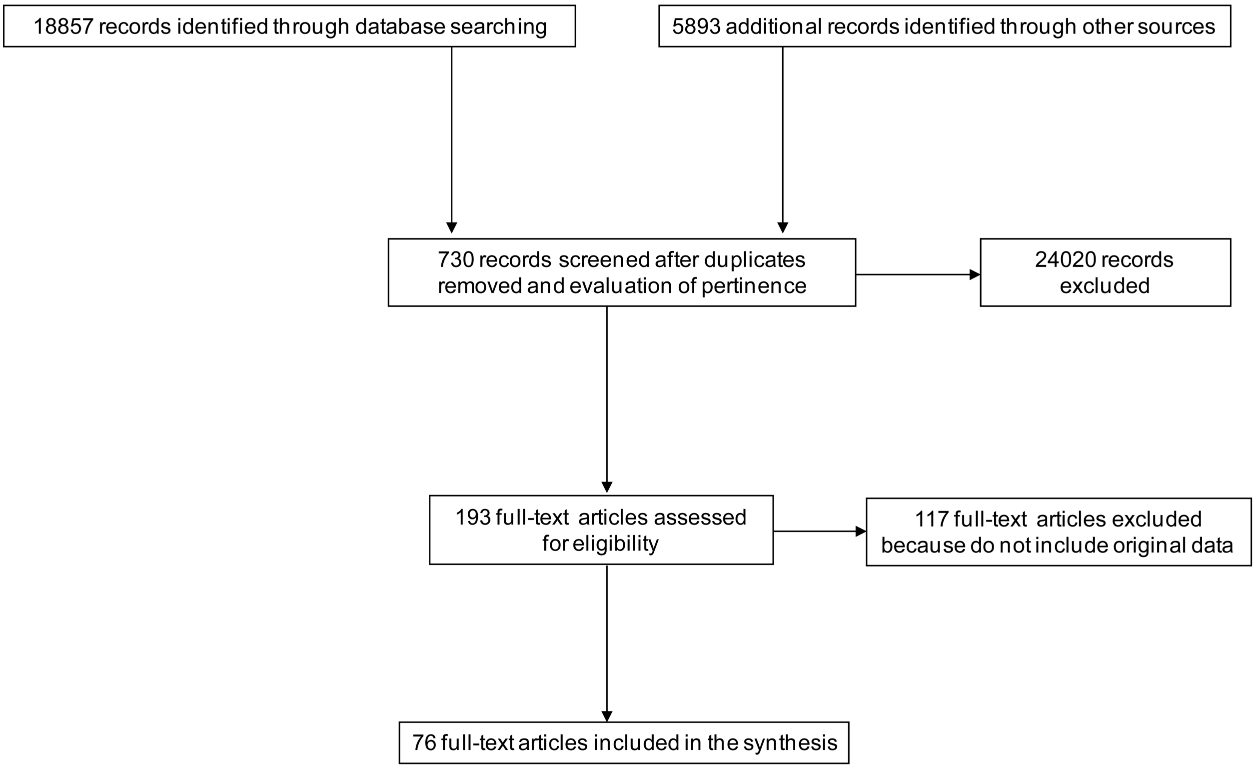
Methods

Search Strategy

The literature search covered the period up to January 2011. The search was performed on 3 sources: electronic databases, search engines, and institutional websites. First, we searched, without any limitations as to publication date, the following electronic databases: PubMed [32], ISI Web of Knowledge [33], Science Direct [34], and PsycInfo [35]. Second, we searched for gray literature on the Internet using the Google search engine [36] and its tool Google Scholar [37]. We used Google because it is the most widely used search engine [38,39]. Lastly, we investigated the institutional websites of the WHO [40], the WHO European Region [41], the WHO American Region [42], the US Centers for Disease Control and Prevention [43], the FDA [44], the European Centre for Disease Prevention and Control [45], and the European Medicines Agency [46]. We investigated all the results obtained by the databases and the institutional websites but considered only the first 500 results for each keyword appearing in Google and Google Scholar, because the number of relevant articles declined substantially after the first 300 results and because this search engine displays results by relevance using a link analysis system or algorithms [47]. We used the following search terms for each website and database analyzed: “drugs and internet,” “drug/s online,” “online pharmacy/ies,” and “internet pharmacy/ies.” We scanned the reference lists for relevant articles up to the second level, and we considered the “related articles” of relevant ones in the PubMed database.

The database search identified 18,857 records, and other sources (search engines and institutional websites) gave 5893 additional records. Screening of these 24,750 records led to 730 articles, excluding duplicates and nonpertinent results. It is important to point out that such a drop in numbers depends mainly on the use of multiple key words, which are often very similar, which were used in order not to miss any pertinent studies. This resulted in a notable “noise effect,” thereby decreasing the specificity and increasing the sensitivity of our search strategy. An in-depth analysis of the 730 selected articles produced 193 eligible ones that were pertinent to the study and fit the inclusion criteria. Of these, 117 where excluded from the analysis and are listed in Appendix 1, giving a final sample of 76 full articles for study. Figure 1 shows the selection process.

Figure 1. Paper selection algorithm.



Inclusion Criteria and Coding of Contents

We included all articles relevant to the subject of the research—namely, online pharmacies, their characteristics, their products, and their consumers. We selected only articles dealing with the sale of prescription-only drugs and with websites that presented themselves as pharmacies: the purpose was to stay within the sphere of substances that are supposed to involve a doctor–patient relationship. Articles regarding only over-the-counter medicines, complementary medicines, herbal remedies, supplements, and drugs of abuse were excluded. If the researchers analyzed websites selling prescription and over-the-counter drugs, we considered websites selling prescription drugs if it was possible to identify them. We decided to deal only with prescription drugs, although over-the-counter substances can also have negative effects on people's health, despite the no-harm claims made by their producers, as several clinical cases demonstrate [48,49]. Although examining nonprescription drugs was not an objective of this review, it should be borne in mind that some case reports showed the presence of prescription drugs even in products that did not claim to contain them [50,51]. This means that prescription-only drugs can be distributed through channels in which the active substances do not appear.

As an additional inclusion criterion, we selected articles in English that had the abstract or the full text available. We included only scientific articles, which means that we excluded popular articles published in daily newspapers, and in weekly and monthly magazines.

We classified the articles according to whether they reported original data. We selected only the articles reporting original data, which means that we excluded articles lacking original data, which means those with only a speculative discussion about the problem or only citing data from other studies; these are, for example, editorials, letters, comments, articles about regulation issues, and reviews. However, to make this debate easily available to the reader, we have listed all articles without original data in Multimedia Appendix 1. A discussion of all the reviews we found is included in the introduction.

The original data are described according to 3 main subjects: types and characteristics of online pharmacies, drugs purchased online, and online pharmacy consumer data, which included case reports on complications occurring in consumers of drugs purchased online. Some articles with original data covered more than 1 of these subjects and were consequently allocated to more than 1 group. Each of these categories is described below.

Types and Characteristics of Online Pharmacies

If they were available, we recorded the number of online pharmacies analyzed in each study, year of data collection, willingness to dispense pharmaceuticals with or without a prescription, availability of a physician's assistance or online medical consultation, disclosure of contact details, geographical location, delivery conditions, types of medicines available, availability of drug information, prices of online drugs and overall costs, sales-promotion strategies, how long websites were accessible, privacy and disclaimer statements, date of last website update, and presence of quality certifications (for instance, Verified Internet Pharmacy Practice Sites [VIPPS] by the NABP; the Health on the Net Foundation [HON] code; and the Joint Commission on Accreditation of Healthcare Organizations [JCAHCO]).

Quality of Drugs Purchased Online

We recorded studies in which the researchers bought prescription drugs online and evaluated the actual purchase and its characteristics. We summarized data regarding the type of drug ordered, the response rate, the quality of the process, and the drugs purchased. Regarding process characteristics, we recorded prescription requirements, management of the online questionnaire, money transactions, and subsequent advertising; with regard to drug quality, we recorded data about packaging and instructions, and chemical composition.

Consumers Buying From Online Pharmacies

We described articles dealing with the number of people purchasing drugs online, which was estimated by means of questionnaires or interviews. Researchers attempted to list the most frequently requested drugs, the main reasons for buying pharmaceutical products online, the importance of the location of online pharmacies they bought from, and the perceived risks related to this practice. In addition, we classified in this section articles that reported clinical cases of adverse effects to active substances and drugs purchased via the Internet as an indicator of this phenomenon.

Results

We selected 193 relevant articles: 76 articles with original data (39%), and 117 articles without original data (editorials, regulation articles, or the like) including 5 reviews. Articles with original data concerned samples of online pharmacies in 47 cases, online drug purchase in 13, consumer characteristics in 15, and case reports of adverse effects of online drugs in 12.

Types and Characteristics of Online Pharmacies

We selected 47 articles about online pharmacies. All of them are shown in Table 1, except 5 that had no data in addition to the number of online pharmacies found (3 articles), or did not clearly discuss the theme of selling drugs on the Internet (2 articles). The first 3 articles are by Schifano et al [52], who found in the Psyconaut 2002 EU Project 165 websites offering the possibility to purchase drug-related items, Schepis et al [53], who assessed the availability of stimulants over the Internet as a function of specific search terms used in the search engine, and Lott and Kovarik [54], who assessed the availability of the dermatological medications isotretinoin and terbinafine over the Internet from illicit commercial sites. We do not show in Table 1 the other 2 articles, which focus on the assessment of community pharmacy websites in Turkey and Switzerland [55,56]. The former made only a passing reference to the presence of e-commerce services, but it was not possible to determine whether they were actually selling prescription drugs; the latter also assessed the presence of e-commerce services without referring to what was actually sold online.

Table 1. Contents of articles about online pharmacies, listed in alphabetical order according to first author; the presence of each item is indicated when studied and the percentage is reported when comparable; “X” indicates that the item was analyzed but could not be tabled

First author, year of publication	Prescription requirement (%)	Online questionnaire (%)	Contact details (%)	Geographical location (%)	Delivery	Drugs offered	Drug information (%)	Prices	Marketing strategies	Quality	How long websites were accessible	Privacy policy
Armstrong, 1999 [57]	0	50		100		X[a]	55	X	X		X	X
Arruanda, 2004 [58]	30		99		X	X	35	X		X	X	
Bate, 2010 [59]					X	X[a]		X				
Bessel, 2002 [60]	81	12	66	79	X	X			X	X		X
Bloom, 1999 [61]	X[b]	X[b]		11		X		X				
Bloom, 2006 [62]	X[b]	X[b]		92		X		X				X
CASA[c], 2008 [63]	15	41			X	X[a]					X	
Cicero, 2008 [64]						X[a]		X				
European Alliance, 2008 [65]	10	16	42	16					X	X		X
Eysenbach, 1999 [66]	9	50		X		X[a]		X	X			X
Forman, 2003 [67]	0[a]			X		X[a]						
Forman, 2006 [68]	0[a]			X		X[a]						
Forman, 2006 [69]	0[a]	50				X[a]						
Forman, 2006 [70]	0[a]	52				X[a]			X	X		
Gallagher, 2010 [71]	7	59	100	59								
GAO[d], 2000 [72]	58	28	81				61					X
GAO[d], 2004 [73]	34	40		X		X						
Gernburd, 2007 [74]	0						100	X	X		X	
Gurau, 2005 [75]	34	59	100						X			X
Holmes, 2005 [76]							X[e]					
Koong, 2005 [14]						X						
Kunz, 2010 [77]										X		
Kuzma, 2011 [78]										X		

First author, year of publication	Prescription requirement (%)	Online questionnaire (%)	Contact details (%)	Geographical location (%)	Delivery	Drugs offered	Drug information (%)	Prices	Marketing strategies	Quality	How long websites were accessible	Privacy policy
Levaggi, 2009 [79]	19	67		44	X	X		X	X			
Littlejohn, 2005 [80]	10	X				X		X				
Mahé, 2009 [81]	X					X[a]		X	X			
Mainous, 2009 [82]	0[a]	64			X	X[a]						
Makinen, 2005 [27]	X	X		X[a]	X	X				X		
Memmel, 2006 [83]	0	75		X		X[a]		X				
NABP[f], 2010 [10]	4	58		47		X				X		
Orizio, 2009 [84]	19	56		43	X	X	X		X			
Orizio, 2009 [85]	0[a]	100[a]		28								
Orizio, 2010 [86]	22	45				X	X		X	X	X	
Peterson, 2001 [87]	88		67	X[a]		X	X			X		X
Peterson, 2003 [88]				X[a]		X				X	X	X
Quon, 2005 [89]				X[a]		X		X				
Raine, 2009 [90]	17	41		43	X[a]	X[a]	X					
Schifano, 2006 [91]	0	10				X[a]						
Soares Gondim, 2007 [92]			81	X[a]			X			X		
Tsai, 2002 [93]	0[a]	81	37	98		X[a]	X	X		X		
Veronin, 2007 [94]				X[a]								
Wagner, 2001 [95]	100		100	X[a]	X	X[a]	100	X				

[a] See other specific inclusion criteria, fourth and fifth column in Multimedia Appendix 2, which lists the characteristics of the studied samples of online pharmacies.

