Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 15 (2013), Issue 3 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Original Paper

Storing and Using Health Data in a Virtual Private Cloud

Nathan Regola[1], BS, MS, PhD; Nitesh V Chawla[1], BE, MS, PhD

Interdisciplinary Center for Network Science and Applications, Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN, United States

Corresponding Author:
Nitesh V Chawla, BE, MS, PhD
Interdisciplinary Center for Network Science and Applications
Department of Computer Science and Engineering
University of Notre Dame
384 Fitzpatrick Hall
Notre Dame, IN, 46556
United States
Phone: 1 574 631 1090
Fax: 1 574 631 9260
Email: nchawla@nd.edu

Abstract

Electronic health records are being adopted at a rapid rate due to increased funding from the US federal government. Health data provide the opportunity to identify possible improvements in health care delivery by applying data mining and statistical methods to the data and will also enable a wide variety of new applications that will be meaningful to patients and medical professionals. Researchers are often granted access to health care data to assist in the data mining process, but HIPAA regulations mandate comprehensive safeguards to protect the data. Often universities (and presumably other research organizations) have an enterprise information technology infrastructure and a research infrastructure. Unfortunately, both of these infrastructures are generally not appropriate for sensitive research data such as HIPAA, as they require special accommodations on the part of the enterprise information technology (or increased security on the part of the research computing environment). Cloud computing, which is a concept that allows organizations to build complex infrastructures on leased resources, is rapidly evolving to the point that it is possible to build sophisticated network architectures with advanced security capabilities. We present a prototype infrastructure in Amazon's Virtual Private Cloud to allow researchers and practitioners to utilize the data in a HIPAA-compliant environment.

(J Med Internet Res 2013;15(3):e63) doi:10.2196/jmir.2076

KEYWORDS

medical informatics; HIPAA

Introduction

Intended Audience

This paper presents an overview of current challenges in the research community as health care data are utilized to explore personalized medicine and other information technology–related advancements. This work attempts to address an appropriate network and systems architecture to support compliance with Health Insurance Portability and Accountability Act (HIPAA) regulations in the United States in a research environment. The intended audience of the article is computer scientists, network administrators, principal investigators, and managers of research programs that utilize, or want to utilize, protected health information (PHI) in their research programs. The reader is expected to be familiar with basic computer science principles and terms, such as subnetwork, firewall, router, virtual local area network (VLAN), access control lists, and event logging. Additionally, the reader should have a basic understanding of computer system configuration and network architectures. Where appropriate, the paper references material that may be helpful for readers to acquire indepth knowledge of the concepts, but we attempt to explain the necessary knowledge to understand the paper.

Background

The growing adoption of electronic health records provides a rich source for data mining in order to identify patterns and trends in health care data. Many communities across the United States have or are forming health information exchanges (HIE). Health information exchanges serve as a common hub for data exchange. These organizations enable health care organizations to transfer data by contacting a common hub, the HIE, instead of maintaining connections with numerous peer organizations.

Health information exchanges can also provide data to researchers. We expect that this trend will grow since universities (and other research organizations) can reach a legal agreement with the HIE for data exchange instead of negotiating with numerous organizations. The fundamental challenge in enabling such data partnerships is to provide a secure environment for data sharing and adhering to HIPAA, as necessary. A rarely discussed but relevant aspect of creating "a secure environment" in the research community is the ability to define the network and systems architecture to suit the relevant requirements. Organizations, such as hospitals and large medical research complexes, likely have architected their entire information technology infrastructure to support protected health information.

However, in general research environments or enterprise environments, the information technology systems are designed to support several stakeholder groups and their reconfiguration is not undertaken lightly (in our experience). A research university often has at least two types of network security zones for workstations and other computing clients: a zone to support financial systems, student information systems, and general business functions and a zone to support research systems that is generally less restrictive in use (and occasionally open to external parties). Often workstations are placed into one zone. Servers and other centralized services are typically placed in a series of private zones within a data center network, with appropriate access to the other zones to permit, for example, employees to access the financial databases and for students and faculty to access files stored on central servers.

A server to support sensitive health data does not fit into this model, as it is neither an enterprise computing system, such as a financial database, nor a research system that should be accessed by internal and external parties. A system to support the data mining of health information should be a private server that is accessible to only a few authorized individuals, with appropriate support to implement auditing and other requirements for HIPAA compliance. Frequently, provisioning new servers in a data center is a slow process due to the necessity to configure new networks and coordinate the physical installation of hardware. Amazon's Virtual Private Cloud (VPC) service offering allows researchers to rapidly provision networks and servers. The features of the Amazon VPC cloud service allow network architects to build sophisticated networks, just as they would in a physical environment. Ultimately, universities and other organizations are likely to set up private clouds that allow researchers to rapidly provision networks and servers that suit a variety of needs. However, the concepts and commentary are largely applicable regardless of whether the researchers utilize a private cloud or a public offering such as Amazon.

Even if the data are anonymous/deidentified, it is important to implement the relevant privacy, access, and security safeguards as a matter of information security best practices. In this paper, we present our experience in setting up Amazon's "Virtual Private Cloud offering" [1], a part of Amazon EC2, for health data exchange. We provided a server with Mirth Connect [2], an open source tool that can connect to health information systems (see Appendix 1). We used Mirth Connect to do a pilot run with a health exchange that provided us anonymized and scrubbed data about patients using the HL7 message format [3,4] over an SSL connection (see Appendix 2). We were able to use Mirth Connect to integrate real-time health data. Obviously, security and privacy were of paramount concern, and we wanted to explore the feasibility of creating a HIPAA compliant environment so that other institutions, researchers, and practitioners could potentially work with protected health information (PHI). Our goal with this paper is to demonstrate the necessary steps to set up a Mirth Connect server (although other health data exchange tools could be used) within the Amazon VPC environment, addressing HIPAA compliance where necessary. We should note that three groups of students from the Healthcare Analytics course at the University of Notre Dame successfully incorporated the Mirth Connect framework into their group projects. The students successfully completed the relevant Human Subjects training as part of the curriculum as well. While the health exchange provided simulated data to the students, we believe the infrastructure allows for anonymized patient data.

HIPAA Requirements

Broadly speaking, HIPAA (1996) requires the US Department of Health and Human Services to adopt national standards for health care transactions and code sets, unique health identifiers, privacy, and security [5]. The transaction and code set standards discuss the content and format for various types of transactions including claim filings, claim status, payment advice, and other types of electronic data transactions that occur between specific types of health providers, health plans, and data clearinghouses. The rules that govern transactions and code sets are maintained by the Centers for Medicare & Medicaid Services [6]. These rules primarily concern software vendors, data clearinghouses, and health plans and apply regardless of the computational environment. This paper assumes that the organization is formatting transactional data according to the standards for data exchange, as these transactions are created within applications and their implementation is independent of the server and network architecture. This work does not discuss the various data exchange formats since the authors are not developing data exchange software or proposing alternative data exchange standards. This work focuses on the privacy and security rule, as their implementation depends on the computational environment that is utilized to host the servers and network. The privacy rule and security rule mandate standards for the protection of data and accompanying monitoring and auditing to ensure that the protections are functioning adequately. Compliance with the privacy rule and security rule became mandatory in 2003 and 2005, respectively, with a 1-year extension for small health plans.

Privacy Rule Overview

The intent of the Privacy Rule was to ensure that patients' privacy was respected as electronic transactions increasingly shared data across many organizations. The Privacy Rule [7] component of HIPAA broadly concerns the use of individually identifiable health information, which is frequently referred to as "protected health information" (PHI). The law is designed to ensure that consumers' PHI is handled appropriately within the health organization and only shared with outside entities

according to the uses permitted by law. The Privacy Rule largely focuses on organizational and legal issues, such as requirements for disclosure accounting, permissible PHI disclosures, and other issues surrounding the privacy and disclosure of PHI. In a research environment, this paper assumes that the organization is entitled to the data and does not need to exchange the data with other entities. This paper does not focus on the organizational and legal requirements surrounding data disclosure to third parties. It is worth noting that the Privacy Rule and Security Rule mandate that an organization appoint a privacy officer and a security officer. Additionally, the organization is required to maintain policies that are consistent with HIPAA (and any state and local laws). The Privacy Rule and Security Rule requirements are applicable regardless of whether an organization maintains their own infrastructure or utilizes cloud computing. However, this paper discusses the more interesting information technology aspects of implementing HIPAA compliance in a cloud computing environment—the challenges of addressing the safeguards in the Security Rule.

Security Rule Overview

The security of the server depends on the physical security of the server and network, the operators and users of the server, and the configuration and management of the applications, operating system, and network. Examining this from a HIPAA compliance standpoint, there are several sections in the Security Rule that we address, including "Risk Analysis and Management", "Administrative Safeguards", "Physical Safeguards", and "Technical Safeguards".

Use Cases

Several use cases serve to highlight the obstacles to utilizing enterprise and research computing environments.

Research Health Data

Assume that a researcher from organization O_1 wants to obtain PHI from organization O_2 for a research project and wants to collaborate on the project with collaborator C_1 and collaborator C_2 at organization O_2 and O_3, respectively. Also assume that the requisite permission for sharing the PHI is available. It is determined by the Information Security department at the researcher's organization that the PHI should be stored in encrypted files on a server with server access granted to the researcher and collaborator C_1 and C_2. Personnel at O_1 could create accounts for O_2 and O_3, but all remote users (through a virtual private network V_1) at O_1 have network access to all systems at O_1 on the default network N_1. The ideal solution would be to create a new network N_2 and a new remote user group, V_2 and allow only users in V_2, a separate VPN (virtual private network) group, to access N_2. While these changes would be possible at many organizations, they would require the cooperation of personnel in various departments and manual changes, since the ability to create this architecture is not exposed to the researcher. Utilizing a cloud solution, the researcher could set up a network N_2 at Amazon and an accompanying VPN V_2 for the 3 users—the researcher, C_1, and C_2. This setup should be provisioned rapidly and easily deleted when the project was complete.