[b] Does not specify whether the prescription is an original one from the customer's physician or an online one.

[c] National Center on Addiction and Substance Abuse.

[d] US Government Accountability Office.

[e] The study aimed to evaluate the responses provided by the “ask the pharmacy” service.

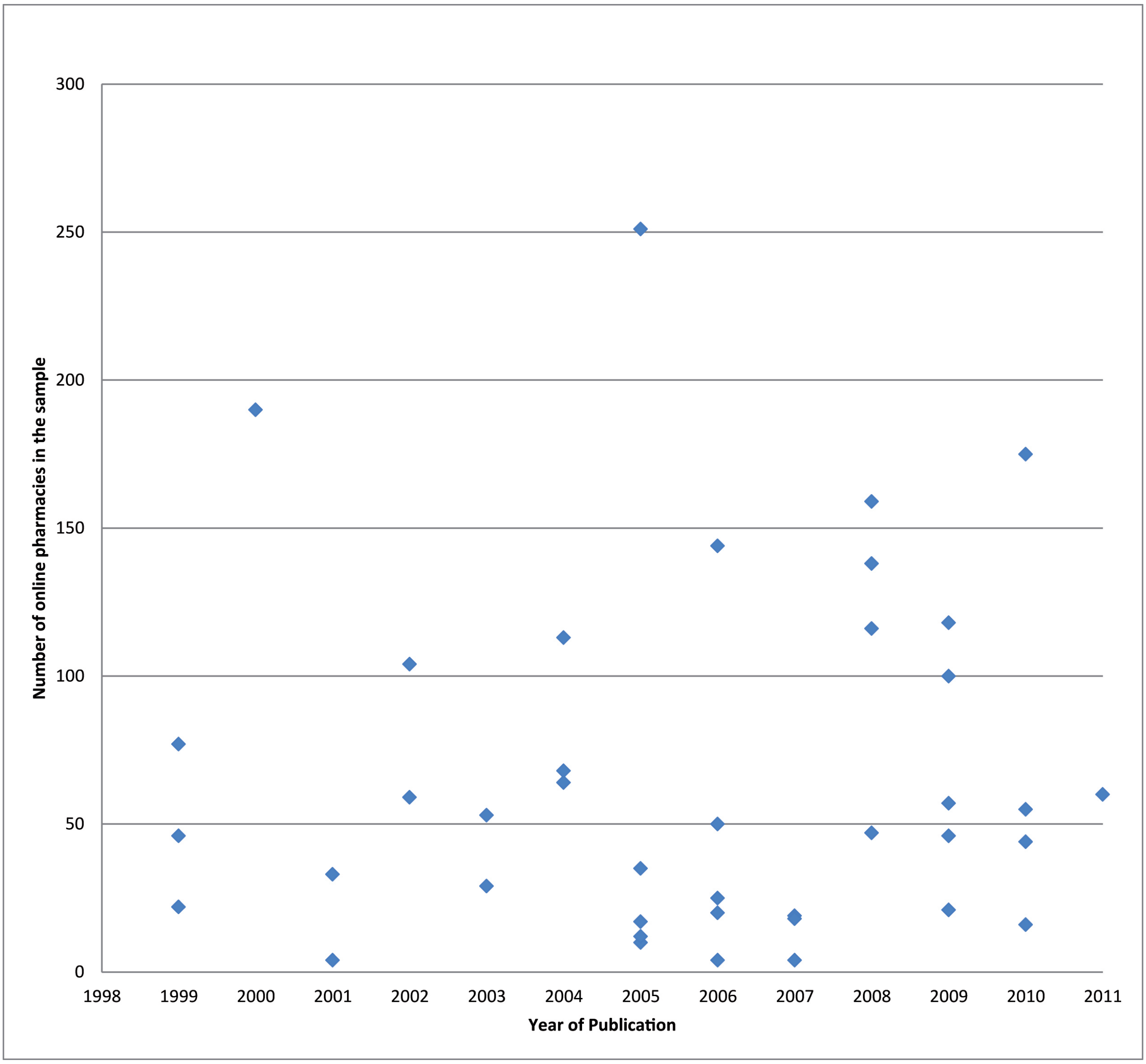
[f] National Association of Boards of Pharmacy.

The remaining 42 articles [10,14,27,57-95] dealing with online pharmacy characteristics are listed in Table 1, which shows the first author, year of publication, and the main features studied; due to the descriptive nature of the data it was not possible to report numerically all the results presented in the following subsections. When the numeric percentages of presence of the item were available, they are given in the table; the letter X means it was not possible to include data in the table or compare them (when, for instance, the groups they refer to are different and hence not comparable). The kind of information represented by the X is always described in detail in the text. Appendix 2 gives the sample size and the methods of selection regarding these 42 articles.

The works were published between 1999 [57,61,66] and 2011 [78], as shown in Figure 2, which displays the articles described in this section by year of publication and number of online pharmacies in the study sample. Less than half of the studies analyzed had more than 50 online pharmacies in their samples. Samples were obtained using a variety of methods, the most frequent being to look for online pharmacies using online search engines via various keywords. Some researchers used different sample selection methods: one found websites from received spam [74], while others looked for the websites indicated by other sources, such as the first one listed in the Best Online Pharmacy guide by epharmacyfinder.com [14], or the VIPPS list [59,76], the Pharmacychecker.com list [89], or the Top 100 Retailer [77].

Online pharmacy samples varied hugely in size, from 4 [83,94,95] to 5859 [10]. The sample selection had different inclusion criteria: many studies focused on online pharmacies selling numerous kinds of drugs, while others selected online pharmacies offering a single medication or 1 class; some selected only online pharmacies not asking for a prescription, and others only online pharmacies based in specific countries. These particular inclusion criteria are shown in Multimedia Appendix 2 and are described in the subsection dealing with each issue.

Figure 2. Articles on online pharmacy characteristics (included in Table 1) by year of publication and number of online pharmacies in the study sample (data not reported because out of range: NABP sample of 5859 online pharmacies [10]).



Prescription Requirement

One of the most controversial and widely studied features is prescription requirement. For the first time, online pharmacies are providing easy access to traditionally controlled (according to each country's regulations) substances, such as pharmaceutical products, which in regulated systems need an original medical prescription before they can be bought. The peculiar distribution chain for prescription drugs is due to the unique nature of pharmaceutical products; the philosophy behind prescription is that consumers are not skilled enough to make their own choice, but need to be given a prescription by a health professional trained to make a risk–benefit evaluation [2,24]. In addition, the medical follow-up of a drug treatment is crucial. Furthermore, there are problems regarding strict quality control for substances that can determine life or death. Last but not least, there are economic implications in ruled systems where drug prices are negotiated and partially funded by the health system [27].

Due to the different methods of sample selection, distribution varies greatly in the articles. Some online pharmacy samples were specifically selected to have only websites not asking for a prescription from the customer's physician [57,67-70, 82,93]. In the samples that did not use prescription as a criterion for inclusion or exclusion, the prescription requirement varied widely: all websites asked for an original medical prescription in Wagner et al [95] (who selected only US-based online pharmacies), more than half did so in other samples [60,72,87], less than half [10,58,63,65,66,71,73,75,80,84,86,90] in still others, and none in a few [74,83,91]. In Bloom and Iannacone [61,62] it was not possible to determine whether the required prescription was an original one by the customer's physician or replaced by an online health evaluation as explained below.

Presence of an Online Questionnaire

Some of the online pharmacies not asking for a prescription replaced it with an online health status evaluation, performed by means of an "online questionnaire." Researchers reported a percentage of online pharmacies offering and/or requiring an online prescription in a range of 10% to 81% [91,93]. Those who looked deeper into this issue found that the identity of the professional who made the prescription was not usually provided [57], whereas they often declared that the questionnaire reviewer was a physician [57,61,62,66,84] or, more rarely, a pharmacist [27,84]. Gallagher and Chapman [71] reported that, of the 26 websites using a questionnaire, 1 involved multiple-choice check-box answers and drop-down menu answers, and the others used questionnaires that allowed users to type in an answer. Eysenbach [66]—who attempted to buy the products—found that the identity of the physician who made out the prescription was revealed in 2 out of 10 orders. The US Government Accountability Office (GAO) [72,73] found the collaboration of licensed physicians, although when they verified these licenses [72] almost all of them were nonexistent or not valid for all the US states where they were declared to be. No study found the use of a validated or standardized questionnaire. Orizio et al [85] did a subanalysis of a previous study [84] that focused on the characteristics of online questionnaires; it reported that the questionnaire was already filled in with negative answers in 70% of online pharmacies using this tool, and that in only 53% of cases were different questions asked for different products ordered. Since the questionnaires were often incomplete, the authors concluded that they appeared to aim more at giving the consumer a false sense of health assurance than at actually assessing health status. The questionnaire frequently investigated personal characteristics, allergies, medical conditions, current therapies, and medical history [57,61,62,66,72,73,85].

Contact Details

Contact details were revealed in the majority (ranging from 100% [71,75,95], or a little less among online pharmacies when this item was analyzed [58,60,72,75,87,92,95], to 66% [60]). Only Tsai et al [93] found just 37% of pharmacies giving a telephone number for costumers' enquiries, and the European Alliance for Access to Safe Medicines' [65] analysis found that 42% provided a working telephone number. Bessel et al [60] pointed out that only 35% of their sample published the owners' or the director's names, and European Alliance [65] found that 94% of the sample did not have a named verifiable pharmacist to answer questions.

Geographical Location

Declaration of geographical location is an important feature with regard to transparency. Apart from the studies that focused only on websites based in a specific area, the sampled websites declared a geographical location ranging from 11% to 100% [47,61]; according to the results of Armstrong et al [57], Bessel et al [60], Bloom and Iannacone [62], Gallagher and Chapman [71], and Tsai et al [93], more than half of the samples mentioned where the company was based; the others found that less than half did [10,61,65,84,90]. Orizio et al [84] found that online pharmacies asking for a prescription were significantly more likely to declare their geographical location than were online pharmacies that did not, which were more often just virtual interfaces. Online pharmacies were most frequently based in the United States in virtually all the studies that investigated samples that were not restricted geographically [10,57,60,61,62,66,73,84]. Other studies focused on online pharmacies based in specific countries, such as the United States [87,88,95], the United States and Europe [27], the United States and Canada [89], Canada [94], or Brazil [92]. Some researchers looked for where the websites were registered, and they found the United States to be the most frequent location of domain registration [10,67,68,84,93]. Orizio et al [84], who compared the declared physical location with the registration domain, found that in only 55% of online pharmacies declaring their physical location did it correspond to the area of domain registration. Armstrong et al [57] found that US-based websites were more likely to ask for medical information and provide information about risk of treatment or its efficacy, and the GAO [73] found a better quality in websites based in the United States.

Delivery: Where From and Where To

The National Center on Addiction and Substance Abuse (CASA) [63] found that 24% of shipments were declared to be from the United States, and 36% of the sample gave no indications. In Orizio et al [84] 42% of the websites declared where they deliver from, the most frequent place being Asia (47%); agreement between declared physical location and delivery location was found in 31% of online pharmacies providing both details. The attempt by Bate and Hess [59] to purchase drugs showed several cases of shipments from a different location from what was indicated on the website.

Both Arruanda [58] and Bessel et al [60] found that two-thirds of the online pharmacy samples sold internationally, whereas all the websites sampled by Schifano et al [91] did. Raine et al [90] selected only online pharmacies delivering to the United Kingdom. Makinen et al [27] related the fact of shipping abroad to the prescription requirements: online pharmacies that asked for a prescription were more likely than those that did not to sell only in the country they were based in, and delivered anywhere in the world. Mainous et al [82] found that almost all the websites sold to the United States, followed by the United Kingdom (84%) and Canada (80%). None of the small sample of 4 US-based online pharmacies analyzed by Wagner et al [95] delivered internationally.