Health Data Class Projects

Assume that a researcher wants to set up a system to enable a class of 20 students to work in teams of 4 in miniature data mining projects, utilizing electronic health care records (EHR), albeit deidentified. Each group is allowed access to some element of EHR and not to all, and no group is allowed access to the entire EHR. To ensure that groups do not attempt to access other groups' data and code and to ensure that "the principle of least privilege" is implemented, five separate networks, named N_1 through N_5 will be created, and five separate VPNs will be created, named V_1 through V_5. Network N_1 will be accessible only via VPN V_1, and network N_2 will be accessible only via VPN V_2, and so on, since users assigned to network N_2 do not need to access N_1. At the conclusion of the competition, the networks and VPNs will be deleted. Purchasing a physical server for one semester is an expensive proposition (assuming that five virtual machines are utilized so that only one physical server is required). Requesting the provisioning of five networks and five VPN groups, along with the associated configuration, for a semester is likely to be considered a large request, since the system is not a permanent enterprise system. Utilizing cloud computing, this task can be provisioned rapidly and even scripted for repeated use (ie, for the next time the course is offered). The virtual machines are leased for the exact time that they are needed, while the onsite solution requires the upfront purchase of a physical server.

Why Use the Cloud?

While it is possible to build similar architectures in private data centers, the lead time required to support "research" activity in an enterprise data center can be significant, since the enterprise's resources and policies are aimed toward supporting enterprise applications, such as accounting, human resources, etc. Often, research computing environments, such as campus clusters, are configured in a lower security setting (since they are optimized for maximum performance and rapid troubleshooting) and do not support the physical security standards necessary for HIPAA since various individuals may be permitted to access data centers and servers. Research computing data center space is typically optimized for rapid troubleshooting and high performance, with many physically unlocked racks accessible by various individuals. Strong physical security standards would likely dictate that sensitive systems are placed in locked racks, separate from the main data center space, so that physical access to the sensitive systems can be audited. Additionally, research computing environments typically do not allow researchers to provision private networks, configured to their needs, within a few days. From a technical perspective, research computing centers could, and probably will, provide a middleware software solution that will eventually enable researchers to create their own customized networks and virtual servers (a private cloud). However, this type of solution is not currently available at many institutions. After private clouds are available at research institutions, the discussion raised in this work should aid researchers in determining if their organization's solution is appropriate for hosting data regulated by HIPAA.

If a research computing center does not have a private cloud, researchers are forced to acquire physical servers, assuming that appropriate physical security can be arranged. However, the time required to obtain physical servers and coordinate their installation can range from a few days to a month or more (including the time necessary to arrange meetings and plan connectivity). In this situation, researchers are left with the choice of requesting space in the enterprise data center or attempting to increase the security posture of the research computing environment. The authors have typically had to pursue the former option, as increasing the security posture of a research computing environment is a significant investment of resources and would serve only a few users.

Enterprise data centers at research institutions often have many internal networks in order to partition the enterprise wide area network into many private networks for specific purposes. For example, payment processing systems often have their own network. Additionally, internal databases (ie financial databases) and other types of back office servers are often segmented from web servers and public facing systems. Access control lists in routers and firewalls control the traffic between various private networks. The ability to dynamically provision networks could be exposed to researchers as part of a private cloud but would require extensive access control within the middleware to prevent malicious users from interfering with the configuration of mission critical networks and services. Given the purpose of enterprise networks and services, the managers of such systems have strong incentives to not expose administrative interfaces to researchers and other end users to enable a "private cloud" utilizing the same administrative systems. In the future, it is possible that robust administrative systems, commonly referred to as middleware, will be developed that parallel the functionality of the most sophisticated cloud providers. When sophisticated middleware is available and research organizations adopt it, "private clouds" will be available for researchers within their own institutions or regional consortiums, assuming that it is cost effective.

Sophisticated cloud computing middleware is largely proprietary although open source platforms, such as OpenStack, do exist. However, since many organizations are already utilizing enterprise virtualization platforms, such as VMware, the task of enabling private clouds for research is more complicated than it might appear. Several questions that might be considered are:

- Should organizations set up a secondary private cloud platform for research workloads or utilize one platform for enterprise and research workloads?
- If one platform is selected, should this platform be a proprietary platform or an open source platform?
- Does the platform support an extensible interface to enable end users to configure network and virtual machines?
- How many physical servers should be devoted to each platform if two platforms are utilized?
- Will the private cloud platform and data centers be audited to any standards such as SAS 70 Type II or ISO 27001?

Most of these questions appear to be unresolved at many research organizations as the cloud platform software is still evolving and migrating existing systems is an expensive proposition. When platform software matures, private clouds may be constructed at research organizations, or research organizations may elect to utilize regional cloud providers or providers such as Amazon, depending on the cost and workload.

Commercial cloud providers offer an alternative to physical servers in local data centers by leveraging the organization's architectural assistance to aid the design to ensure that it meets organizational policies. Advantageously, the systems are placed in an appropriate environment that is separate from the organization's enterprise IT systems and research computing systems.

Environment

Several major public cloud computing environments exist, such as Rackspace, Amazon's EC2 service, and Verizon. These providers, and many smaller ones, allow customers to run an operating system of their choice and maintain full control of their operating system and network environment. These types of providers can broadly be classified as infrastructure as a service (IaaS) providers [8]. Platforms such as Microsoft Azure, often termed platform as a service (PaaS) [8], provide an application hosting environment, and the system administrators do not have direct control over the operating system or the network. Maintaining full control of the operating system allows system administrators of the leased compute environment to configure a host-based firewall, set up event logging, and perform any other customizations that are typical for their environment and threat model. This paper does *not* attempt to argue that it is impossible to achieve HIPAA compliance with platform as a service providers. We merely note that maintaining full control of an operating system and network allows the system administrators to set up servers in an environment that closely mirrors their internal network and data centers. For example, if a health organization typically uses software encryption to protect entire directories on a server, encryption of specific directories could also be set up in the operating systems hosted by IaaS providers, since the organization maintains full control.

While it may be possible to achieve HIPAA compliance when using PaaS providers, if the providers facilitate appropriate logging capability and server isolation, we believe that it is simpler to demonstrate HIPAA compliance with the traditional model of an organization controlling their own servers and network. Additionally, the legal and regulatory environment is more familiar with the concept of an organization maintaining full control of their operating system and network as opposed to new models where customers merely maintain control of the application layer. Information security guidance typically centers on a "defense in depth" strategy for protecting data, including the network, operating system, and application running on top of the operating system. Given the current regulatory environment, and the similarity between IaaS and organizations' owned private data centers, we chose to use Amazon's EC2 "Virtual Private Cloud" (VPC) as our platform. Amazon's VPC has several important features that are noteworthy from a security perspective:

- Private subnets in Amazon's environment with nonroutable Internet protocol (IP) addresses

- Option to assign public IP addresses to servers
- Ability to create a virtual private network (VPN) connection to Amazon to incorporate servers into the wide area network of an organization
- Combination of public IP addresses and VPN to create multitiered services
- Inbound and outbound stateful packet inspection firewall rules for each server group
- Ability to set up a host-based firewall (if supported in the operating system)
- Network access control lists (ACL) to control incoming and outgoing traffic at the subnet level
- Multiple network interfaces for each server to build sophisticated network architectures

These features, when combined, enable system architects to create flexible networks that can support a range of security requirements. If implemented correctly and accompanied by appropriate plans and monitoring, we think that these features allow a system to address the privacy and security rules mandated by HIPAA for electronic personal health information (e-PHI). Obviously, HIPAA is a complex set of laws, interpretations of the law, and technical implementations of the safeguards and consequently involves sophisticated risk analysis. The authors are not lawyers and advise the reader to seek appropriate legal counsel. The steps necessary to "ensure confidentiality, integrity, and availability of all e-PHI they create, receive, maintain or transmit", part of the HIPAA Security Rule [9] for a computerized physician order entry (CPOE) system with 1000 users, is a significant undertaking. Ensuring the confidentiality, integrity, and availability of a research database for 5 users containing a limited subset of e-PHI data from a HIE is likely to be far simpler than designing and supporting an environment for a 1000-user CPOE. For example, 20 minutes of unscheduled downtime can likely be considered an acceptable risk for a research database. However, 20 minutes of unscheduled downtime for a CPOE could have disastrous results for a large medical center. The design of a scalable, available, and secure system for a CPOE in the cloud to support 1000 users is well beyond the scope of this paper. The intent of this paper is to analyze Amazon's EC2 platform for use as a hosting environment for research datasets and applications [10], potentially containing e-PHI, referencing the US Department of Health and Human Services documents [7,9] where appropriate. Several other providers, such as Rackspace [11], also support private networking and security groups. The concept of VPC was proposed by researchers [12] before Amazon released their "Virtual Private Cloud" as a public product.

Method

Addressing the Security Rule

Risk Analysis and Management

Risk analysis and management affects the tradeoffs that must be made when selecting appropriate safeguards. First, the evaluation of the likelihood and impact of potential risk to e-PHI needs to be performed. For instance, outside parties may only obtain a copy of e-PHI data from the HIE, and the original e-PHI is not at risk for modification. Second, risk remediation safeguards need to be implemented. Third, a risk analysis and management plan should be documented and updated as the environment changes. Risk analysis and management should be familiar to information security departments, and a similar analysis process should be followed, regardless of the location of the virtual server (whether an infrastructure as a service provider such as Amazon is utilized or an organization's private cloud).

Administrative Safeguards

The administrative safeguards follow from the "Risk Analysis and Management" section above. The organization with e-PHI must identify and analyze the risks to e-PHI and implement security measures to mitigate the risks. Additionally, the organization should designate a security official that is responsible for developing and implementing its security policies. An information access management policy should ensure that the Privacy Rule standards are met by enforcing the minimum disclosure necessary to support the task. Additionally, access to e-PHI data should be granted based on the user's or recipient's role. The security rule also requires that an organization must provide appropriate authorization and supervision of workforce members with access to e-PHI, train all members with access to e-PHI (according to the organization's policies and procedures), and apply appropriate sanctions against members who violate the policies. HIPAA mandates that organizations periodically assess the effectiveness of their policies and procedures.

Physical Safeguards

Physical safeguards are composed of two broad items—"Facility Access and Control" and "Workstation and Device Security". "Facility Access and Control" refers to "limiting physical access to its facilities while ensuring that authorized access is allowed."

Infrastructure as a service offerings provide acceptable "Facility Access and Control" for the network (the network that supports the virtual servers) and virtual servers if the organization considers the providers' standards acceptable. Amazon EC2's service has completed a SAS 70 Type II audit, obtained ISO 27001 certification, and PCI level DSS validation as a level 1 service provider.

Our prototype's environment is Amazon's VPC, a configurable infrastructure as a service offering of Amazon's EC2 service. The VPC enables customers to set up a private multitier network architecture, utilizing a subset of instance types available in EC2. The micro and the cluster compute instance types are currently unavailable in the VPC. This feature is not typically available to researchers at most research organizations unless they have a highly configurable cloud platform with extensive security and auditing tools.

Workstation and Device Security is also an important component of the security plan, and its importance cannot be overstated. If users' workstations are in unsecured areas, without appropriate physical security, this presents a risk that has serious consequences. For example, if SSH keypairs are utilized without passphrases (or with weak passphrases) and stored on local file systems without any encryption, a basic attack could retrieve

the private SSH key from a workstation when the user is not present, and the attacker could use the stolen credentials to impersonate the user. Similarly, keyboard logging software could be used to retrieve a user's private key passphrase, or standard SSH password. Once a user's credentials are impersonated, the attacker can gain access to e-PHI data on remote servers, potentially comprising e-PHI data. If user impersonation occurs, the ability to determine who inappropriately accessed e-PHI data is lost. One possible technical solution to reduce the risk of credential theft would be to implement smart cards.

There are many other types of attacks that can be launched from within an organization, and they are beyond the scope of this paper. The risk mitigation strategies are also beyond the scope of this paper. In addition, it would be impossible for the authors to comment on the readers' environments because various organizations maintain different volumes of data, are subject to numerous types of laws (local, state, and obviously the federal HIPAA laws which mandate a minimum federal standard) and contracts, and experience their own threats. However, these examples serve to motivate the importance of workstation and device security.

Cloud computing does not remove the need to provide a trustworthy computing environment for workstations and other devices in use at an organization. It is paramount for access control, auditing, and integrity controls (below) to provide a trustworthy computing environment for user workstations so that user identity can be ascertained. User authentication is vital so that appropriate authorization decisions can be made by the servers containing e-PHI data. The organization's information security and technology departments should provide a trustworthy environment for the organization's devices and assist with the necessary risk and mitigation strategies that are appropriate for their specific environment.

The need for physical security also applies to the cloud administrator, privacy manager, security manager and any other key positions that have administrative control over systems with e-PHI data. Obviously, if the user credentials of the cloud manager are compromised, the trustworthiness of the audit records, e-PHI data, and the virtual servers is suspect. Our lab maintains a private data handling room for use by key personnel so that tampering with "important" workstations requires a higher level of determination and sophistication compared to a workstation in a public area shared by many individuals.

The workstation and device security section also includes requirements for the protection of e-PHI on electronic media. For example, when electronic media are transferred, removed, disposed, or re-used. Cloud-based block storage, such as Amazon's Elastic Block Store (EBS) [13], should be treated the same as a magnetic or solid state disk. Before the volume is "disposed" by deleting it from an Amazon EC2 account, the volume should be sanitized using an approved wiping tool. This sanitization process can be achieved in the cloud by:

1. powering off the server S, but not terminating it
2. detaching the EBS volume, named J, from the original server instance S
3. attaching the EBS volume J as an additional volume to the wiping instance, W
4. mounting the volume, J, from within the operating system of the wiping instance, W
5. wiping the EBS volume, J, from within W using an appropriate tool

Next Steps for Workstation and Device Security

The organization should put in place a robust workstation and physical device security program (or re-evaluate their current program). This program should establish various classes of users and ensure an appropriate environment for each user class. For example, the user class that has administrative oversight of the cloud computing environment should be granted appropriate attention.

Technical Safeguards

Technical safeguards are perhaps the most interesting of the Security Rule safeguards in an infrastructure as a service environment. Technical safeguards in a HIPAA environment consist of four primary categories: access control, audit controls, integrity controls, and transmission security. Access control ensures that only authorized persons are able to access e-PHI data. This is implemented in our environment with SSH2 RSA keypairs to link Unix login names to actual users, as SSH password login can be vulnerable to timing attacks [14]. Users created a keypair with a passphrase to protect the private key and were trained to store the private key on secure storage. Access control was strengthened by implementing host-based firewall rules (this also allows the capture of firewall audit records for audit and monitoring purposes), Amazon Security Group firewall rules (a stateful packet inspection firewall) that are applied to individual servers (see Tables 1 and 2), and network access control rules (nonstateful ACL rules) to control traffic between networks in the virtual private cloud (see Appendices 3-6). Our network configuration is presented in Figure 1. The audit server has its own private network and the corresponding security group is configured according to Table 1. The analytic servers are segmented to their own network, and the security group rules are presented in Table 2. For those interested in accessing SSH and other services over a VPN connection, an IPsec VPN can be constructed to the Amazon Virtual Private Cloud so that SSH and other traffic from an organization is tunneled over an encrypted connection. As a worst case scenario, assume that a vulnerability in an application level protocol is identified (such as SSH), and the password can be guessed based on the pattern of VPN traffic (which is itself based on the application level traffic, such as SSH) and passwords are obtained for this application. The malicious individual would not be able to use these credentials unless they could access the VPN connection to the VPC, bypass Amazon's firewall and network access lists (assuming that all credentials obtained were used to access resources in the virtual private cloud and the passwords were not shared with any other systems external to the virtual private cloud), or collaborate with an internal user.

Table 1. Audit network security group rules.[a]

Item #	Direction	Source or Destination Address	Port/Protocol	Purpose
1	IN	Data Handling Room	22/TCP	SSH traffic for management
2	IN	Data Handling Room	8000/TCP	Splunk SSL connection for management of monitoring software
3	IN	10.0.1.0/24	10514/TCP	Syslog traffic from servers
4	OUT	0.0.0.0/0 (any network)	80/TCP	Retrieve operating system patches
5	OUT	10.0.0.2/32	53/UDP	DNS lookup
6	OUT	10.0.0.1/32	67/UDP	DHCP server in Amazon Virtual Private Cloud
7	OUT	0.0.0.0/0	123/UDP	NTP servers for time synchronization

[a]The Security Group is composed of two rule sets, inbound and outbound. The order of the rules is not important, but they are numbered for convenience in the table. When the direction is "IN", the address field represents a source address. When the direction is "OUT", the address field represents a destination address. Amazon's web-based tools automatically populate the type of address field since each type of security group (IN or OUT) is stateful and automatically has an accompanying rule in the opposite direction to enable the traffic specified in that particular rule. An "OUT" rule in a stateful firewall can only control traffic to a destination.

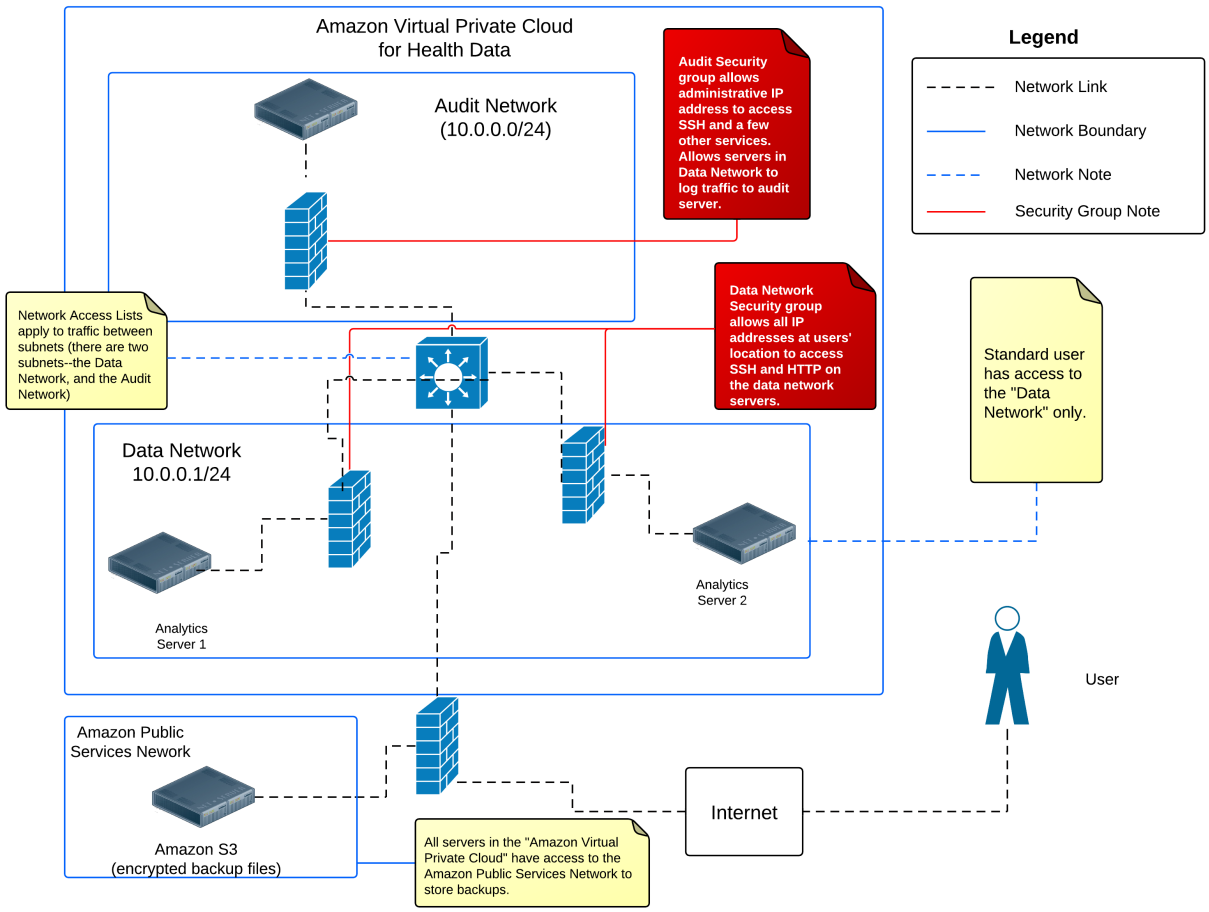
Table 2. Security group for analytics servers.[a]

Item #	Direction	Source or Destination Address	Port/Protocol	Purpose
1	IN	129.74.0.0/16	22/TCP	SSH traffic for users at our campus
2	IN	129.74.0.0/16	443/TCP	SSL access for Apache web server to support web applications
3	OUT	HIE_IP/32	443/TCP	SSL access to retrieve data from HIE
4	OUT	10.0.0.10/32	10514/TCP	Syslog traffic to audit server
5	OUT	10.0.0.2/32	53/UDP	DNS lookup
6	OUT	10.0.0.1/32	67/UDP	DHCP server in Amazon Virtual Private Cloud
7	OUT	0.0.0.0/0	123/UDP	NTP servers for time synchronization

[a]The Security Group is composed of two rule sets, inbound and outbound. The order of the rules is not important, but they are numbered for convenience in the table. When the direction is "IN", the address field represents a source address. When the direction is "OUT", the address field represents a destination address. Amazon's web-based tools automatically populate the type of address field since each type of security group (IN or OUT) is stateful and automatically has an accompanying rule in the opposite direction to enable the traffic specified in that particular rule. An "OUT" rule in a stateful firewall can only control traffic to a destination.