Drugs Offered by Online Pharmacies

What do online pharmacies have to offer? From an analysis of the literature it appears that online pharmacies have become more and more complex as time passes; whereas 10 years ago they tended to sell principally lifestyle drugs such as sildenafil, it looks today as if they offer virtually anything. Back in 1999, Bloom and Iannacone [61] studied a sample of online pharmacies in which the majority sold 1 or 2 drugs only, one for erectile dysfunction and the other for alopecia; in the same year Armstrong et al [57] found an entire sample of 77 websites selling sildenafil. In 2003 Arruanda [58] found that their sample was equally divided into 3 groups: selling 1 drug only, between 2 and 29 drugs (average 7), and selling 30 or more. The various studies reported several specific drug offers: sildenafil [59,60,80,84,86,87], benzodiazepines [84,86,87], painkillers [80,84,88], antibiotics [87,88], insulin [87,88], female hormones [87,88], antidepressants [80,84,86], alopecia medications [14,27,80], and obesity medications [59,80]. Bloom and Iannacone [62] reported that 160 separate medications were offered in their sample, which gives an idea of how widely differentiated the drug offer is. Quon et al [89] compared the offers of 12 Canadian-based Internet pharmacies with 3 US-based drug chain pharmacies for 44 different drugs. The GAO [73] found that some drugs were more widely available and easier to purchase (Celebrex, Lipitor, Viagra, and Zoloft) than others, which were available from fewer sources or were more difficult to obtain (Accutane and Clozaril). As we will see in more detail in the next section on actual purchase, what is hard to obtain—despite appearing to be available—is US class II and III opioid analgesics, as found in Peterson and colleagues' studies in 2001 [87] and 2003 [88] and confirmed by the GAO 2004 [73]. In the United States, "Under the Controlled Substances Act all substances that are regulated under existing federal law are placed in one of five schedules on the basis of the substances' medicinal value, harmfulness, and potential for abuse or addiction. Schedule I is reserved for the most dangerous drugs that have no recognized medical use, while Schedule V is the classification used for the least dangerous drugs." [72]. The NABP [10] survey showed that 14% of websites dispensed controlled substances as defined by the Act cited above, and 40% foreign or non-FDA-approved drugs.

Some studies focused on specific classes of drugs. They looked for Parkinson disease medications [95], 3 types of controlled substances (opioids, and central nervous system depressants and stimulants) [63], opiates [64,67-70]), analgesics [90], dextropropoxyphene (a painkiller) [91], antibiotics [82], and specifically ciprofloxacin [93], contraceptives [83], the erectile dysfunction medication sildenafil [57,66,71], and psoriasis medications [81].

Littlejohn et al [80] and Makinen et al [27] linked the drugs offered to the type of online pharmacy: "legitimate pharmacies" did not supply opioids or ritalin [80], selling only over-the-counter products and herbal, hygiene, and cosmetic products [27]; "lifestyle pharmacies" supplied erectile dysfunction and alopecia medications; and "no-prescription pharmacies" supplied virtually everything, including opioids [80] and unapproved pharmaceuticals [27]. Both generic and brand drugs were available on the market [14,58,79,89,95].

Presence of Information About the Drugs for Sale

In 1999 Armstrong et al [57] found that 55% of their sample included drug information. Several researchers who attempted to evaluate the presence of information on side effects found that a fairly consistent portion of online pharmacies, ranging from a quarter to a third, declared none of them [84,86,90,92,93]. Arruanda [58] found that 35% of his sample provided a service allowing buyers to consult experts about the use of medicine. In the GAO [72] sample, 61% of websites gave drug information. Wagner et al [95] found in all the online pharmacies in their small sample (4 US-based online pharmacies) more comprehensive information than received from the community store. Peterson [87] did not find a statistically significant difference between the types of pharmacies and the provision of drug information, although it should be noted that his sample was small for performing a group comparison (33 online pharmacies). Gernburd and Jadad's [74] research based on spam offers found that all of the websites in their sample made benefit claims and warned about potential side effects. Interestingly, an attempt to count the declared side effects of 4 drugs tracked in online pharmacies that asked for a prescription and those that did not revealed that no-prescription online pharmacies declared more side effects for amytriptiline, fluoxetine, and tramadol, but fewer for sildenafil, which—probably not by accident—is one of the most widely offered "lifestyle drugs" in online pharmacies [84,86].

Holmes et al [76] evaluated the quality of online pharmacy "ask the pharmacy" services regarding adverse effects, drug interaction, risk factors, drug information, and directions for use. They found that answers were received to only 51% of all questions submitted to the websites, and the percentage of correct answers provided for each of 22 response components ranged from 7% to 96%, with few differences in quality between VIPPS-approved and -unapproved websites.

Prices

Arruanda [58] evaluated the presence of drug price lists (found in 96% of the sample) and of "price comparisons with competitive pharmacies" (1%).

A comparison of online pharmacy and retail pharmacy prices had different results. Drugs offered online were more expensive in Bloom and Iannacone's studies [61,62], in Tsai and colleagues' [93] analysis, and in Cicero and colleagues' [64] purchase. Wagner et al [95], who compared prices of US retail versus US online pharmacies, found that the latter were cheaper for both generic and brand-name medications. Levaggi et al [79] found that drugs bought without a prescription cost more than with one. In Bate and Hess's [59] purchase, larger orders generally tended to have lower per-tablet/per-capsule prices, and were more prevalent among online pharmacies that had not been approved by the NABP. Interestingly, they found that Viagra offered on noncredentialed websites was on average far more expensive than from credentialed ones, and for all the drugs purchased (Lipitor, Celebrex, Nexium, and Zoloft), except Viagra, prices were higher at physical-location pharmacies. Quon et al [89] compared US retail prices with those of Canadian online pharmacies, and they found savings when purchasing from Canadian online pharmacies for the majority of brand drugs (3 exceptions in the 44-drug sample were medications for erectile dysfunction), whereas generics were more expensive. The same result is confirmed by an aggregate macroeconomic analysis based on IMS Incorporated data on prices, patents, and cross-border Internet pharmacies between online Canada and retail US pharmacies [15]. Memmel et al [83] found lower prices for contraceptives purchased online. Mahé et al [81] found that, with the exception of tazarotene, the average price of all the online psoriasis medicines analyzed was higher than the French retail price.

Additional costs have to be considered: the need to join clubs or member groups and pay a nonrefundable fee, with the risk of not finding the wanted drug after joining [64,80]; and delivery costs, with free delivery in some online pharmacies or for some purchases, but with charges depending on the type of shipping in other cases (standard, express, overnight) [57,58,61,62,66,74,79,89,95]. Online prescription could be another added cost [57]. It is interesting to note that, although a prescription is not necessary in marketing messages, Levaggi et al [79] found in their price analysis that the prescription has a value on the market; indeed, when you buy a drug without a prescription, the drug costs more.

Marketing Strategies

Only a few articles focused on the marketing strategies of online pharmacies. Levaggi et al [79] and Orizio et al [84] disclosed the persuasive statements more frequently used by websites to promote their products, and identified arguments regarding privacy, service and drug quality, price offers, reassurances that buying drugs online is legal, and the suggestion that you can obtain a drug while avoiding a visit to the doctor. In particular, privacy issues were about the use of personal data and discreet packaging; service quality statements regarded short delivery times, online tracking of the state of the orders, and—indirectly—displaying testimonials by people who had already bought online; and price offers referred to encouragement to buy bulk purchases (found by Armstrong et al [57], European Alliance for Access to Safe Medicines [65], and Mahé et al [81] as well), lower prices than in "brick and mortar" pharmacies, fidelity bonuses, free delivery, and special discounts [79]. Armstrong et al [57] reported that one-third of their sample pointed out the advantages of ordering online, including confidentiality, ease of ordering, and lower cost. Forman and Block [70] analyzed the implied legitimacy and credibility claims, which they divided into 3 types: medical legitimacy, found in 82% of online pharmacies (eg, pictures of lab coats, rx/health symbols, and pharmaceutical logos); legal legitimacy, found in 72% (eg, government logos, explicit and implicit claims of being legal, and FDA approval statements); and retailer legitimacy, found in 24% (eg, customer testimonials). They also reported claims about shipping regarding security (52%), secretiveness (70%), reshipment if seized (2%), risk of seizure by US customs (6%), free delivery (24%), and the delivery company (52%). All the websites in Gernburd and Jadad's [74] sample put forward benefit claims. One website in Eysenbach's [66] sample even offered to ship cimetidine together with a Viagra order because it causes "a 56% increase in plasma sildenafil concentration when co-administered". Gurau [75] pointed out that providing contact information may be a way to reduce the perceived risk of online transactions; moreover, statements about price, convenience, choice, and discreetness of service are more frequently used by online pharmacies not asking for a prescription than by those asking for one. Advertising prescription-only medicines was also found by Bessel et al [60] in 20% of online pharmacies. Orizio et al [86] pointed out that the marketing strategies adopted by online pharmacies enhance consumers' peripheral reflection: by analogically playing with the sale of other commodities, they magnify aspects of online trading that consumers might find convenient, but overshadow the nature and risks of the actual products they sell.

Quality

The presence of at least 1 quality certification was found to range from 12% to 13% [60,70,86] to 28% [58], but Tsai et al [93] did not find any websites displaying certificates. Researchers found certificates about the quality of the health contents and/or about security in the money transaction.

Regarding health content quality, Arruanda [58] found that the most frequent quality certificate was VIPPS [96] (9%), followed by the HON code [97] (5%). Bessel and colleagues' [60] sample had quality certification on 12% of the websites: the national pharmacy authority, the HON, and the JCAHCO [98]. Makinen et al [27] reported finding the HON code and VIPPS seal. Peterson's sample in 2001 [87] had 12% of websites with the VIPPS seal, rising to 18% in 2003 [88]. The European Alliance for Access to Safe Medicines [65] study found that 4% of the sample was licensed by a board of pharmacy or had an appropriate pharmacy listing, while 20% had a "stamp approval" from a recognized society or association, but they found when clicking on them that 86% gave a link to a bogus "approval" webpage. Soares Gondim and Borges Falcao [92] found that 15 of the 16 websites analyzed lacked the Brazilian National Health Surveillance Agency seal.

As regards the security of the money transaction, Arruanda [58] found the VeriSign seal [99] (8%), Forman and Block [70] found Verified by Visa and Master Card securicode logos, and Makinen et al [27] reported finding the TRUSTe seal [100]. As well, 24% of Peterson's sample in 2001 [87] incorporated secure socket layer technology, rising to 33% in 2003 [88].

A special study was performed by Kuzma [78] about the Web vulnerability of a random sample of 60 online pharmacies. She chose as her testing tool the N-Stalker Web Application Security Scanner 2009 Free Edition 7.0, which showed that a majority of worldwide online pharmacies do not provide adequate protection for their consumers, especially in cross-site scripting. The NABP [10] found that 17% of the websites in their sample did not have a secure site.

Kunz and Osborne [77] assessed the readability of 16 online pharmacies using Storytoolz; they found that the majority of the information provided on direct-to-consumer pharmaceutical websites is written at a level far higher than that which the average consumer can understand.

Website Time of Existence

Arruanda [58] found that the sample online pharmacies had been online for a length of time ranging from just over 1 year to 7 years. Armstrong et al [57] reported that in the 10 days between website identification and data collection 9 (12%) ceased operating. In the CASA study [63] only 2% of the 152 non-VIPPS anchor sites identified in 2004 were still operating in 2008. When Orizio et al [86] rechecked in 2008 the 118 websites found in 2007, 75% were working. The follow-up of Peterson's [87] study found that 1 year later 88% of the initial 33 online pharmacies were still working [88]. Gernburd and Jadad [74], who worked on links addressed by advertising spam, found that only 58% of the active link in health-related messages received during the first week of the study remained active at the end of the second week, whereas 26% were active at the end of the month.

Privacy Policy and Disclaimers

In several of the papers analyzed we found that some websites required the consumer to relieve the companies from all liability, from 100% of websites [62,66], to 68% [57], 50% [65], and 33% [72]. Bessel et al [60] found privacy statements, information disclaimers, and return policies on 40%, 31%, and 37% of the websites, respectively, while 37% of websites displayed none of these policies. The GAO [72] found privacy statements on 23% of the sample. Gurau [75] found that websites displayed their privacy policy more frequently when they asked for a prescription (98%) than if they asked for an online questionnaire to be filled out (86%) or asked for nothing (56%). Peterson [87] found that a privacy policy was present in a significantly different way in the various types of pharmacies that he identified (chain extension 100%, mail order 16%, online 64%, independent extensions 25%); the same author reported in 2003 [88] that several websites had added a privacy policy.

Last Website Update

Only 1 study investigated the date of the online pharmacies' last update, which was displayed by 4% of them (data not shown in Table 1) [60].