Amazon provides three methods of interacting with their management interface, which is utilized to create and terminate virtual servers and manage networking (including Security Group firewall rules and network access control lists). These methods are currently the Amazon AWS web-based management console, Amazon API Access Keys, and X.509 certificates. Since obtaining access to the API access key or X.509 private certificates would typically only require access to the cloud manager's workstation (unless smartcards are utilized to store an X.509 certificate) , we elected to disable both of these access methods and exclusively utilize the AWS web-based client. The web-based client permits the use of multifactor authentication, which requires two pieces of information to authenticate the cloud manager—the cloud manager's Amazon password (which only he/she knows) and the one-time password token from the cloud manager's token device carried on a keychain. This approach is superior to one piece of information to prove identity and is likely warranted for the cloud manager and other personnel with high-level access to the configuration of the VPC. Impersonating the cloud manager would grant the attacker the ability to terminate the servers or at a minimum compromise their integrity by adjusting firewall rules, terminating the audit server, or perform other catastrophic acts. It is worth noting that it is possible to configure access lists to restrict API access to specific IP address ranges and times of day, complicating an attack using stolen API keys or X.509 certificates. The IP address restrictions could be used in combination with multifactor authentication to limit access for privileged accounts. However, this complicates failure modeling. For example, if a server is compromised and the cloud manager is on vacation and the company VPN is experiencing a failure, then the cloud manager would be unable to login to Amazon and power down a server or adjust the VPC configuration (assuming that the API calls are only accessible from within the organization).

Figure 1. Our prototype environment for HIPAA data, utilizing two subnets, the “Data Network” for servers, and the “Audit Network” for an audit and monitoring server.



Next Steps for Authentication and Authorization

Amazon provides many options for authentication to the EC2 management console, such as the Identity Access and Management services. These services allow organizations to utilize their own authentication systems to grant access to Amazon’s EC2 resources through identity federation for EC2 managers. The organization should determine the roles of the individuals that require access to the Amazon EC2 management console and the servers themselves. It is likely that organizations can readily utilize their own authentication and authorization systems for server and database login if they configure the private cloud at Amazon as an extension of their own network. Integrating internal user accounts with the EC2 management interface will likely require custom software development and may be unnecessary if only a limited number of users require access to the management console for VPC management.

Audit controls are processes and systems that collect access and activity information from systems that contain e-PHI. Our prototype implemented a central syslog server that was devoted to collecting and monitoring events from the project servers. SSH logins and firewall activity (from the host-based firewall) were logged to the syslog server. Our prototype system, containing deidentified health information, logged audit events according to Table 3. Splunk [15], a log management and monitoring tool, was installed on the logging server to monitor the audit event records. Databases that store e-PHI could potentially incorporate logging capability through stored procedures or utilize the database engine’s native logging capability (if it exists) and send these audit messages to the central logging server. For example, if e-PHI is accessed, deleted, or modified, an audit record could be generated and sent to the central logging server. The granularity of auditing should be selected with the consultation of the appropriate risk management and legal personnel.

Table 3. Audited events from servers.

Item #		Source	Event	Notes
1	SSH login	iptables host based firewall	SSH connection containing the source IP address	Traffic is logged before being accepted, ideally capturing any login attempts that cause the SSH daemon to fail.
2	SSH login	SSH daemon	SSH login, including type of authentication, and username	Provides more detail than 1, but occurs after the TCP connection is allowed.
3	Standard Redhat Linux System Events	Various Applications and System Services		
4	HTTPS request to HIE	Iptables host-based firewall		Establish baseline volume of requests and monitor for abnormal behavior

Next Steps for Auditing

Events should be defined according to the legal and business needs of the organization. Once these events are defined, the auditing and alerting software should be configured to alert the appropriate individuals when interesting events are detected.

Software Configuration

The server image was built with Amazon Linux AMI 2011.09 [16], and then we proceeded to install the Mirth Connect Server (using the command line installer), Apache, PHP, and MySQL, leaving these services turned off in chkconfig. The server image is simply a server without user accounts or other personalized information that can be readily cloned to create an arbitrary number of servers. After these tools were installed, we shut down the image machine and created an S3 snapshot of the machine so that we could launch instances based on this image (without a common MySQL password or SSL certificate). After instances were created and started from the image, we configured each instance with a unique MySQL password and Apache SSL certificate and created user accounts.

Next Steps for Software Configuration

The individual or team responsible for creating servers should obtain the necessary software and validate its authenticity. If multiple servers will be utilized for the same task the server can be cloned by creating an S3 image.

Discussion

This paper presents a tutorial on how to duplicate our environment in Amazon's EC2 "Virtual Private Cloud" to obtain data from a HIE. We believe this environment can be HIPAA compliant if best practices are followed as Amazon provides many features (not necessarily configured by default) that can be set up to suit a variety of security requirements. Information security references [17] have discussed HIPAA compliance for some time, and their methodologies can also be applied to new environments, such as the cloud.

Amazon's documentation, to the best of our knowledge, does not clarify how the private networks are implemented. For instance, are they implemented using 802.1Q trunks, a proprietary overlay network, or something else entirely? The choice of an implementation presumably included the analysis of tradeoffs between performance, security, scalability, cost, and potentially other metrics of interest to Amazon. This tradeoff decision is of interest to system administrators so that they can determine if Amazon's security model is an appropriate fit for their project.

Several improvements could be made to the VPC service related to the interaction of "Security Groups" and the services that are utilized by the Amazon Linux servers. One noticeable problem that we experienced was the dynamic nature of the update mirrors that are used to patch Amazon Linux. Ideally, there would be a small set of servers, with Amazon IP addresses, provided in a list, so that system administrators could easily implement outgoing traffic filtering in security groups and allow the update server traffic. Currently, since several mirrors are used, the process of identifying the IP addresses that could potentially be used for updates is haphazard and requires manual trial and error on the part of the system administrator, or alternatively the exemption of all traffic destined to port 80 during the update interval. A similar issue exists when attempting to use NTP servers, since public NTP server pools are typically outside of Amazon, and a given DNS hostname in the NTP pool can resolve to many addresses. Overall, the functionality of the "Security Groups" in the VPC is excellent, since it allows for inbound and outbound filtering, and its effectiveness would be enhanced if these issues are corrected.

Another issue that has more serious implications for auditing and contractual compliance, in our opinion, is the inability for the customer to save or otherwise access the events that are denied by the "Security Group" or "Network Access Control Lists". This information could be valuable for information security professionals attempting to determine the volume and/or source of traffic that is targeted at their network. For example, port scans or other basic activities could be a sign of a pending attack or interest in the network by malicious users. While the host-based firewall rules could be used to signal suspicious behavior by internal users, the addition of audit records from the "Security Groups" or "Network ACLs" could corroborate the system level events in any legal action against internal users.

HIPAA does not mandate specific audit checklists for the setup of servers and networks and instead states that the security should be appropriate for the risk. HIPAA guidelines would be more useful to system administrators if additional guidance was

provided regarding minimum standards. For example, what type of data should be audited to satisfy the requirement to "examine access and other activity in information systems that contain or use e-PHI" [9]? If these requirements were more clearly outlined, then it is likely that cloud providers would readily adopt the minimum requirements so that their customers could use their services for HIPAA compliance. The lack of clarity in HIPAA standards was noted by Wafa in a law review article [18].

There are several broad security considerations that should be mentioned when discussing a migration to cloud computing. First, cloud computing providers typically utilize virtualization to provide isolation [19]. If the underlying physical server is used to host computing resources for multiple customers, several types of side-channel attacks are possible [20]. Amazon offers an option to host virtualized compute servers on hardware that is dedicated (not shared) to a given customer's resources for an additional cost. It is also important to note that side-channel vulnerabilities that exist because of the virtualization software (typically termed a hypervisor) will also likely exist in private data centers and private clusters, assuming that the "attacking" server (often shared on the same physical hardware as the target) is able to be accessed by an internal individual that wants something from the target server.

The reliability of the organization's Internet connection is another important concern if the cloud based servers are primarily accessed from the organization. Based on our experience at our campus, our wide area network (and connection to local data centers) is more reliable than our connection to the commodity Internet. The organization should make the appropriate investments in redundant equipment and Internet connections if they decide to leverage cloud computing as a mission critical service. Amazon recently introduced [21] dedicated connections (1-10Gbps) to customers in specific geographic areas, which would decrease the reliance on commodity Internet connections, since the dedicated connection could be connected directly to the organization's wide area network. It is likely that other providers will release similar products or at least attempt to address this issue.

The system architect must determine the risk and cost of the resulting event, should a side-channel attack occur in a large cloud, given the exposure of the system. For example, one might base the risk decision on whether the server is connected to the public Internet or accessible only through a VPN and the number and type of services that the server provides. For example, assuming that someone is able to demonstrate and carry out a side-channel attack to obtain RSA keys for server login, if the system is accessible only from a VPN, the other organization would not be able to easily utilize the stolen keys unless they could also gain access to the organization's network and VPN.

Conclusion

Ultimately, the system architect must attempt to answer the following questions. First, is their computing infrastructure safer at a cloud provider or in their own data center? Second, what is the budget that is allocated to a "more secure" environment, assuming the basic environment is HIPAA compliant? If the organization's private data center frequently has power outages or only maintains simple backup tapes in the same facility as the servers, the system architect may decide that the benefits of cloud computing (demonstrable physical security, scalable and simple backups hosted in multiple locations, up to date patching of the hypervisor code) far outweigh the drawbacks (the remote possibility of sophisticated side-channel attacks). We suspect that the debate on these security issues is just starting and will continue.

In our experience, cloud computing has the capability to enable rapid provisioning of resources that can be configured to suit a variety of security requirements. Amazon's Virtual Private Cloud offers one possibility to quickly provision sensitive data networks for research purposes. We believe that our prototype can be implemented in a HIPAA compliant manner (for research purposes) and that the discussion in this paper will provide a list of suggested improvements and considerations should readers wish to explore the implementation of a mission critical HIPAA workload in the cloud.

Acknowledgments

The authors would like to thank Andrew K. Rider from the University of Notre Dame and Nicholas Rupley from the Michiana Health Information Network for the instructions related to Mirth Connect.

Conflicts of Interest

NR also provides contract consulting services related to the design and implementation of distributed systems. NVC is the founder of Aunalytics, Inc., focused on big data analytics.