Drugs Purchased Online: Process Characteristics and Drug Quality

We found 13 studies [59,64-66,73,74,82,83,94,101-104] in which the researchers bought prescription drugs online and evaluated the actual dispatch and its characteristics, and the quality of the products received (Table 2). The response rate ranged from 30% [66] or over (39%) [74] to almost 100% [65,101] or 100% [59,73,74,102], except for the research by Cicero et al [64], who received nothing after ordering opioids, whereas his only order for tramadol was successful. It should be noted that the various groups of researchers ordered and purchased different types of drugs from different samples of online pharmacies. Some were interested in specific categories of drugs [64,66,82,83,94,101,102,103], whereas others ordered several active ingredients [59,65,73,74,104]. Cicero and colleagues' attempts to buy class II and III opioid analgesics did not result in their dispatch [64], but hormonal contraceptives and simvastatin were always delivered when ordered [83,94], despite a much lower number of total orders. The GAO [73], which received 75% of orders, showed the same result, as they found that top-selling drugs such as Celebrex, Lipitor, Viagra, and Zoloft were readily available from multiple Internet pharmacies, whereas other drugs, such as those with special safety restrictions (Accutane and Clozaril) and narcotic painkillers (Percocet, OxyContin, and Vicodin), were offered for sale by fewer Internet pharmacies or were otherwise more difficult to obtain.

Table 2. Articles on the quality of drugs purchased online, listed in alphabetical order according to the first author

First author, year of publication	Year(s) of data collection	Type of drugs ordered	Response rate (products received/number of orders)	Drug purchase characteristics	Drug quality characteristics
Bate, 2010 [59]	2009	Lipitor, Viagra, Celebrex, Nexium, Zoloft	Response rate not computable; 152 ordered drugs were received	Prescription requirement, money transaction	Packaging, chemical analysis
Cicero, 2008 [64]	2006	Opioid analgesics	0% (0/47) of “opioid scheduled” orders, 100 (1/1) purchase of tramadol	Prescription requirement, money transaction, subsequent advertisement	
Dean, 2010 [102]	Not declared	Dapoxetine	100% (1/1)		Chemical analysis
European Alliance, 2008 [65]	Not declared	18 different active ingredients[a]	94% (34/36)	Prescription requirement	Packaging, instructions, chemical analysis
Eysenbach, 1999 [66]	1999	Viagra	30% (3/10)[b]	Prescription requirement, management of online questionnaire, money transaction	
GAO[c], 2004 [73]	2004	13 different active ingredients[d]	75% (68/90)	Prescription requirement, money transaction	Packaging, instructions, chemical analysis
Gernburd, 2007 [74]	2006	13 different active ingredients[e]	39% 5/13[b]	Prescription requirement, money transaction	
Mainous, 2009 [82]	2008	Antibiotics	100% (1/1)	Prescription requirement	Instructions
Memmel, 2006 [83]	2004–2005	Hormonal contraceptives	100% (10/10)[f]	Prescription requirement, management of the online questionnaire	Packaging, instructions
Miller, 2001 [101]	1999–2000	Prescription and nonprescription contraceptives	96% (9/10 in 1999 and 15/15 in 2000)	Prescription requirement, money transaction	
Veronin, 2004 [103]	Not declared	Simvastatin	Response rate not computable; 5 ordered samples were received		Chemical analysis
Veronin, 2007 [94]	2006	Simvastatin	100% (4/4)[g]		Chemical analysis
Westenberger, 2005 [104]	Not declared	Fluoxetine, levothyroxine sodium, metformin hydrochloride, phenytoin sodium, warfarin sodium	Response rate not computable; 20 ordered samples were received		Packaging, instructions, chemical analysis

[a] The drugs purchased were Cialis, Levitra, Viagra, Propecia, Lipitor, Plavix, Seretide, Coversyl, Micardis, Spiriva, Zyprexa, Efexor, Risperdal, Aricept, Reminyl, Zoton, Reductil, and Mirapex.

[b] Orders made only on websites asking for an “online questionnaire” to be filled in to obtain an online prescription. The online questionnaire was completed by a fictitious patient with clear contraindications for sildenafil.

[c] US Government Accountability Office.

[d] The drugs purchased were Accutane, Celebrex, Clorazil, Combivir, Crixivan, Epogen, Humulin N, Lipitor, and OxyContin.

[e] The drugs purchased were Ambien, Celebrex, Cialis, Meridia, Nexium, Propecia, Soma, tramadol, Valium, Viagra, Xanax, Zithromax, and Zoloft.

[f] Orders made with different risk profiles.

[g] Orders from Canadian websites only.

Drug Purchase Characteristics

The actual purchase of drugs without having an original medical prescription was verified by all authors except for Dean et al [102], Veronin and Youan [103], Veronin et al [94], and Westenberger et al [104], who did not specify whether a prescription was used. Two studies evaluated the effectiveness of online questionnaires for assessing health status. The results highlighted the low performance of this tool, which allowed the purchase of contraindicated products by fictitious consumers in some or all orders [66,83].

Several authors found a lack of reliability in the business practices of some online pharmacies, linked to completed money transactions without actual receipt of the drugs [66,73,74,101] or having to join clubs, only to find no drugs available after having paid the necessary fee [64]. One study reported subsequent advertising by email and phone for more than 4 months after ordering [64].

Drug Quality Characteristics

Regarding drug quality characteristics, researchers evaluated different features: the packaging of the drugs purchased, the instructions included, and the chemical composition.

As to packaging characteristics, packaging showed problems in more than half of the drug samples in the GAO [73] study. Memmel et al [83] found that all packaging appeared legitimate and products were within the expiration date, but patient education material (instructions) did not always match the product. GAO [73] researchers found better quality instructions in samples from US and Canadian Internet pharmacies, which in all 47 samples included dispensing pharmacy labels that generally provided patient instruction for use and in 87% of cases included warning information, compared with other foreign pharmacies, whose labeling and facility of manufacture were not FDA approved in the majority (19/21) of cases. Westenberger et al [104] found that only 1 of 20 samples had final packaging, including package insert, similar to that of the US products. The final packaging of the remainder consisted of bubble wrap inside a paper envelope, a Styrofoam sheet inside a paper envelope, loose blister packs, capsules or tablets in clear plastic bags without labels, and capsules or tablets in opaque plastic containers or boxes with labels. Bate and Hess [59] reported that many drugs, including some that did not fail in spectrometry testing, had an unclear or problematic origin or problematic packaging. In the European Alliance for Access to Safe Medicines study [65], 47% of the drugs purchased had no packaging, and in 6% the packaging had been tampered with; half of the sample supplied a patient leaflet, which was in English in 80% of cases. Mainous and colleagues' [82] purchase arrived with no instructions.

The GAO [73] found that the chemical composition of 4 samples out of 68 was not comparable with the product ordered: in 2 cases a counterfeit version contained less active ingredients, and in 2, a significantly different chemical composition from the product ordered. An analysis of counterfeit dapoxetine (a short-acting selective serotonin-reuptake inhibitor), used against premature ejaculation, found that the tablets contained undisclosed sildenafil [102]. Bate and Hess's [59] analysis showed that 2.5% (3/121) of the tablet sample failed Raman spectrometry. Westenberger et al [104] found that 2 out of 20 samples failed in terms of dissolution and chromatographic purity; Veronin and Youan in 2004 [103] found differences in simvastatin tablet formulation, obtained from 5 countries via the Internet, using near-infrared spectroscopy chemical imaging methods to assess blend uniformity. In contrast, Veronin et al in 2007 [94] found a quality standard comparable with that of the American manufacturer and the Canadian generic drug product tested. The European Alliance for Access to Safe Medicines study [65] study found that 62% of the products received were counterfeit, substandard, or unapproved medications (68% of these were generic and 32% were branded).

Consumers Purchasing From Online Pharmacies

The most difficult task in connection with online pharmacies is attempting to establish the number of people buying and the volume of money traded. Except for cases when this practice is a legal requirement, there are no official data on the issue and it was not possible to obtain details from the Internet as in the sections above.

The scientific evidence about consumers comprises 2 types of data: population surveys and case studies on the adverse effects of drugs purchased via the Internet.

Population Surveys

Table 3 shows the 15 articles [12,64,75,105-116] dealing with consumer characteristics and based on surveys. Most studies were US based, except for 4, which were conducted in Europe [75,105,106] and South America [107]. The findings were published between 2003 [108] and 2010 [105-107,109-112]. Seven studies investigated the general population [12,75,106,108,109,113,114], while the remainder were about specific groups, described in Table 3 [64,105,107,109,111,112,115,116].

Table 3. Articles about consumers buying from online pharmacies, listed in alphabetical order according to the first author

First author, year of publication	Year(s) of data collection	Country where the study was performed	Population investigated	Study design	Percentage of people buying prescription drugs online
Atkinson, 2009 [113]	2005	US	Sample of general population	HINTS 2005 survey[a]	13% (715/5586) (bought medicines or vitamins)
Baker, 2003 [108]	2001–2002	US	Sample of general population	Internet survey	5% (over 3668 respondents)
Bechara, 2010 [107]	2009	Argentina	Healthy young men	Questionnaires on use of phosphodiesterase type 5 inhibitors	2.9% (2/321)
Cicero, 2008 [64]	2006	US	Prescription drug abusers	Questionnaires	6% (41/685)
Cohen, 2010 [109]	2009	US	Sample of general population aged 18–64 years	NHI Survey 2009[b]	6% of 7192
Fox, 2004 [12]	2004	US	Sample of general population	Telephone interviews	4% (93/2200)
Gordon, 2006 [115]	2003–2004	US	Drug-dependent inpatients	Semistructured interviews	6% (6/100)
Gurau, 2005 [75]	2004	UK	Sample of general population	Semistructured questionnaires	34% (102/300) (people buying or intending to buy online)
Harte, 2010 [110]	2006–2007	US	Male college and university students	Online questionnaires on use of phosphodiesterase type 5 inhibitors	12% (8/77) of users
Inciardi, 2009 [116]	Varies with source	US	Drug abusers, students, street sex workers, and “club culture” population	RADARS System[c], NSDUH[d], Delaware School Study, Miami street studies, and qualitative studies	1%–6%
Inciardi, 2010 [111]	Varies with source	US	Drug abusers, students and young adults	RADARS System[c], NSDUH[d], MTF[e]	0.5%–3%
Mazer, 2010 [112]	2007	US	Sample of emergency department patients	Questionnaires	5.4% (89/1654)
Rajamma, 2009 [114]	Not declared	US	Sample of general population born 1946–64	Online questionnaires to a sample from the consumer panel by Common Knowledge Research Services	Not applicable[f]
Schnetzler, 2010 [105]	2008	UK, Germany, Italy	Sexually active men	Online questionnaires on use of phosphodiesterase type 5 inhibitors	32%[g]
Wiedmann, 2010 [106]	2008	Germany	Sample of general population	Face-to-face interviews	Not applicable[f]

[a] Health Information National Trends Survey by the National Cancer Institute.

[b] National Health Interview Survey.

[c] Researched Abuse Diversion and Addiction-Related Surveillance System.

[d] National Survey of Drug Use and Health.

[e] Monitoring The Future.

[f] Evaluation of online drug shopping attitude related to cognitive characteristics.

[g] 32% of Viagra users obtained the drug from sources outside the health system, including the Internet.

In studies about the general population, the percentage of people buying drugs online was between 4% and 6% in the United States in the studies by Fox [12], Baker et al [108], and Cohen and Stussman [109] (who analyzed the health information technology questions of the National Health Interview Survey). Atkinson et al [113] (who analyzed the data of the US national representative sample of the Health Information National Trends Survey [HINTS] by the National Cancer Institute) found a higher percentage (13%), probably because the purchase of vitamins was included in the estimate. The UK-based survey considered both having bought and the intention to buy, so a third of the sample gave a positive response [75].

As in the general population studies, Mazer et al [112] found that 5.4% patients at an emergency department had bought drugs online.

Both surveys on prescription drug abusers [64,115] reported that 6% of the interviewees had used the Internet to purchase prescription medications for their addiction. The 2 papers by Inciardi et al [111,116] were based on several sources of information investigating different populations, including Internet-savvy high school and college students, chronic drug users, and members of the general population. Inciardi et al [116] tried to reveal the “black box” of drug diversion (the transfer of a prescription drug from a lawful to an unlawful channel of distribution or use) using several sources of information: the Researched Abuse Diversion and Addiction-Related Surveillance (RADARS) System, the National Survey of Drug Use and Health (NSDUH), the Delaware School Survey, several Miami street studies, and 2 qualitative studies. The authors concluded that, although the Internet is indeed a source for prescription drugs, the overwhelming volume of purchases is probably at the wholesale level, since few end users report accessing the Internet for drugs. In the various population groups analyzed in this study, the percentage of people accessing prescription drugs via the Internet ranges from 1% to 6%. The results of the 2010 work by Inciardi et al [111] seem to confirm the findings of the previous paper: based on the RADARS System, the NSDUH, and the Monitoring The Future (MTF) survey, they still found that the Internet is a source of prescription opioid acquisition for 0.5%–3% of the investigated populations.