Multimedia Appendix 1

Instructions for installing Mirth Connect on Redhat Enterprise Linux 5 or Redhat Enterprise Linux 6.

[PDF File (Adobe PDF File), 79KB - jmir_v15i3e63_app1.pdf]

Multimedia Appendix 2

Instructions for setting up Mirth Connect for an SSL connection to a HIE.

[PDF File (Adobe PDF File), 64KB - jmir_v15i3e63_app2.pdf]

Multimedia Appendix 3

Amazon Network Access Control List for incoming traffic to data network.

[PDF File (Adobe PDF File), 73KB - jmir_v15i3e63_app3.pdf]

Multimedia Appendix 4

Amazon Network Access Control List for outgoing traffic leaving data network.

[PDF File (Adobe PDF File), 76KB - jmir_v15i3e63_app4.pdf]

Multimedia Appendix 5

Amazon Network Access Control List for incoming traffic to the audit network.

[PDF File (Adobe PDF File), 72KB - jmir_v15i3e63_app5.pdf]

Multimedia Appendix 6

Amazon Network Access Control List for outgoing traffic leaving audit network.

[PDF File (Adobe PDF File), 73KB - jmir_v15i3e63_app6.pdf]

References

1. Amazon Web Services LLC. Online. 2011 Jul 15. User Guide (API Version 2011-07-15) URL: http://docs.amazonwebservices.com/AmazonVPC/2011-07-15/UserGuide/ [accessed 2011-12-05] [WebCite Cache ID 63iAjjWZB]
2. Mirth Corporation. Mirth Connect, The Leading Open Source HL7 Interface Engine. URL: http://www.mirthcorp.com/products/mirth-connect [accessed 2011-12-05] [WebCite Cache ID 63iBSGoC2]
3. iNTERFACEWARE Inc. Components of an HL7 Message. URL: http://www.interfaceware.com/hl7_message_components.html [accessed 2012-01-26] [WebCite Cache ID 64zFu01hw]
4. Eveland J. HL7 Segment : PID : Patient Information. 2005 Jul 26. URL: http://jwenet.net/notebook/1777/1295.html [accessed 2012-01-26] [WebCite Cache ID 64zF9KZY7]
5. US Department of Health & Human Services. HIPAA Administrative Simplification Statute and Rules. URL: http://www.hhs.gov/ocr/privacy/hipaa/administrative/index.html [accessed 2012-12-21] [WebCite Cache ID 6D5mAVS8p]
6. US Centers for Medicare & Medicaid Services. Transaction & Code Sets Standards. 2012 Mar 22. URL: http://www.cms.gov/Regulations-and-Guidance/HIPAA-Administrative-Simplification/TransactionCodeSetsStands/index.html?redirect=/TransactionCodeSetsStands/02_TransactionsandCodeSetsRegulations.asp [accessed 2012-12-21] [WebCite Cache ID 6D5mo19QQ]
7. US Department of Health & Human Services. Summary of the HIPAA Privacy Rule. URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/index.html [accessed 2011-11-22] [WebCite Cache ID 63OO4PSSi]
8. Rimal BP, Eunmi C, Lumb I. A Taxonomy and Survey of Cloud Computing Systems. In: Proceedings of the Fifth International Joint Conference on INC, IMS and IDC. 2009 Presented at: Fifth International Joint Conference on INC, IMS and IDC; Aug. 25-27, 2009; Seoul, Korea p. 41-51. [doi: 10.1109/NCM.2009.218]
9. US Department of Health & Human Services. Summary of the HIPAA Security Rule. URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/srsummary.html [accessed 2011-11-22] [WebCite Cache ID 63ONtnYSX]
10. 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 [FREE Full text] [Medline: 21749966]
11. Rackspace, US Inc.. Managed Private Clouds. URL: http://www.rackspace.com/managed_hosting/private_cloud/ [accessed 2011-12-12] [WebCite Cache ID 63smLXdaI]
12. Wood T, Gerber A, Ramakrishnan KK, Shenoy P, Van der Merwe J. The Case for Enterprise-Ready Virtual Private Clouds. In: Proceedings of the 2009 Conference on Hot Topics in Cloud Computing. 2009 Presented at: 2009 Conference on Hot Topics in Cloud Computing; 2009; Berkeley, CA URL: http://static.usenix.org/event/hotcloud09/tech/full_papers/wood.pdf
13. Amazon Web Services LLC. Amazon Elastic Block Store (EBS). URL: http://aws.amazon.com/ebs/ [accessed 2011-12-05] [WebCite Cache ID 63iGg08AP]
14. Song DX, Wagner D, Tian X. Timing analysis of keystrokes and timing attacks on SSH. In: Proceedings of the 10th USENIX Security Symposium. 2001 Presented at: USENIX Security Symposium; Aug. 13-17, 2001; Washington, DC.

15. Splunk Inc. Operational Intelligence, Log Management, Application Management, Enterprise Security and Compliance. URL: http://www.splunk.com/ [accessed 2012-02-02] [WebCite Cache ID 659e4erRI]
16. Amazon Web Services LLC. Amazon Linux AMI 2011-09 - Release Notes. 2011. URL: http://aws.amazon.com/amazon-linux-ami/latest-release-notes/ [accessed 2011-12-14] [WebCite Cache ID 63vttHqTS]
17. MacLeod D, Geffert B, Deckter D. HIPAA 201: A Framework Approach to HIPAA Security Readiness. In: Tipton H, Krause M, editors. Information Security Management Handbook. Hoboken: Taylor & Francis Ltd; 2004.
18. Wafa T. How the Lack of Prescriptive Technical Granularity in HIPAA Has Compromised Patient Privacy. North Illinois University Law Review 2010;30(3) [FREE Full text]
19. Li A, Yang X, Kandula S, Zhang M. CloudCmp: comparing public cloud providers. In: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement. New York: ACM; 2010 Presented at: 10th ACM SIGCOMM Conference on Internet Measurement; Nov. 1-3, 2010; Melbourne, Australia p. 1-14. [doi: 10.1145/1879141.1879143]
20. Ristenpart T, Tromer E, Shacham H, Savage S. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In: Proceedings of the 16th ACM Conference on Computer and Communications Security.: ACM; 2009 Presented at: 16th ACM Conference on Computer and Communications Security; 2009; Chicago p. 199-212. [doi: 10.1145/1653662.1653687]
21. Amazon Web Services LLC. AWS Direct Connect. URL: http://aws.amazon.com/directconnect/ [accessed 2011-12-14] [WebCite Cache ID 63vuck4UQ]

Abbreviations

ACL: access control list
CPOE: computerized physician order entry
DNS: domain name system
EBS: elastic block service
EHR: electronic health record
e-PHI: electronic personal health information
HIE: health information exchange
HIPAA: Health Insurance Portability and Accountability Act
IaaS: infrastructure as a service
NTP: network time protocol
PaaS: platform as a service
PHI: personal health information
SSH: secure shell
VLAN: virtual local area network
VPC: virtual private cloud
VPN: virtual private network

Edited by G Eysenbach; submitted 06.02.12; peer-reviewed by C Lv; comments to author 25.06.12; revised version received 02.10.12; accepted 19.12.12; published 13.03.13.

Please cite as:
Regola N, Chawla NV
Storing and Using Health Data in a Virtual Private Cloud
J Med Internet Res 2013;15(3):e63
URL: http://www.jmir.org/2013/3/e63/
doi:10.2196/jmir.2076
PMID:23485880

Original Paper

Internet-Based Photoaging Within Australian Pharmacies to Promote Smoking Cessation: Randomized Controlled Trial

Oksana Burford[1*], BPharm; Moyez Jiwa[2*], MA, MD, MRCGP, FRACGP; Owen Carter[3], DPsych; Richard Parsons[4], PhD; Delia Hendrie[5], MA

[1]Curtin Health Innovation Research Institute, School of Pharmacy, Curtin University, Perth, Australia

[2]Curtin Health Innovation Research Institute, Curtin University, Perth, Australia

[3]Office of the Pro-Vice Chancellor (Health Advancement), Edith Cowan University, Perth, Australia

[4]Curtin Health Innovation Research Institute, School of Pharmacy, Occupational Therapy and Social Work, Curtin University, Perth, Australia

[5]Curtin Health Innovation Research Institute, Centre for Population Health Research, Curtin University, Perth, Australia

[*]these authors contributed equally

Corresponding Author:
Oksana Burford, BPharm
Curtin Health Innovation Research Institute
School of Pharmacy
Curtin University
GPO Box U1987
Perth, 6845
Australia
Phone: 61 8 9266 7201
Fax: 61 8 9266 2769
Email: O.Burford@curtin.edu.au

Abstract

Background: Tobacco smoking leads to death or disability and a drain on national resources. The literature suggests that cigarette smoking continues to be a major modifiable risk factor for a variety of diseases and that smokers aged 18-30 years are relatively resistant to antismoking messages due to their widely held belief that they will not be lifelong smokers.

Objective: To conduct a randomized controlled trial (RCT) of a computer-generated photoaging intervention to promote smoking cessation among young adult smokers within a community pharmacy setting.

Methods: A trial was designed with 80% power based on the effect size observed in a published pilot study; 160 subjects were recruited (80 allocated to the control group and 80 to the intervention group) from 8 metropolitan community pharmacies located around Perth city center in Western Australia. All participants received standardized smoking cessation advice. The intervention group participants were also digitally photoaged by using the Internet-based APRIL Face Aging software so they could preview images of themselves as a lifelong smoker and as a nonsmoker. Due to the nature of the intervention, the participants and researcher could not be blinded to the study. The main outcome measure was quit attempts at 6-month follow-up, both self-reported and biochemically validated through testing for carbon monoxide (CO), and nicotine dependence assessed via the Fagerström scale.

Results: At 6-month follow-up, 5 of 80 control group participants (6.3%) suggested they had quit smoking, but only 1 of 80 control group participants (1.3%) consented to, and was confirmed by, CO validation. In the intervention group, 22 of 80 participants (27.5%) reported quitting, with 11 of 80 participants (13.8%) confirmed by CO testing. This difference in biochemically confirmed quit attempts was statistically significant (χ^2_1=9.0, P=.003). A repeated measures analysis suggested the average intervention group smoking dependence score had also significantly dropped compared to control participants (P<.001). These differences remained statistically significant after adjustment for small differences in gender distribution and nicotine dependence between the groups. The mean cost of implementing the intervention was estimated at AU $5.79 per participant. The incremental cost-effectiveness ratio was AU $46 per additional quitter. The mean cost that participants indicated they were willing to pay for the digital aging service was AU $20.25 (SD 15.32).