A US survey on the recreational use of erectile dysfunction medications in undergraduate men showed that 12% of users had bought them on the Internet [110]. Bechara et al [107] in Argentina investigated the recreational use of PDE5Is by healthy young men and found that 2.9% of the interviewees had bought the drug through the Internet. A cross-sectional Europe-based observational study examining the purchasing patterns for the same substances found that 11% of the 11,889 subjects reported current use of PDE5Is; 32% of these reported obtaining their PDE5Is from sources outside the health care system without prior health care personnel interaction (eg, Internet, friends) [105].

Some reports suggest widespread use of online pharmacies in people over 35 years of age [75,113] and in women [109]. Mazer et al [112] found no difference in age between those who bought drugs online and those who did not, and no difference in student status, but patients on multiple medications and those with prescription plans used online pharmacies more frequently. Fox [12] found that the most frequently bought products were drugs for chronic conditions (75%), followed by weight loss and sexual performance substances (25%). The most frequent reasons quoted by interviewees for buying or intending to buy online were convenience and saving money [12,75], followed by information anonymity and choice [75]. Regarding location of online pharmacies, the majority of buyers chose or would have chosen sites based in their own countries or in economically developed countries [12,75]. Regarding risk perception, Gurau’s [75] interviewees reported being worried by a lack of a license on the part of the pharmacy (31%), privacy issues (27%), security of online payment (26%), additional charges, drug quality, and superficial prescription. It is interesting to note that health-related risks (drug quality and prescription requirement) rank last in consumers’ perception. In Fox’s [12] survey 68% agreed that online purchasing makes it too easy to obtain drugs illegally.

Rajamma and Pelton [114] explored the potential impact of consumers’ cognitive characteristics on their decision making as it relates to procuring pharmaceutical products via online retail channels. They sent 350 consumers a self-administered electronic questionnaire. Their analysis showed that male, higher-educated, and higher-income consumers had a greater propensity to procure medication online, whereas insurance status did not have any influence. Wiedmann et al [106] investigated the consumer-perceived values and risks related to online shopping attitude and behavior in an e-pharmacy context; they used their model to identify 4 clusters labeled as “enthusiastic experts” (29%), “risk-averse traditionalists” (24%), “convenience-oriented rationalists” (20%), and “inexperienced opponents” (28%).

Clinical Case Reports

We found 12 published papers on clinical cases related to prescription drugs obtained via the Internet, which are shown in Multimedia Appendix 3. One regarded the Internet purchase and injection of gamma-butyrolactone by an 18-year-old woman that led to admission to a pediatric intensive care unit [117]. Romero et al [118] reported florid withdrawal delirium following the discontinuation of Fioricet (a combination of butalbital, which is a barbiturate, with caffeine and acetaminophen indicated for muscle contraction headaches) that a 37-year-old woman purchased online and self-administered in escalating doses for the treatment of chronic headaches in a context of a history of depression. A 40-year-old woman was treated at the emergency department for severe tetany-like spasms, probably due to the ingestion of haloperidol and bentazepam purchased on the Internet, which were displayed on the online pharmacy under “sleep aids” [119]. Neuberg et al [120] reported a case of life-threatening thyroid hormone abuse in a 56-year-old woman encouraged and enabled by unconventional health advice and nonprescribed medication obtained via the Internet. Carisoprodol (a muscle relaxant) withdrawal after Internet purchase was reported by Eleid et al [121]. Another report described a 43-year-old woman who underwent surgery for brain cancer; after hospital discharge she researched adjunctive treatments of cancer on the Internet and self-initiated a 10-day course of cesium chloride, ending up with an acquired long QT syndrome [122]. A 55-year-old man with squamous cell carcinoma of the maxillary sinus, who declined to undergo surgery, radiation, and chemotherapy, decided to treat his cancer using hydralazine sulfate, obtained online; he died from fatal hepatorenal failure, probably caused by the hydralazine sulfate [123]. A series of case studies was reported by Lineberry and Bostwick [124] regarding a suicide attempt by a 35-year-old man using amitriptyline purchased online unbeknownst to his psychiatrist, who had prescribed him paroxetine, and 3 stories of opiate and painkiller addiction in a 37-year-old man, a 42-year-old man, and a 29-year-old woman facilitated by online purchase. Levesque [125] reported tardive dyskinesia in a 67-year-old man associated with the online purchase of the older antipsychotic drugs he probably received when he requested a tranquilizer. All the cases reported so far occurred in the US, although other evidence shows that the phenomenon exists in Europe as well. A case of prolonged hypogonadotropic hypogonadism caused by anabolic steroids purchased on the Internet by a 34-year-old Italian man was reported by Pirola et al [126]. In the United Kingdom, orlistat-induced subacute liver failure was reported in a 57-year-old woman [127]. Also in the United Kingdom, acute coronary syndrome was diagnosed in a 41-year-old man who had taken Viagra for erectile dysfunction [128].

Discussion

We synthesized the scientific literature on online pharmacies by performing an up-to-date and comprehensive review scanning several scientific and institutional databases, with no publication time limits, focusing on the broader picture of online pharmacies. We thought it was necessary to implement a new review because we found no recent summarizing material in the scientific literature, since the second of the only 2 available reviews with a general approach cited just 14 papers regarding online pharmacy supply and was up-to-date to 2007. This is quite a considerable time in a rapidly changing world such as the Internet. In addition, 63% (48/76) of all the studies we found with original data were published in or after 2005, which seems to indicate that the phenomenon of purchasing medications online is increasing.

The main challenge in conducting this review was that the works we further analyzed and report were fairly difficult to compare owing to the widely differing methods used to select and assess samples, and often the works were written in answer to multiple research questions. In order not to miss the multiple aspects of these works we elaborated different research questions—in our opinion the only way to analyze these data. This could be seen as a limit of the study in terms of coherence, but we prefer to view it as an added value of this review, since we wanted to impress on the reader the complexity of the papers' methods and data, mirroring the complexity of the phenomenon of online pharmacies.

Despite this complexity, we made an effort to identify a common denominator in all the research questions scholars have used in tackling the issue, and we related them back to the broad issue of consumer safety in its multiple variations. Ultimately, the tangible side of this consumer safety framework is clinical reports of health damage caused by drugs purchased online, the last link in the chain, when the feared dangers have already occurred. Consumer studies have tried to estimate the number of exposed people and to identify at-risk groups. The majority of these studies focused on the general population or specific groups, and found that 6% or less of the sample had bought drugs online. As depicted in Figure 3, all the contents that we systematized in our research can be linked to consumer safety, in terms of drug misuse, denial and delay of care, transparency issues, drug accessibility, drug quality, and consumer data protection.

One of the major risks posed by online pharmacies is drug misuse. Prescription requirement and use of online questionnaires can be linked to avoiding the physician and hence to the possible misuse of drugs. All the random samples with no specific limits regarding prescription requirements found that at least some websites sold drugs without a prescription, and that online questionnaires could be used as a substitute for prescriptions. This issue leads to the risk of one of the most feared consequences: the possible disruption of the doctor–patient relationship, which has been widely discussed. In the context of the doctor–patient relationship, drugs purchased online can have acute effects, and even chronic and irreversible ones, and nowadays doctors should always investigate the use of nonprescribed substances in their anamnesis [125]. Also connected to drug misuse is the theme of self-medication: in this context drug information displayed by online pharmacies is supposed to be a tool to help consumers be aware of the risks they are exposed to when taking a specific drug. If drug information is not available, this could minimize risk awareness. Risk awareness can also be minimized by marketing strategies, as the findings have often shown that online pharmacies tend to market their products as if they were any other commodity. Indeed, inflation of drug demand has been suggested as an effect by the papers that have analyzed aspects related to marketing strategies adopted by online pharmacies. The demand for drugs is enhanced not just by advertising economic advantages (the true nature of which appears confused and controversial when comparing works that tackled the issue), showing that there are many triggers causing a person to buy online, not just a cheaper purchase, especially in countries other than the United States where drug prices are often negotiated; the other factors probably involved are confidentiality and willingness to avoid the doctor [27,79]. Whatever the reasons, the phenomenon is likely to increase, in a context in which people are becoming increasingly accustomed to online commerce, which is increasing day by day in terms of sales volumes and the number of people engaging in it. Being more accustomed does not mean always being more aware of risks: an experimental study about risk perception in young US consumers regarding “rogue” online pharmacies showed a worrying inability to see multiple signs of danger and a tendency to be misled by online sellers that use professional design, veil untrustworthy features, and mimic reputable websites [129]. The risk of getting unnecessary drugs is also linked to pressure from marketing strategies.

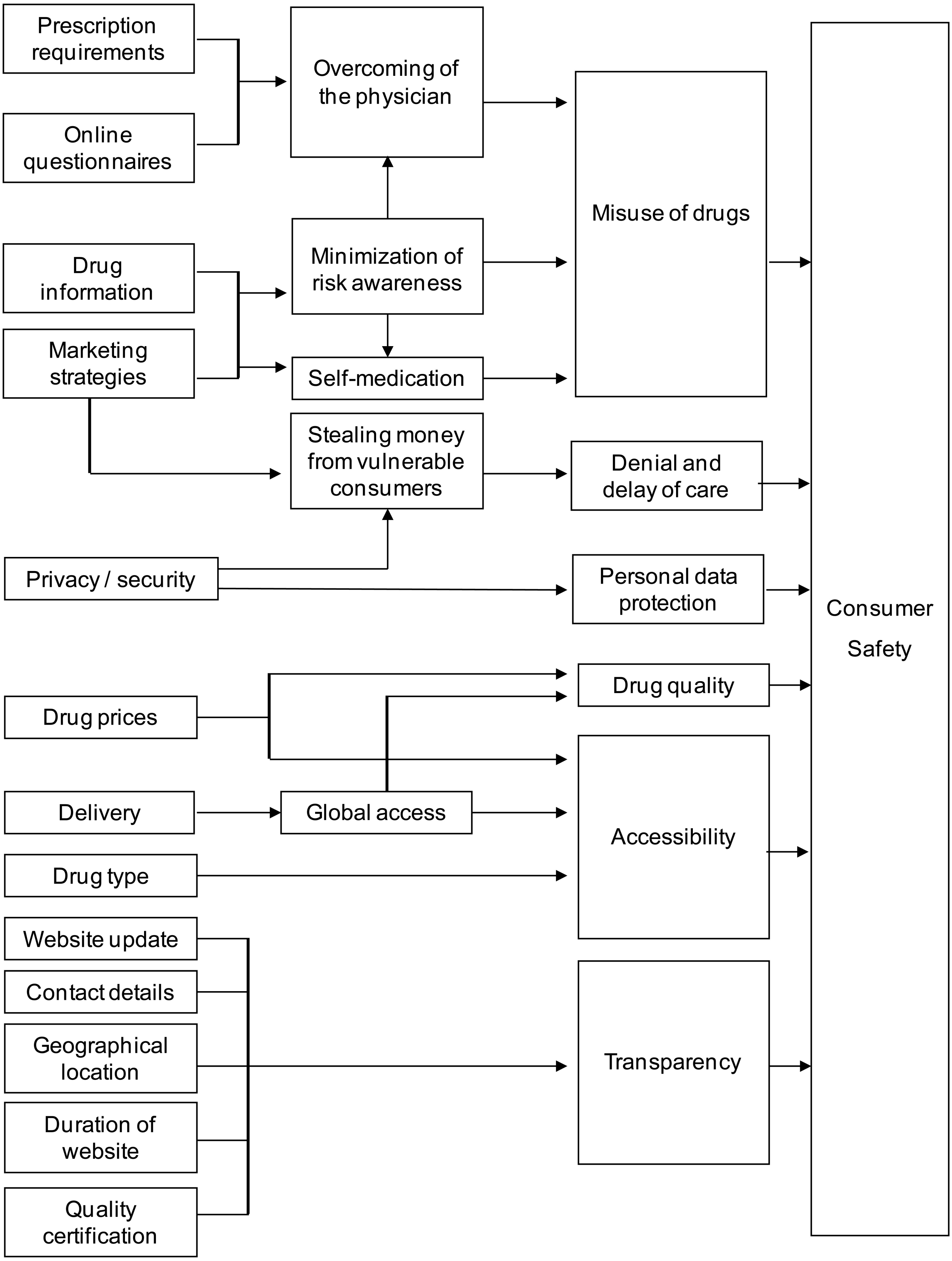
Privacy issues are about the confidentiality of consumers' data and personal data protection. The evidence suggests that online pharmacies (or ones presumed to be such) could be a tool for data fishing or fraud when they do not deliver the products (sometimes charging the consumer anyway) or send something different from what was ordered. It is not just a matter of privacy; it becomes a matter of security. This could even mean that stealing money from vulnerable consumers could lead to them not being able to afford the drug or having to wait to obtain it elsewhere, thus posing problems of denial of care and delay of care.