Conclusions: Demonstrating the detrimental effects on facial physical appearance by using a computer-generated simulation may be both effective and cost-effective at persuading young adult smokers to quit.

Trial Registration: Australian New Zealand Clinical Trials Registry: ACTRN12609000885291; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12609000885291 (Archived by WebCite at http://www.webcitation.org/6F2kMt3kC)

(J Med Internet Res 2013;15(3):e64) doi:10.2196/jmir.2337

KEYWORDS

smoking; tobacco use disorder; skin aging

Introduction

Tobacco smoking leads to premature death or morbidity and places a drain on national resources. Consequently, health professionals and governments stress the importance of smoking cessation and a reduction in exposure to tobacco smoking [1,2].

The younger people are when they start smoking, the greater the risk of illness or death caused by smoking [3]. Approximately half of smokers die prematurely from their habit, with half of these in middle age [4]. Smoking reduces life expectancy by approximately 7 years, with significant morbidity in the final years of a shortened life [4,5]. Even those who smoke between 1 and 4 cigarettes per day triple their long-term risk of dying from cardiovascular disease or lung cancer [6].

Currently in Australia, 19.7% of males and 16.3% of females aged 20 to 29 years smoke on a daily basis [7]. The detrimental long-term health effects of smoking, such as cardiovascular diseases and a variety of cancers, are generally well known in Australia [8]. However, health promotion research shows that, in isolation, knowledge about the hazards of smoking is insufficient to deter smoking behaviors [9]. Young adults who smoke are generally not concerned about the long-term health consequences of smoking because they may believe they will give up the habit while still young [10]. A number of previous studies have investigated the potential of personalized, computer-generated, facial aging software to prompt quit attempts in young adult smokers. These have found facial aging interventions to have some impact [11-14].

The objectives of this randomized controlled trial (RCT) were to test the efficacy and cost-effectiveness of an intervention based on personalized, vivid illustrations of “smoker’s face” among young smokers (18-30 years of age). Smokers face includes wrinkling of the face, gauntness of facial features, and a gray and plethoric complexion. Efficacy was assessed by comparing successful quitting, number of quit attempts, and change in smoking dependence (assessed by the Fagerström score) between the intervention and control groups. The study also aimed to explore the value (feasibility and cost) of an unfunded intervention within pharmaceutical practices.

Methods

Study Design and Population

This study was a RCT (Trial ID number: ACTRN12609000885291) that recruited 160 participants (80 participants assigned to both control and intervention groups) from 8 metropolitan community pharmacies located geographically around Perth city center, Western Australia, when presenting to collect prescribed medications or over-the-counter (OTC) medications.

Eligibility criteria included (1) age range of 18-30 years old (self-report); (2) smoker (defined as smoking 1 or more cigarettes per day via self-report); (3) able to give consent; (4) available for follow-up at 6 months; (5) no beards, mustaches, or non-removable facial accessories; (6) no body dysmorphia (participants screened using the Body Dysmorphic Disorder Questionnaire [BDDQ)]) [15]; and (7) not using nicotine replacement therapy (NRT) or taking oral drugs for nicotine dependence.

Sample Size and Strategy

The sample size of 80 participants per group was calculated to observe a medium effect size (d=0.5), with 80% power and a type I error probability of 5%, and allowing for a 50% attrition rate. The anticipated effect size and attrition rates were based upon the results of a pilot study [16]. At each pharmacy, participants were recruited and assigned by the researcher to the different arms of the study on alternate weeks to minimize contamination between intervention and control participants. The study aimed to recruit 10 participants from each of the 8 pharmacies to each treatment arm (intervention or control). This stratification by pharmacy was performed in an attempt to avoid any bias due to socioeconomic factors.

The Intervention

The APRIL Face Aging software is an Internet-based 3-dimensional age progression software package that creates a stream of aged images of faces from a standard digital photograph (the wrinkling/aging algorithms based upon normative data from people of a broad variety of ages, ethnicities, lifestyle habits, as well as published data regarding facial changes associated with aging). Additionally, the resulting aged images can be adjusted to compare how a person will age as a smoker versus as a nonsmoker (Figures 1 and 2).

Figure 1. Current age photo (25 years) and future digitally aged photo (65 years) of a female current smoker.

Figure 2. Digitally aged photos of female participant at 65 years as a nonsmoker (left) and as a smoker (right).

Data Collection

At recruitment, all participants were asked to complete a baseline questionnaire consisting of demographic data, the Fagerström Smoking Dependence Scale (score from 0-10) [17], questions concerning attitudes toward personal appearance, opinions about health risks associated with smoking, and perceived barriers to quitting smoking. Participants were recruited only if they were not using NRT and not taking oral drugs for nicotine dependence. Participants in both the intervention and control groups received standard 2-minute smoking cessation advice from the pharmacist.

Participants in the intervention group were also screened for body dysmorphia using the BDDQ. In addition, they were photographed and their images were digitally aged as both a smoker and a nonsmoker (using the Internet-based APRIL Face Aging software), and invited to view the age-processed images (Figure 3). They were also asked to complete a questionnaire about their willingness to pay (WTP) for the digital aging service. The digitally aged photograph was sent to their email address within 24 hours of the intervention. Follow-up surveys were undertaken via telephone at 1, 3, and 6 months, each taking approximately 3 minutes to complete.

If participants stated at the 6-month follow-up that they had quit smoking, they were required within 48 hours to undertake a carbon monoxide (CO) breath test to validate their nonsmoking status. The CO monitor utilized was a portable, battery-operated Smokerlyzer (Bedfont Scientific Ltd, Kent, England, UK) that provided a CO level reading in parts per million (ppm).

Figure 3. Intervention being delivered by pharmacist.

Primary Outcomes Measured

The primary outcomes measured were (1) the effect of the intervention by using successful quitting, quit attempts, and progression along the transtheoretical stages of change model, and (2) nicotine dependence using the Fagerström scale. These were measured at baseline and at 1-, 3-, and 6-month follow-ups.

The demographic and baseline smoking habit profiles of the recruited participants were compared between groups using Fisher's exact test and Pearson's chi-square test for categorical variables, and the Student's *t* test for continuous variables. The primary endpoints of the study at the 6-month follow-up were analyzed using chi-square tests to compare percentages of quitters in each group, or *t* tests to compare smoking dependence levels. Percentages of quitters in each group were compared as both self-reported values and as CO-validated values. A logistic regression model was used to analyze the percentage of quitters in the 2 groups after adjustment for any possible differences between groups on the basis of demographic or baseline data. A repeated measures analysis (random effects regression model) was used to identify any changes in the Fagerström dependence score over the entire course of the study using 1- and 3-month follow-up surveys in addition to baseline and 6-month data. Data were analyzed using SAS v9.2 software (SAS Institute, Inc, Cary, NC, USA) with $P<.05$ taken to indicate a statistically significant association.

Secondary Outcomes Measured

The secondary outcomes measured were (1) the cost-effectiveness of the intervention from a health sector perspective in terms of the incremental cost per additional quitter and per additional lifetime quitter, and (2) the business viability of delivering the intervention in a community pharmacy. These were calculated at the conclusion of the study.

Two perspectives were adopted: a health sector perspective and the perspective of a community pharmacy on the assumption that the intervention was not government funded. The direct costs of providing the digital aging service over and above providing standard cessation advice were calculated based on the time taken to provide the service and the cost to a pharmacy of purchasing tokens to use the online software to photoage participants. The cost of a pharmacist's time was valued based on a published recommended rate of pay in Western Australia [18] and tokens were costed based on market price [19]. Time taken that was protocol driven was excluded. Potential cost offsets from a reduction in health care costs of quitters were used to calculate net intervention costs. Cost offsets were based on the Quit Benefits Model, which is a tool developed in Australia to predict the difference in health care costs of smokers and nonsmokers for males and females by age group after 10 years follow-up [20]. This follow-up period was considered long enough to show the beneficial impact of quitting, but short enough to remain within the time frame of policy makers. Cost

offsets were discounted at a rate of 3% as recommended by the US Panel on Cost-effectiveness in Health and Medicine [21]. All costs were expressed in 2011 Australian dollars. The cost of the tokens was converted from American dollars to Australian dollars based on the average exchange rate in 2011 [22]. The number of lifetime quitters was calculated assuming a long-term smoking relapse rate of 37% within 10 years [23]. Smoking relapse after 10 years of abstinence has been found to be less than 1% per year [24].

To assess the robustness of the study results, a scenario sensitivity analysis with the best-case and worst-case scenarios were performed [25]. The parameters varied were the pharmacist's time spent providing the service, the exchange rate for converting the cost of tokens from American dollars to Australian dollars, and the discount rate (Table 1).

In the best-case scenario, the pharmacist's time was adjusted down by 25%, the exchange rate for converting American dollars to Australian dollars was varied to the lowest level in the past 5 years, and a discount rate of 0% was used [22]. In the worst-case scenario, the pharmacist's time was adjusted up by 25%, the exchange rate was varied to the highest level in the past 5 years, and a discount rate of 5% was used. The quantitative data from the customer survey (WTP questionnaire) were analyzed using SPSS v17 software (SPSS Inc, Chicago, IL, USA). Customers' perceptions about the value of the intervention and its impact on loyalty intentions and potential future sales were analyzed by using simple descriptive statistics.

Table 1. Parameter values for base case and sensitivity analyses.

Item[a]	Base case	Scenario sensitivity analysis	
		Best case	Worst case
Pharmacist time per participant to deliver service (mins)	4.8	3.6	6.0
Award wage rate per week for a pharmacist (AU$)	907.40	-	-
Cost of a token (AU$)	3.87	3.63	6.53
Exchange rate (AU$)	0.9687	0.9067	1.6321
Discount rate (%)	3	0	5

[a] Compared to US $1.

Results

Study Design and Population

Customers were screened for eligibility to the RCT from 8 community pharmacies (Figure 4).

Sample Size and Strategy

In total, 1259 customers were screened for eligibility; 213 customers were eligible and 160 were recruited, the others declined the invitation to participate for a range of very different reasons. Eighty participants were recruited to the control group and 80 to the intervention group.

The Intervention

The smoker's face simulations were created using a digital photograph (6.0 megapixels) taken of the intervention participants and uploaded to APRIL Face Aging software version 2.5 on a laptop computer.

Data Collection

The RCT was conducted between January 2010 and December 2010 and all follow-up surveys were completed by June 2011. The final 6-month follow-up showed a response rate of 77.5% for the control group and 72.5% for the intervention group. The demographic and baseline smoking behaviors of recruited participants are shown and compared between groups (intervention versus control) in Table 2.

There were more females and lighter smokers (smoking up to 5 cigarettes per day) in the intervention group; however, there were no statistically significant differences between the control and intervention groups on demographic or smoking dependence variables at baseline. No participants were revealed to have body dysmorphia. A number of questions on the survey were designed to gather the respondents' opinions of self-perceptions and their attitudes toward their smoking behavior. These questions were taken from an earlier survey [11], and showed that the groups were generally well matched. However, a greater proportion of the intervention group appeared to be concerned about their physical appearance (82.5% versus 67.5%, χ^2_1=4.8, *P*=.03), and believed that facial wrinkles were associated with smoking (98.8% versus 85.0%, χ^2_1=10.1, *P*=.002). There was no difference in the proportion of participants in each group who had made at least 1 attempt to quit smoking in the past (68.4% versus 70.9%, χ^2_1=0.1, *P*=.73).

Figure 4. Profile of the randomized controlled trial (using CONSORT guidelines).

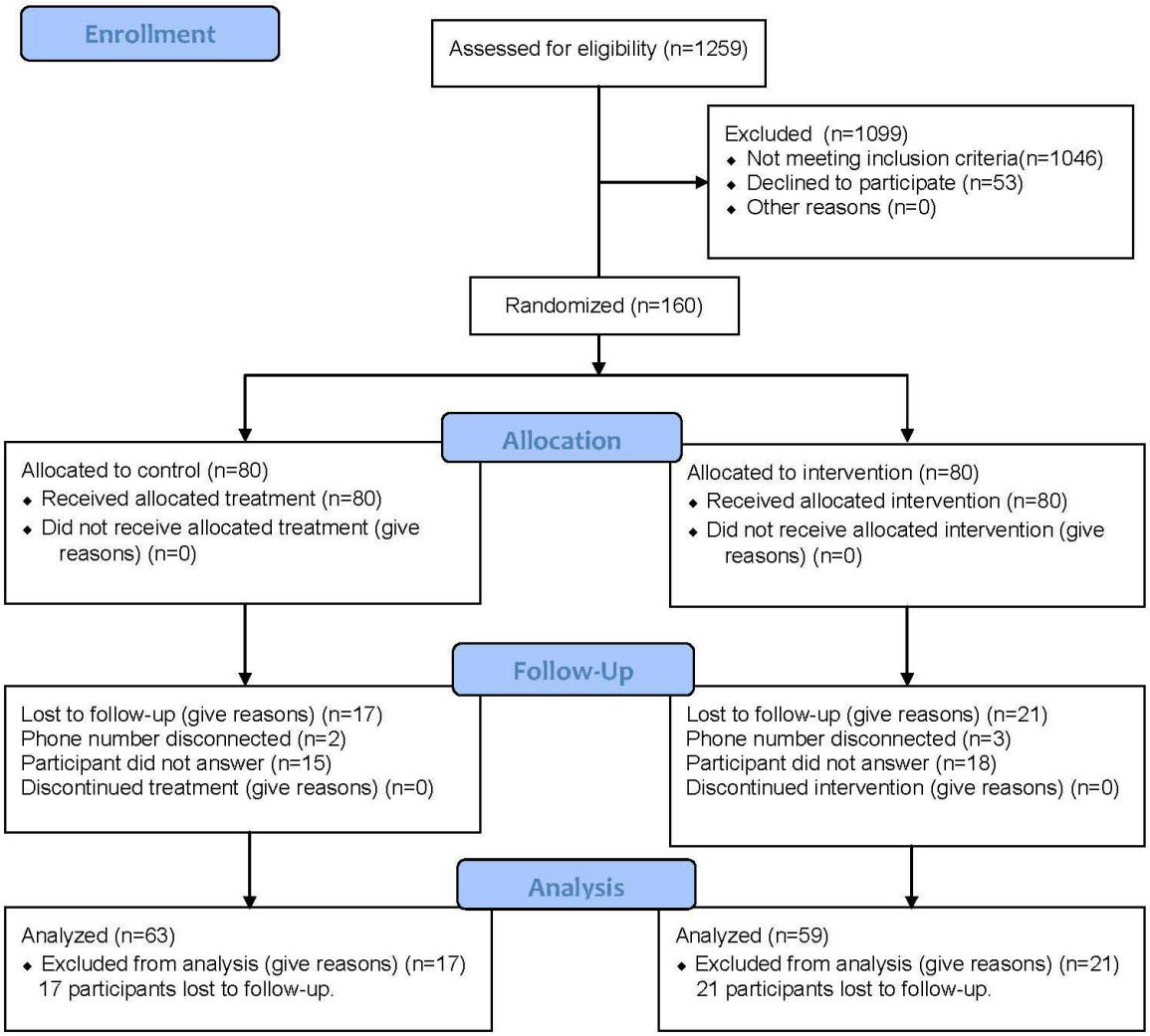


Table 2. Demographic and baseline smoking profile of study participants (N=160).

Variable	Control group (n=80)	Treatment group (n=80)	*P* value[a]
Gender, n (%)			
Male	35 (43.8)	25 (31.3)	.10
Female	45 (56.2)	55 (68.7)	
Age, mean (SD)	25.1 (4.1)	24.2 (4.1)	.16[b]
Education, n (%)			.71
Year 10 high school	15 (19.0)	17 (21.3)	
Year 12 high school	31 (39.2)	29 (36.3)	
Technical and further education (TAFE) qualifications	17 (21.5)	22 (27.5)	
Degree (university/college)	16 (20.3)	12 (15.0)	
Cigarettes per day over past 30 days, n (%)			.35
1	11 (13.8)	19 (23.8)	
2-5	9 (11.3)	10 (12.5)	
6-10	21 (26.3)	14 (17.5)	
11-20	27 (33.8)	29 (36.3)	
>21	12 (15.0)	8 (10.0)	
Fagerström score, mean (SD)	2.96 (2.52)	2.87 (2.48)	.82[b]
Fagerström dependency score, n (%)			.92
0-2	39 (48.8)	39 (49.4)	
3-4	19 (23.8)	18 (22.8)	
5	8 (10.0)	10 (12.7)	
6-7	10 (12.5)	10 (12.7)	
8-10	4 (5.0)	2 (2.5)	

[a] From chi-square test (unless otherwise marked) comparing the treatment groups.

[b] From *t* test.

Primary Outcomes Measured

Table 3 shows the response rates to the follow-up surveys, and the change in smoking behavior over the study duration. There was a significant difference in the proportion of participants self-reporting to have successfully quit smoking by the 6-month survey. Assuming that participants who failed to complete the final follow-up survey continued to smoke, only 1 of 80 control participants (1.3%, 95% CI 0-6.7) were confirmed nonsmokers compared to 11 of 80 participants in the intervention group (13.8%, 95% CI 7.8-22.9). This difference in confirmed quitting is statistically significant (χ^2_1=9.0, *P*=.003). The intervention group contained a larger proportion of participants responding to the question: “I care about how people think I look.” A logistic regression model was used to investigate the association between treatment group and self-reported quitting after adjustment for this difference as well as the small differences between groups in gender and nicotine dependence. The difference remained statistically significant after adjustment for these potentially confounding variables (*P*=.003).

A similar model using confirmed quitting as the dependent variable showed an adjusted *P* value for the treatment group of *P*=.03.

Table 3. Pattern of survey completion and change in smoking behavior at 6-month follow-up (N=160).

Variable	Control group (n=80)	Treatment group (n=80)	*P* value[a]
Response to follow-up questionnaires, n (%)			
All surveys completed	56 (70.0)	48 (60.0)	.38
Incomplete (last survey completed), n (%)			
6 month	6 (7.5)	10 (12.5)	
3 month	8 (10.0)	14 (17.5)	
1 month	3 (3.8)	4 (5.0)	
No follow-up	7 (8.8)	4 (5.0)	
Quit smoking at 6 months, n (%)			
Self-report (questionnaire)	5 (6.3)	22 (27.5)	<.001
Confirmed (CO-validated)	1 (1.3)	11 (13.8)	.003
Change in Fagerström smoking dependence score at 6 months, n (%)			<.001[b]
Reduced dependence	11 (13.8)	41 (51.3)	
No change	68 (85.0)	39 (48.8)	
Increased dependence	1 (1.3)	0	
Change in mean Fagerström score from baseline			<.001[c]
At 1-month follow-up	–0.14	–0.83	
At 3-month follow-up	–0.38	–1.34	
At 6-month follow-up	–0.26	–1.88	

[a] From chi-square tests unless otherwise specified.

[b] From Fisher's Exact test.

[c] Obtained from a repeated measures analysis including all available surveys.

Table 3 also shows changes in the Fagerström smoking dependence score. The 6-month score was grouped into the 5 broad dependence level categories and compared with baseline data. There was a significant difference in change in smoking dependence between groups (χ^2_2=26.2, *P*<.001), with 14% of the control group moving to a lower category compared to 51% of the intervention group doing so.

A random effects regression model was used to model the mean change in Fagerström score from baseline by using data from all follow-up surveys The control group did not experience a significant drop in Fagerström score over the study (*P*=.36), whereas the participants in the intervention group dropped by an average of approximately 1.9 points (*P*=.002). The change in mean scores over the whole study was significantly different between treatment and control groups (*P*<.001).

Although there were no differences between participants at baseline, the regression models were extended to adjust for the gender and age of the participant, and the number of cigarettes smoked at baseline. The models were fitted to the control and intervention group separately because it was clear that changes in score appeared only in the intervention group. For the control group, there were no associations between change in score and age (*P*=.14), gender (*P*=.72), or baseline consumption (*P*=.49). However, for the intervention group, age (*P*<.001) and baseline consumption (*P*<.001) were significantly associated with change in score, whereas gender (*P*=0.34) was not associated. Older participants were less likely to reduce their score than younger participants (*P*=.001), suggesting that the intervention may have a greater effect on the younger participants. Participants who smoked more than 10 cigarettes per day showed a significant drop in score on the Fagerström scale of at least 1 point (*P*<.001) independent of age. Participants smoking 6 to 10 cigarettes per day obtained a lower score, but this change was not statistically significant (*P*=.07), whereas light smokers (0-5 cigarettes per day) showed no change in score.

Secondary Outcomes Measured

Total costs of implementing the intervention from a health sector perspective were AU $463, or the equivalent of AU $5.79 per participant (Table 4).

Table 4. Economic analysis of photoaging service.