In terms of drug quality, when the researchers bought drugs online, they very frequently found inappropriate packaging and labeling, whereas the chemical composition was not as expected in a minority of the samples of studies, except for one [65].

Drug accessibility is a core issue regarding online pharmacies. Worldwide delivery eliminates national barriers for consumers; but the place of dispatch can be indicative of the place of production, and therefore it could be linked to drug quality. The analysis of the drugs on offer showed that an online consumer can get virtually anything, which is a matter of risk as well as accessibility, since some drugs are more intrinsically dangerous than others. Lastly, prices can modify drug accessibility and could be linked to drug quality.

Another important area associated with consumer safety is transparency in giving consumers details of the company they are buying from; this aspect can be assessed by analyzing the contact details, geographical location, time websites were accessible, quality certifications, and last website update. Geographical characteristics showed that this information is concealed on many websites, and that US-based websites tend to behave better than others. Studies that investigated the presence of quality certification found it in a minority of the websites.

Figure 3. Consumer safety as a common denominator for studying online pharmacies.



Conclusions

From a policy point of view, online pharmacies are only partially regulated due to intrinsic difficulties linked to the impalpable and evanescent nature of the Web and its global dimension, with no national barriers. The legal implications are really challenging, since the virtual “brave new world” created by the Internet poses issues never faced before. The fragmentary picture of online drug trading regulations is a recognized issue, and a noteworthy attempt to regulate the phenomenon is the “Implementation of the Ryan Haight Online Pharmacy Consumer Protection Act” by the US Department of Justice in 2009 [130-132]. Liang and Mackey [133] proposed a statute that includes drug access costs, prohibition of Internet sales, a legal reform to give several federal agencies the authority to destroy contraband drugs, a pharmacy verification and licensure system, search engine accountability, and prevention of illicit transactions.

Given the technical difficulty of reducing the risks from an enforcement point of view, the role of the consumer becomes essential. The role of the patient as an active partner in health care, and not just a passive object of diagnostic testing and medical treatment, is widely accepted. As this view is accepted, providing information to patients becomes a very crucial issue. Attempts to create some sort of labeling to distinguish trustworthy from rogue websites are valuable, but they cannot be very effective as long as people are not aware of these tools and of the risks involved in buying medication online. As described by Eysenbach [134], medicine 2.0 includes the concept of a shift from an “intermediation environment” (based on the power of the experts) to an “apomediation environment” (based on the empowerment of users). The latter is “desirable by older adolescents and adults, experienced or information literate consumers”, otherwise the risks of increased autonomy can far outweigh the benefits in a context like that of free access to drugs.

In conclusion, online pharmacies are a case where major conflicts occur between the concept of individuals being able to decide their purchases in their own interests on the one hand, and on the other the demand that the state must prevent people from harming themselves and must use public resources fairly and efficiently, as well as the value of social solidarity [23]. In order to enhance the benefits and minimize the risks of online pharmacies, a 2-level approach could be adopted. The first level should focus on policy, with laws regulating the phenomenon at an international level, filling the existing legislative vacuum, although, as stated above, this would be very difficult, costly, and only partially effective. The second level needs to focus on the individual. This approach should aim to increase health literacy, which is the foundation of critical thinking, a skill required for making appropriate health choices, recognizing risks, and making the most of the multitude of opportunities offered by the world of medicine 2.0.

Acknowledgments

The authors wish to thank the anonymous reviewers for their valuable comments and insightful suggestions to improve the quality of the paper.

Authors’ Contributions

GO participated in the conception and the design of the study, checked the collected data, analyzed and interpreted the data, and drafted the article; AM collected and assembled the data; PJS participated in the conception and the design of the study and made a critical revision of the article; UG participated in the conception and the design of the study and collection and interpretation of data, and continuously revised the article during drafting. All authors approved the final article.

Conflicts of Interest

None declared

Multimedia Appendix 1

List of articles about online pharmacies without original data.

[PDF File (Adobe PDF File), 68KB - jmir_v13i3e74_app1.pdf]

Multimedia Appendix 2

List of articles about online pharmacies, sorted in alphabetical order according to first author; the characteristics of each study’s sample is shown in terms of year of data collection, selection method, inclusion criteria, and number of online pharmac

[PDF File (Adobe PDF File), 63KB - jmir_v13i3e74_app2.pdf]

Multimedia Appendix 3

Articles about case reports involving medicines bought on an online pharmacy, listed in alphabetical order according to the first author.

[PDF File (Adobe PDF File), 43KB - jmir_v13i3e74_app3.pdf]

References

1. IMDb. The Internet Movie Database. 2011. Memorable Quotes for The Matrix URL: http://www.imdb.com/title/tt0133093/quotes [accessed 2011-05-31] [WebCite Cache ID 5z5tPKkFp]
2. Orizio G, Gelatti U. Public e-health and new scenarios in terms of risks and opportunities: a specific focus on cyberpharmacies. Social Semiotics 2010;20(1):29-41. [doi: 10.1080/10350330903438394]

3. Gallagher JC, Colaizzi JL. Issues in Internet pharmacy practice. Ann Pharmacother 2000 Dec;34(12):1483-1485. [Medline: 11144707]
4. US Food and Drug Administration. US Department of Health and Human Services. 2010. Buying Medicines Over the Internet URL: http://www.fda.gov/Drugs/ResourcesForYou/Consumers/BuyingUsingMedicineSafely/BuyingMedicinesOvertheInternet/default.htm [accessed 2011-01-05] [WebCite Cache ID 5vVm1xu37]
5. World Health Organization. ATLAS eHealth Country Profiles: Based on the Findings of the Second Global Survey on eHealth. Global Observatory for eHealth series, volume 1. Geneva: WHO; 2010.
6. World Health Organization. WHO. 2010 Jan. Medicines: Counterfeit Medicines. Fact sheet no 275 URL: http://www.who.int/mediacentre/factsheets/fs275/en/ [accessed 2011-01-05] [WebCite Cache ID 5vVnfDxhn]
7. World Health Organization. International Medical Products Anti-Counterfeiting Taskforce — IMPACT. 2009. URL: http://www.who.int/impact/en/ [accessed 2011-08-23] [WebCite Cache ID 5wFRo91BL]
8. Manchikanti L. Prescription drug abuse: what is being done to address this new drug epidemic? Testimony before the Subcommittee on Criminal Justice, Drug Policy and Human Resources. Pain Physician 2006 Oct;9(4):287-321 [FREE Full text] [Medline: 17066115]
9. MarkMonitor Press Release. MarkMonitor. 2011. MarkMonitor Finds Online Drug Brand Abuse is Growing URL: http://www.markmonitor.com/pressreleases/2009/pr090928-bji.php [accessed 2011-01-05] [WebCite Cache ID 5vVnrlWST]
10. National Association of Boards of Pharmacy. Progress report for State and Federal Regulators: April. 2010. Internet drug outlet identification program URL: http://www.nabp.net/news/assets/InternetReport1-11.pdf [accessed 2011-02-04] [WebCite Cache ID 5wFSXLwH3]
11. Fox S. Online Health Search. Washington, DC: Pew Internet & American Life Project; 2007. URL: http://www.pewinternet.org/topics/Health.aspx?start=161&x=x%252cx%252cx%252cx&x=x#ListContinue [WebCite Cache ID 5vVrjWRQ8]
12. Fox S. Prescription drugs online. Washington, DC: Pew Internet & American Life Project; 2004 Oct 10. URL: http://www.pewinternet.org/~/media//Files/Reports/2004/PIP_Prescription_Drugs_Online.pdf.pdf [accessed 2011-08-23] [WebCite Cache ID 619YkxnaM]
13. Bernath P. Regulation of online pharmacy: an Australian perspective. J Law Med 2003 Feb;10(3):339-363. [Medline: 12650004]
14. Koong KS, Koong LY, Liu LC, Yu M. An examination of selected drug availability at online pharmacies. Int J Electron Healthc 2005;1(3):291-302. [Medline: 18048211]
15. Skinner BJ. The Fraser Institute. Price Controls, Patients, and Cross-Border Internet Pharmacies. Risks to Canada's Drug Supply and International Trading Relations URL: http://www.fraserinstitute.org/research-news/display.aspx?id=13315 [accessed 2011-01-13] [WebCite Cache ID 5vhxn9v9e]
16. Castronova JR. Operation cyber chase and other agency efforts to control internet drug trafficking. The “virtual” enforcement initiative is virtually useless. J Leg Med 2006 Jun;27(2):207-224. [doi: 10.1080/01947640600716440] [Medline: 16728353]
17. Adams JU. An Op-Ed concerning steroids and the law: how the Internet has changed illegal drug trade and its prosecution. Physiol Behav 2010 Jun 1;100(3):205-207. [doi: 10.1016/j.physbeh.2010.01.014] [Medline: 20093133]
18. Charatan F. Kansas sues internet medicine suppliers. BMJ 1999 Jun 26;318(7200):1720 [FREE Full text] [Medline: 10381702]
19. Bruckel K, Capozzoli EA. Internet pharmaceutical sales: attributes, concerns, and future forecast. J Hosp Mark Public Relations 2003;15(1):61-76. [Medline: 15683019]
20. Fung CH, Woo HE, Asch SM. Controversies and legal issues of prescribing and dispensing medications using the Internet. Mayo Clin Proc 2004 Feb;79(2):188-194 [FREE Full text] [Medline: 14959914]
21. Henney JE. Cyberpharmacies and the role of the US Food And Drug Administration. J Med Internet Res 2001;3(1):E3 [FREE Full text] [doi: 10.2196/jmir.3.1.e3] [Medline: 11720945]
22. Montoya ID, Jano E. Online pharmacies: safety and regulatory considerations. Int J Health Serv 2007;37(2):279-289. [Medline: 17665724]
23. Nuffield Council on Bioethics. Medical Profiling and Online Medicine: The Ethics of 'Personalised Healthcare' in a Consumer Age. London, UK: Nuffield Council on Bioethics; 2010.
24. Glover-Thomas N, Fanning J. Medicalisation: the role of e-pharmacies in iatrogenic harm. Med Law Rev 2010;18(1):28-55. [doi: 10.1093/medlaw/fwp026] [Medline: 20133321]
25. Wilkinson E. Is the online drugs market putting patients at risk? Nurs Times 2006;102(46):23-24. [Medline: 17144224]
26. Weiss AM. Buying prescription drugs on the internet: promises and pitfalls. Cleve Clin J Med 2006 Mar;73(3):282-288 [FREE Full text] [Medline: 16548451]
27. Mäkinen MM, Rautava PT, Forsström JJ. Do online pharmacies fit European internal markets? Health Policy 2005 May;72(2):245-252. [doi: 10.1016/j.healthpol.2004.09.007] [Medline: 15802158]
28. Nielsen S, Barratt MJ. Prescription drug misuse: is technology friend or foe? Drug Alcohol Rev 2009 Jan;28(1):81-86. [doi: 10.1111/j.1465-3362.2008.00004.x] [Medline: 19320680]
29. Jackson G, Arver S, Banks I, Stecher VJ. Counterfeit phosphodiesterase type 5 inhibitors pose significant safety risks. Int J Clin Pract 2010 Mar;64(4):497-504. [doi: 10.1111/j.1742-1241.2009.02328.x] [Medline: 20088883]

30. Kahan SE, Seftel AD, Resnick MI. Sildenafil and the Internet. J Urol 2000 Mar;163(3):919-923. [Medline: 10688016]
31. Baert B, De Spiegeleer B. Quality analytics of Internet pharmaceuticals. Anal Bioanal Chem 2010 Sep;398(1):125-136. [doi: 10.1007/s00216-010-3912-4] [Medline: 20582403]
32. US National Library of Medicine, National Institutes of Health. PubMed.gov. 2011. PubMed URL: http://www.ncbi.nlm.nih.gov/pubmed [accessed 2011-05-31] [WebCite Cache ID 5z5uhQcZh]
33. Thomson Reuters. Web of Knowledge. 2010. ISI Web of Knowledge URL: http://apps.isiknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=R2iP4@FOaEGobKP9NI5&preferencesSaved= [WebCite Cache ID 5z5ul668q]
34. Elsevier. SciVerse. 2011. Science Direct URL: http://www.sciencedirect.com/ [accessed 2011-08-23] [WebCite Cache ID 619ZHtuXx]
35. American Psychological Association. APA. 2011. PsycInfo URL: http://www.apa.org/pubs/databases/psycinfo/index.aspx [accessed 2011-05-31] [WebCite Cache ID 5z5vIGvkT]
36. Google. 2011. Google Italia URL: http://www.google.it/ [accessed 2011-05-31] [WebCite Cache ID 5z5v4ujKg]
37. Google. Google.com. 2011. Google scholar URL: http://scholar.google.it/ [accessed 2011-05-31] [WebCite Cache ID 5z5vBMqs4]
38. SEO Consultants. Seoconsultants.com. 2010 Aug 28. Top Ten Search Engines URL: http://www.seoconsultants.com/search-engines/ [accessed 2011-06-01] [WebCite Cache ID 5z7O3nhQ6]
39. Checkit. Checkit B.V. 2011. Checkit Nationale Search Engine Monitor URL: http://www.checkit.nl/nationale-search-engine-monitor.aspx [accessed 2011-01-05] [WebCite Cache ID 5vVsRxteS]
40. World Health Organization. WHO. 2011. Homepage URL: http://www.who.int/en/ [accessed 2011-05-31] [WebCite Cache ID 5z5tiHubG]
41. World Health Organization. WHO. 2011. Regional Office for European URL: http://www.euro.who.int/en/home [accessed 2011-05-31] [WebCite Cache ID 5z5vLoxqD]
42. Pan American Health Organization. WHO. 2010. Regional Office of the World Health Organization URL: http://new.paho.org/ [accessed 2011-05-31] [WebCite Cache ID 5z5vQF8dj]
43. US Centers for Diseases Control and Prevention. CDC. 2011 May 31. Health & Safety Topics URL: http://www.cdc.gov/ [accessed 2011-05-31] [WebCite Cache ID 5z5vTEYn9]
44. US Food and Drug Administration. US Department of Health and Human Services. 2011. URL: http://www.fda.gov/ [accessed 2011-05-31] [WebCite Cache ID 5z5vX1gJn]
45. European Centre for Disease Prevention and Control. ECDC network. 2011. Homepage URL: http://www.ecdc.europa.eu/en/Pages/home.aspx [accessed 2011-05-31] [WebCite Cache ID 5z5vZkE1G]
46. European Medicines Agency. European Union. 2011. Homepage URL: http://www.ema.europa.eu/ema/index.jsp?curl=/pages/home/Home_Page.jsp [accessed 2011-05-31] [WebCite Cache ID 5z5vfbUGP]
47. Wikipedia. 2011 Jan 2. PageRank URL: http://en.wikipedia.org/wiki/PageRank [accessed 2011-01-05] [WebCite Cache ID 5vVsKKhav]
48. Brown CW, Goldstein GD, Birkby CS. Auto-Mohs.com. Dermatol Surg 2001 Nov;27(11):975-978. [Medline: 11737136]
49. Logan BK, Goldfogel G, Hamilton R, Kuhlman J. Five deaths resulting from abuse of dextromethorphan sold over the internet. J Anal Toxicol 2009 Mar;33(2):99-103. [Medline: 19239735]
50. Müller D, Weinmann W, Hermanns-Clausen M. Chinese slimming capsules containing sibutramine sold over the Internet: a case series. Dtsch Arztebl Int 2009 Mar;106(13):218-222. [doi: 10.3238/arztebl.2009.0218] [Medline: 19471631]
51. Nguyen MH, Ormiston T, Kurani S, Woo DK. Amphetamine lacing of an Internet-marketed neutraceutical. Mayo Clin Proc 2006 Dec;81(12):1627-1629 [FREE Full text] [Medline: 17165642]
52. Schifano F, Deluca P, Baldacchino A, Peltoniemi T, Scherbaum N, Torrens M, et al. Drugs on the web; the Psychonaut 2002 EU project. Prog Neuropsychopharmacol Biol Psychiatry 2006 Jun;30(4):640-646. [doi: 10.1016/j.pnpbp.2005.11.035] [Medline: 16458404]
53. Schepis TS, Marlowe DB, Forman RF. The availability and portrayal of stimulants over the Internet. J Adolesc Health 2008 May;42(5):458-465. [doi: 10.1016/j.jadohealth.2007.11.140] [Medline: 18407040]
54. Lott JP, Kovarik CL. Availability of oral isotretinoin and terbinafine on the Internet. J Am Acad Dermatol 2010 Jan;62(1):153-154. [doi: 10.1016/j.jaad.2009.04.038] [Medline: 20082897]
55. Yegenoglu S, Sozen B, Aslan D, Calgan Z, Cagirci S. An evaluation of the quality of Turkish community pharmacy web sites concerning HON principles. Telemed J E Health 2008 May;14(4):375-380. [doi: 10.1089/tmj.2007.0065] [Medline: 18570568]
56. Zehnder S, Bruppacher R, Ruppanner H, Hersberger KE. Swiss community pharmacies' on the Web and pharmacists' experiences with E-commerce: longitudinal study and Internet-based questionnaire survey. J Med Internet Res 2004 Mar 3;6(1):e9 [FREE Full text] [doi: 10.2196/jmir.6.1.e9] [Medline: 15111275]
57. Armstrong K, Schwartz JS, Asch DA. Direct sale of sildenafil (Viagra) to consumers over the Internet. N Engl J Med 1999 Oct 28;341(18):1389-1392. [doi: 10.1056/NEJM199910283411810] [Medline: 10536133]

58. Arruñada B. Quality safeguards and regulation of online pharmacies. Health Econ 2004 Apr;13(4):329-344. [doi: 10.1002/hec.827] [Medline: 15067671]
59. Bate R, Hess K. Assessing website pharmacy drug quality: safer than you think? PLoS One 2010;5(8):e12199 [FREE Full text] [doi: 10.1371/journal.pone.0012199] [Medline: 20730049]
60. Bessell TL, Silagy CA, Anderson JN, Hiller JE, Sansom LN. Quality of global e-pharmacies: can we safeguard consumers? Eur J Clin Pharmacol 2002 Dec;58(9):567-572. [doi: 10.1007/s00228-002-0519-5] [Medline: 12483449]
61. Bloom BS, Iannacone RC. Internet availability of prescription pharmaceuticals to the public. Ann Intern Med 1999 Dec 7;131(11):830-833 [FREE Full text] [Medline: 10610627]
62. Bloom BS, Iannacone RC. Changing availability and cost of Internet physician consultations and prescription medications. Med Inform Internet Med 2006 Dec;31(4):247-253. [doi: 10.1080/14639230600551405] [Medline: 17135057]
63. CASA White Paper. The National Center on Addiction and Substance Abuse at Columbia University. 2008 Jul. “You’ve Got Drugs!” V: Prescription Drug Pushers on the Internet URL: http://www.casacolumbia.org/download.aspx?path=/UploadedFiles/2zlhtwsh.pdf [accessed 2011-08-23] [WebCite Cache ID 619aLyQ0c]
64. Cicero TJ, Shores CN, Paradis AG, Ellis MS. Source of drugs for prescription opioid analgesic abusers: a role for the Internet? Pain Med 2008 Sep;9(6):718-723. [doi: 10.1111/j.1526-4637.2007.00323.x] [Medline: 18816331]
65. European Alliance for Access to Safe Medicines. EAASM. 2008. The Counterfeiting Superhighway: The Growing Threat of Online Pharmacies URL: http://www.eaasm.eu/Media_centre/News/The_Counterfeiting_Superhighway [accessed 2011-02-04] [WebCite Cache ID 5wFQBxB6n]
66. Eysenbach G. Online prescribing of sildanefil (Viagra) on the world wide web. J Med Internet Res 1999;1(2):E10 [FREE Full text] [doi: 10.2196/jmir.1.2.e10] [Medline: 11720919]
67. Forman RF. Availability of opioids on the Internet. JAMA 2003 Aug 20;290(7):889 [FREE Full text] [doi: 10.1001/jama.290.7.889] [Medline: 12928464]
68. Forman RF, Woody GE, McLellan T, Lynch KG. The availability of web sites offering to sell opioid medications without prescriptions. Am J Psychiatry 2006 Jul;163(7):1233-1238 [FREE Full text] [doi: 10.1176/appi.ajp.163.7.1233] [Medline: 16816229]
69. Forman RF, Marlowe DB, McLellan AT. The Internet as a source of drugs of abuse. Curr Psychiatry Rep 2006 Oct;8(5):377-382. [Medline: 16968618]
70. Forman RF, Block LG. The marketing of opioid medications without prescription over the Internet. J Public Policy Marketing 2006;25(2):133-146.
71. Gallagher CT, Chapman LE. Classification, location and legitimacy of web-based suppliers of Viagra to the UK. Int J Pharm Pract 2010 Dec;18(6):341-345. [doi: 10.1111/j.2042-7174.2010.00061.x] [Medline: 21054594]
72. US Government Accountability Office. GAO. 2000 Oct 19. Internet Pharmacies: Adding Disclosure Requirements Would Aid State and Federal Oversight URL: http://www.gao.gov/products/GAO-01-69 [accessed 2011-01-11] [WebCite Cache ID 5veywvLvl]
73. US Government Accountability Office. GAO. 2004 Jun 17. Internet Pharmacies: Some Pose Safety Risks for Consumers URL: http://www.gao.gov/products/GAO-04-820 [accessed 2011-01-11] [WebCite Cache ID 5veyhcN8a]
74. Gernburd P, Jadad AR. Will spam overwhelm our defenses? Evaluating offerings for drugs and natural health products. PLoS Med 2007 Sep;4(9):e274 [FREE Full text] [doi: 10.1371/journal.pmed.0040274] [Medline: 17880258]
75. Gurau C. Pharmaceutical marketing on the internet: Marketing techniques and customer profile. J Consumer Marketing 2005;22(7):421. [doi: 10.1108/07363760510631165]
76. Holmes ER, Desselle SP, Nath DM, Markuss JJ. Ask the pharmacist: an analysis of online drug information services. Ann Pharmacother 2005 Apr;39(4):662-667. [doi: 10.1345/aph.1E457] [Medline: 15713791]
77. Kunz MB, Osborne P. A preliminary examination of the readability of consumer pharmaceutical web pages. J Marketing Dev Competitiveness 2010;5(1):33-41.
78. Kuzma J. Web vulnerability study of online pharmacy sites. Inform Health Soc Care 2011 Jan;36(1):20-34. [doi: 10.3109/17538157.2010.520418] [Medline: 21208091]
79. Levaggi R, Orizio G, Domenighini S, Bressanelli M, Schulz PJ, Zani C, et al. Marketing and pricing strategies of online pharmacies. Health Policy 2009 Oct;92(2-3):187-196. [doi: 10.1016/j.healthpol.2009.03.010] [Medline: 19394104]
80. Littlejohn C, Baldacchino A, Schifano F, Deluca P. Internet pharmacies and online prescription drug sales: a cross-sectional study. Drugs (Abingdon Engl) 2005 Feb;12(1):75-80. [doi: 10.1080/0968763042000275326]
81. Mahé E, Saiag P, Aegerter P, Beauchet A. Shopping for psoriasis medications on the Internet. J Eur Acad Dermatol Venereol 2009 Sep;23(9):1050-1055. [doi: 10.1111/j.1468-3083.2009.03248.x] [Medline: 19453812]
82. Mainous AG, Everett CJ, Post RE, Diaz VA, Hueston WJ. Availability of antibiotics for purchase without a prescription on the internet. Ann Fam Med 2009 Oct;7(5):431-435 [FREE Full text] [doi: 10.1370/afm.999] [Medline: 19752471]
83. Memmel LM, Miller L, Gardner J. Over-the-internet availability of hormonal contraceptives regardless of risk factors. Contraception 2006 Apr;73(4):372-375. [doi: 10.1016/j.contraception.2005.10.001] [Medline: 16531170]
84. Orizio G, Schulz P, Domenighini S, Caimi L, Rosati C, Rubinelli S, et al. Cyberdrugs: a cross-sectional study of online pharmacies characteristics. Eur J Public Health 2009 Aug;19(4):375-377 [FREE Full text] [doi: 10.1093/eurpub/ckn146] [Medline: 19151103]

85. Orizio G, Schulz P, Domenighini S, Bressanelli M, Rubinelli S, Caimi L, et al. Online consultations in cyberpharmacies: completeness and patient safety. Telemed J E Health 2009 Dec;15(10):1022-1025. [doi: 10.1089/tmj.2009.0069] [Medline: 19929235]
86. Orizio G, Rubinelli S, Schulz PJ, Domenighini S, Bressanelli M, Caimi L, et al. "Save 30% if you buy today". Online pharmacies and the enhancement of peripheral thinking in consumers. Pharmacoepidemiol Drug Saf 2010 Sep;19(9):970-976. [doi: 10.1002/pds.2007] [Medline: 20652863]
87. Peterson AM. A survey of selected Internet pharmacies in the United States. J Am Pharm Assoc (Wash) 2001 Apr;41(2):205-212. [Medline: 11297333]
88. Peterson AM, Dang B, Fisher BA. A follow-up survey of selected Internet pharmacies in the United States. J Am Pharm Assoc (2003) 2003 Oct;43(5):639-642. [Medline: 14626758]
89. Quon BS, Firszt R, Eisenberg MJ. A comparison of brand-name drug prices between Canadian-based Internet pharmacies and major U.S. drug chain pharmacies. Ann Intern Med 2005 Sep 20;143(6):397-403. [Medline: 16172438]
90. Raine C, Webb DJ, Maxwell SR. The availability of prescription-only analgesics purchased from the internet in the UK. Br J Clin Pharmacol 2009 Feb;67(2):250-254. [doi: 10.1111/j.1365-2125.2008.03343.x] [Medline: 19154446]
91. Schifano F, Deluca P, Baldacchino A. Online availability of dextropropoxyphene over time, 2003-2005. Cyberpsychol Behav 2006 Aug;9(4):404-409. [doi: 10.1089/cpb.2006.9.404] [Medline: 16901243]
92. Soares Gondim AP, Falcão CB. [Evaluation of Brazilian online pharmacies]. Rev Saude Publica 2007 Apr;41(2):297-300 [FREE Full text] [Medline: 17384808]
93. Tsai AC, Lurie P, Sehgal AR. An outbreak of web sites selling ciprofloxacin following an outbreak of anthrax by mail. Am J Med 2002 Oct 1;113(5):424-427. [Medline: 12401538]
94. Veronin MA, Lee E, Lewis EN. "Insight" into drug quality: comparison of simvastatin tablets from the US and Canada obtained via the internet. Ann Pharmacother 2007 Jul;41(7):1111-1115. [doi: 10.1345/aph.1H680] [Medline: 17595307]
95. Wagner ML, Alonso J, Mehlhorn AJ. Comparison of Internet and community pharmacies. Ann Pharmacother 2001 Jan;35(1):116-119. [Medline: 11197571]
96. National Association of Boards of Pharmacy. NABP. 2011. Verified Internet Pharmacy Practice Sites URL: http://www.vipps.nabp.net/ [accessed 2011-05-31] [WebCite Cache ID 5z5vuhNR1]
97. Health on the Net Foundation. HON. 2011 May 26. Homepage URL: http://www.hon.ch/ [accessed 2011-05-31] [WebCite Cache ID 5z5w1ede6]
98. The Joint Commission. Joint Commission on Accreditation of Healthcare Organizations. 2011. Homepage URL: http://www.jointcommission.org/ [accessed 2011-05-31] [WebCite Cache ID 5z5w68sdB]
99. VeriSign Authentication Services. Symantec Corporation. 2011. VeriSign EV SSL URL: http://www.verisign.com/ [accessed 2011-05-31] [WebCite Cache ID 5z5wCmoog]
100. TRUSTe. 2011. Homepage URL: http://www.truste.com/ [accessed 2011-05-31] [WebCite Cache ID 5z5wGWl0w]
101. Miller L, Nielsen C. Internet availability of contraceptives. Obstet Gynecol 2001 Jan;97(1):121-126. [Medline: 11152920]
102. Dean J, Klep R, Aquilina JW. Counterfeit dapoxetine sold on the Internet contains undisclosed sildenafil. Int J Clin Pract 2010 Aug;64(9):1319-1322. [doi: 10.1111/j.1742-1241.2010.02436.x] [Medline: 20497263]
103. Veronin MA, Youan BB. Medicine. Magic bullet gone astray: medications and the Internet. Science 2004 Jul 23;305(5683):481. [doi: 10.1126/science.1097355] [Medline: 15273381]
104. Westenberger BJ, Ellison CD, Fussner AS, Jenney S, Kolinski RE, Lipe TG, et al. Quality assessment of internet pharmaceutical products using traditional and non-traditional analytical techniques. Int J Pharm 2005 Dec 8;306(1-2):56-70. [doi: 10.1016/j.ijpharm.2005.08.027] [Medline: 16266793]
105. Schnetzler G, Banks I, Kirby M, Zou KH, Symonds T. Characteristics, behaviors, and attitudes of men bypassing the healthcare system when obtaining phosphodiesterase type 5 inhibitors. J Sex Med 2010 Mar;7(3):1237-1246. [doi: 10.1111/j.1743-6109.2009.01674.x] [Medline: 20136707]
106. Wiedmann KP, Hennigs N, Pankalla L, Kassubek M, Seegebarth B, Reeh MO. Online distribution of pharmaceuticals: investigating relations of consumers' value perception, online shopping attitudes and behaviour in an e-pharmacy context. J Customer Behav 2010;9(2):175-199. [doi: 10.1362/147539210X511362]
107. Bechara A, Casabé A, De Bonis W, Helien A, Bertolino MV. Recreational use of phosphodiesterase type 5 inhibitors by healthy young men. J Sex Med 2010 Nov;7(11):3736-3742. [doi: 10.1111/j.1743-6109.2010.01965.x] [Medline: 20722788]
108. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406 [FREE Full text] [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
109. Cohen RA, Stussman B. CDC/National Center for Health Statistics. 2009 Dec 23. NCHS Health E-Stat: Health Information Technology Use Among Men and Women Aged 18-64: Early Release of Estimates From the National Health Interview Survey, January-June 2009 URL: http://www.cdc.gov/nchs/data/hestat/healthinfo2009/healthinfo2009.htm [accessed 2011-02-04] [WebCite Cache ID 5wFQjshb4]
110. Harte CB, Meston CM. Recreational use of erectile dysfunction medications in undergraduate men in the United States: characteristics and associated risk factors. Arch Sex Behav 2011 Jun;40(3):597-606. [doi: 10.1007/s10508-010-9619-y] [Medline: 20358273]

111. Inciardi JA, Surratt HL, Cicero TJ, Rosenblum A, Ahwah C, Bailey JE, et al. Prescription drugs purchased through the internet: who are the end users? Drug Alcohol Depend 2010 Jul 1;110(1-2):21-29. [doi: 10.1016/j.drugalcdep.2010.01.015] [Medline: 20227199]
112. Mazer M, Deroos F, Shofer F, Hollander J, McCusker C, Peacock N, et al. Medications from the web: use of online pharmacies by emergency department patients. J Emerg Med 2010 Jul 23. [doi: 10.1016/j.jemermed.2010.05.035] [Medline: 20656438]
113. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
114. Rajamma RK, Pelton LE. An empirical investigation of consumers' procurement of pharmaceutical products via online retail channels. Psychol Marketing 2009;26(10):865-887. [doi: 10.1002/mar.20303]
115. M Gordon S, Forman RF, Siatkowski C. Knowledge and use of the internet as a source of controlled substances. J Subst Abuse Treat 2006 Apr;30(3):271-274. [doi: 10.1016/j.jsat.2005.08.013] [Medline: 16616172]
116. Inciardi JA, Surratt HL, Cicero TJ, Kurtz SP, Martin SS, Parrino MW. The "black box" of prescription drug diversion. J Addict Dis 2009 Oct;28(4):332-347. [doi: 10.1080/10550880903182986] [Medline: 20155603]
117. Winickoff JP, Houck CS, Rothman EL, Bauchner H. Verve and Jolt: deadly new Internet drugs. Pediatrics 2000 Oct;106(4):829-830. [Medline: 11015528]
118. Romero CE, Baron JD, Knox AP, Hinchey JA, Ropper AH. Barbiturate withdrawal following Internet purchase of Fioricet. Arch Neurol 2004 Jul;61(7):1111-1112. [doi: 10.1001/archneur.61.7.1111] [Medline: 15262744]
119. Barnes Recker CK, Hoehns JD, Sutherland JE, Opheim EF. An online misadventure with haloperidol. Ann Pharmacother 2003 Oct;37(10):1531. [doi: 10.1345/aph.1D176] [Medline: 14519026]
120. Neuberg GW, Stephenson KE, Sears DA, McConnell RJ. Internet-enabled thyroid hormone abuse. Ann Intern Med 2009 Jan 6;150(1):60-61. [Medline: 19124828]
121. Eleid MF, Krahn LE, Agrwal N, Goodman BP. Carisoprodol withdrawal after internet purchase. Neurologist 2010 Jul;16(4):262-264. [doi: 10.1097/NRL.0b013e3181aa917e] [Medline: 20592570]
122. Dalal AK, Harding JD, Verdino RJ. Acquired long QT syndrome and monomorphic ventricular tachycardia after alternative treatment with cesium chloride for brain cancer. Mayo Clin Proc 2004 Aug;79(8):1065-1069 [FREE Full text] [Medline: 15301336]
123. Hainer MI, Tsai N, Komura ST, Chiu CL. Fatal hepatorenal failure associated with hydrazine sulfate. Ann Intern Med 2000 Dec 5;133(11):877-880 [FREE Full text] [Medline: 11103057]
124. Lineberry TW, Bostwick JM. Taking the physician out of "physician shopping": a case series of clinical problems associated with Internet purchases of medication. Mayo Clin Proc 2004 Aug;79(8):1031-1034 [FREE Full text] [Medline: 15301331]
125. Levesque CA. Tardive dyskinesia associated with Internet drug purchase. Mayo Clin Proc 2004 Dec;79(12):1587-8; author reply 1588 [FREE Full text] [Medline: 15595348]
126. Pirola I, Cappelli C, Delbarba A, Scalvini T, Agosti B, Assanelli D, et al. Anabolic steroids purchased on the Internet as a cause of prolonged hypogonadotropic hypogonadism. Fertil Steril 2010 Nov;94(6):2331.e1-2331.e3. [doi: 10.1016/j.fertnstert.2010.03.042] [Medline: 20416868]
127. Thurairajah PH, Syn WK, Neil DA, Stell D, Haydon G. Orlistat (Xenical)-induced subacute liver failure. Eur J Gastroenterol Hepatol 2005 Dec;17(12):1437-1438. [Medline: 16292105]
128. Solomon H, Man J, Gill J, Jackson G. Viagra on the internet: unsafe sexual practice. Int J Clin Pract 2002 Jun;56(5):403-404. [Medline: 12137453]
129. Ivanitskaya L, Brookins-Fisher J, O Boyle I, Vibbert D, Erofeev D, Fulton L. Dirt cheap and without prescription: how susceptible are young US consumers to purchasing drugs from rogue internet pharmacies? J Med Internet Res 2010;12(2):e11 [FREE Full text] [doi: 10.2196/jmir.1520] [Medline: 20439253]
130. Seeberg-Elverfeldt NJ. Mail-order trade in medicines in Europe: a guide for legislators to protect consumers. Eur J Health Law 2009 Dec;16(4):351-366. [Medline: 19927423]
131. Catalin J. Internet medicine sales and the need for homogeneous regulation. J Med Marketing 2004;4:342-349. [doi: 10.1057/palgrave.jmm.5040185]
132. US Department of Justice, Drug Enforcement Administration. Federal Register. 2009. Implementation of the Ryan Haight Online Pharmacy Consumer Protection Act of 2008; Final Rule URL: http://www.deadiversion.usdoj.gov/fed_regs/rules/2009/fr0406.pdf [accessed 2011-08-23] [WebCite Cache ID 619bcUHD1]
133. Liang BA, Mackey T. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(1):125-184. [Medline: 19534258]
134. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]

Abbreviations

CASA: National Center on Addiction and Substance Abuse
FDA: Food and Drug Administration

GAO: Government Accountability Office
HINTS: Health Information National Trends Survey
HON: Health on the Net Foundation
JCAHCO: Joint Commission on Accreditation of Healthcare Organizations
MTF: Monitoring The Future
NABP: National Association of Boards of Pharmacy
NSDUG: National Survey of Drug Use and Health
PDE5Is: phosphodiesterase type 5 inhibitors
RADARS System: Researched Abuse Diversion and Addiction-Related Surveillance System
VIPPS: Verified Internet Pharmacy Practice Sites
WHO: World Health Organization

Edited by G Eysenbach; submitted 08.03.11; peer-reviewed by L Ivanitskaya, S Rubinelli; comments to author 23.04.11; accepted 06.06.11; published 30.09.11.

Please cite as:
Orizio G, Merla A, Schulz PJ, Gelatti U
Quality of Online Pharmacies and Websites Selling Prescription Drugs: A Systematic Review
J Med Internet Res 2011;13(3):e74
URL: http://www.jmir.org/2011/3/e74/
doi:10.2196/jmir.1795
PMID:21965220

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