Item	Base case	Scenario sensitivity analysis	
		Best case	Worst case
Mean cost per participant of service (AU$)	5.79	5.07	8.93
Total cost of service[a] (AU$)	463	406	714
Incremental cost-effectiveness ratio (ICER)			
Cost per additional quitter (AU$)	46	41	71
Cost per additional lifetime quitter (AU$)	74	64	113
Cost offset from reduction in health care costs (AU$)	2144	2660	1867
Net total cost savings (AU$)	1778	2346	1316
Willingness to pay for service (AU$)			
Mean (SD)	20.25 (15.32)	-	-
Median (IQR)	20.00 (10.00, 20.00)	-	-

[a] For all 80 participants.

With an additional 10 quitters confirmed in the intervention group compared to the control group (11 and 1, respectively), the incremental cost-effectiveness ratio (ICER) was AU $46 per additional quitter, or the equivalent of AU $74 per additional lifetime quitter. Cost offsets of AU $2144 from a reduction in the health care costs of quitters resulted in the intervention potentially generating net total cost savings of AU $1778. In the best-case scenario, the ICER was AU $41 per additional quitter and net total cost savings were AU $2346. Corresponding figures for the worst-case scenario were AU $71 per additional quitter and AU $1316, respectively.

The mean cost that the participants indicated that they were willing to pay for the digital aging service was AU $20.25, which exceeded the mean cost per participant for delivering the service (AU $5.79). The median cost they were willing to pay was AU $20, similar to the mean value. More than 80% of participants said they would be more likely to use the pharmacy to purchase future smoking cessation therapies and to use it more for other purchases. More than 80% of participants also thought their friends would be willing to pay for the service, and all but 2 participants said they would recommend the photoaging intervention to 1 or more friends who were smokers.

Discussion

Summary of Findings

The impact of the photoaging innovation on confirmed quit attempts by the young people recruited to this study was statistically significant (P=.003). The data further demonstrate that the photoaging intervention had a larger influence on younger participants. Also, the participants who did not make a quit attempt, but who smoked more than 10 cigarettes per day, were likely to become less dependent on nicotine.

Strengths and Weaknesses of the Study

The pharmacies selected to take part in the study were chosen to cover a range of socioeconomic areas and the equal number of study participants selected for each treatment group at each pharmacy aimed to diminish any potential biases. However, because of the nature of the intervention, the participants and researcher could not be blinded to the study group. Allocation to groups was not performed as eligible participants were recruited, but according to the treatment being used at the pharmacy during that week. In this setting, there was a substantial risk of contamination between treatment and control groups if participants had been randomized at the point of recruitment rather than by week of attendance at the pharmacy.

The baseline comparisons showed that the 2 groups were very similar on smoking dependence scores, and the 6-month follow-up response rate was high (over 70% for both groups). Follow-up to 12 months may have been preferable, but impractical, in this case. However, follow-up at 6 months was augmented by biochemical verification of tobacco use and cessation [26]. If participants stated they had made a quit attempt at the 6-month conclusion of the study, they were invited to undertake a CO monitor test to validate their nonsmoking status. It was disappointing that so few participants in the control group agreed to CO verification. There are 2 possible reasons for this: it is possible that they continued to smoke, or they were not as engaged in the project as the intervention group and were less amenable to follow-up. Nevertheless the self-reported smoking status data are interesting and although likely to be prone to socially desirable responses, the effect size is still substantial and on a par with other intervention trials.

Although there were more females and light smokers in the intervention group, this was not statistically significant and appeared not to diminish the significant statistical association between treatment group and quitting smoking.

Strengths and Weaknesses in Relation to Other Studies

Although many individualized smoking cessation interventions have been implemented in the past few decades, few have had as marked an impact as reported here. With the advent of digital technology, quit messages can now be delivered by mobile

telephone, email, text messaging, and online social networks [27].

To date, there have been few studies reporting on a personalized photoaging intervention [12-14]; those published have only recruited females, and only 1 of these studies was an RCT that recruited a small number of female smokers who had been referred to a smoking cessation service [14].

Implications for Clinicians and Policy Makers

The economic analysis demonstrates that this personalized smoking cessation intervention is cheap and cost-effective, and it could be readily adopted in community pharmacies. It targets young smokers who are at significant risk of adverse effects of smoking if they continue lifelong smoking.

With an ICER of AU $74 per additional lifetime quitter, the intervention is cost-effective compared with other individualized smoking cessation programs. For example, a systematic review of economic evaluations of a range of smoking cessation interventions found ICERs of between US $260 to US $3263 per lifetime quitter (2002) for counseling or self-help programs versus usual care [28]. These ICERs are of a similar order of magnitude as reported elsewhere for brief advice from a general practitioner to quit smoking and smoking cessation counseling of £196 and £653 per lifetime quitter respectively (1999) [29]. Although these other studies calculated ICERs based on a societal perspective, additional non-health care costs of the photoaging interventions, such as patient time input, are minimal.

Unanswered Questions and Future Research

A review commissioned by the Australian Commonwealth Department of Health and Ageing concluded that interventions delivered by health care providers significantly increased the number and success of quit attempts made in Australia each year [30].

Health care providers, such as pharmacists, are accessible and highly trained [31]. They have an established role in delivering smoking cessation pharmacotherapies and other forms of cessation assistance [32,33]. Could this intervention also be delivered by other health care providers in community settings, such as family medicine or allied health clinics?

A significant development for the use of mass media in delivering antismoking messages is the advent of digital technology. Technologies such as the Internet, social networking sites, and smartphones have the potential to reach large populations of younger people [2,28,34]. Could this Internet technology be delivered to the public without professional facilitation and would it have the same effect?

Further experimental research deploying photoaging technology is needed to answer these questions.

Acknowledgments

Acknowledgement for use of progression software: MJ, OB, and OC thank APRIL Face Aging software for supplying the progression software v2.5 license per gratis for the duration of the study. MJ, OC, and OB conceptualized and designed the study; OB acquired all data; OB, RP, DH, OC, and MJ drafted the manuscript; RP performed all statistical data analyses; DH performed all economic analyses; MJ, OC, RP, DH, and OB critically reviewed the manuscript for intellectual context.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [35].

[PDF File (Adobe PDF File), 988KB - jmir_v15i3e64_app1.pdf]

References

1. World Health Organization. WHO report on the global epidemic, 2011: warning about the dangers of tobacco. URL: http://www.who.int/tobacco/global_report/2011/en/ [accessed 2012-08-09] [WebCite Cache ID 69matV7xr]
2. World Health Organization. 2012 May. WHO Tobacco Fact Sheet No 339 URL: http://www.who.int/mediacentre/factsheets/fs339/en/index.html [accessed 2012-06-15] [WebCite Cache ID 69mbjhQuf]
3. White V, Hayman J. Smoking behaviours of Australian secondary students in 2005. 2006 Jun. URL: http://www.nationaldrugstrategy.gov.au/internet/drugstrategy/publishing.nsf/Content/E1B70590AD4EF56DCA257225000EDCE9/$File/mono59.pdf [accessed 2012-08-09] [WebCite Cache ID 69mcGNbid]
4. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004 Jun 26;328(7455):1519 [FREE Full text] [doi: 10.1136/bmj.38142.554479.AE] [Medline: 15213107]
5. Brønnum-Hansen H, Juel K. Abstention from smoking extends life and compresses morbidity: a population based study of health expectancy among smokers and never smokers in Denmark. Tob Control 2001 Sep;10(3):273-278 [FREE Full text] [Medline: 11544393]
6. Bjartveit K, Tverdal A. Health consequences of smoking 1-4 cigarettes per day. Tob Control 2005 Oct;14(5):315-320 [FREE Full text] [doi: 10.1136/tc.2005.011932] [Medline: 16183982]

7. Australian Institute of Health and Welfare. 2010 National drug strategy household survey report. Canberra: AIHW; 2011 Jul. URL: http://www.aihw.gov.au/publication-detail/?id=32212254712 [accessed 2012-08-09] [WebCite Cache ID 69mcVDdqj]
8. Siahpush M, McNeill A, Hammond D, Fong GT. Socioeconomic and country variations in knowledge of health risks of tobacco smoking and toxic constituents of smoke: results from the 2002 International Tobacco Control (ITC) Four Country Survey. Tob Control 2006 Jun;15 Suppl 3:iii65-iii70 [FREE Full text] [doi: 10.1136/tc.2005.013276] [Medline: 16754949]
9. Droomers M, Schrijvers CT, Mackenbach JP. Educational differences in the intention to stop smoking: explanations based on the Theory of Planned Behaviour. Eur J Public Health 2004 Jun;14(2):194-198 [FREE Full text] [Medline: 15230509]
10. Carter O. Australian tobacco control advertising 1997-2005. In: Audit of Published and Unpublished Formative Research and Campaign Evaluations. Curtin University of Technology: 2006 CBRCC Report 060721; 2005.
11. Hysert PE, Mirand AL, Giovino GA, Cummings KM, Kuo CL. "At Face Value": age progression software provides personalised demonstration of the effects of smoking on appearance. Tob Control 2003 Jun;12(2):238 [FREE Full text] [Medline: 12773741]
12. Weiss C, Hanebuth D, Coda P, Dratva J, Heintz M, Stutz EZ. Aging images as a motivational trigger for smoking cessation in young women. Int J Environ Res Public Health 2010 Sep;7(9):3499-3512 [FREE Full text] [doi: 10.3390/ijerph7093499] [Medline: 20948939]
13. Grogan S, Flett K, Clark-Carter D, Gough B, Davey R, Richardson D, et al. Women smokers' experiences of an age-appearance anti-smoking intervention: a qualitative study. Br J Health Psychol 2011 Nov;16(4):675-689. [doi: 10.1348/2044-8287.002006] [Medline: 21199543]
14. Grogan S, Flett K, Clark-Carter D, Conner M, Davey R, Richardson D, et al. A randomized controlled trial of an appearance-related smoking intervention. Health Psychol 2011 Nov;30(6):805-809. [doi: 10.1037/a0024745] [Medline: 21767016]
15. Phillips KA. The Broken Mirror: Understanding and Treating Body Dysmorphic Disorder. Oxford: Oxford University Press; 2005.
16. Burford O, Smith M, Jiwa M, Carter O. Photoageing intervention (PAINT): a proposal for a randomised controlled trial in Australian primary care. AMJ 2008 Nov 2008;1(7):8-12. [doi: 10.4066/AMJ.2009.108]
17. Quitter's Guide.com. 2008. Fagerstrom Scale: how addicted are you to nicotine? URL: http://www.quittersguide.com/fagerstrom-scale.shtml [accessed 2012-08-09] [WebCite Cache ID 69mcZW6WY]
18. Pharmacy Industry Award 2010 (MA0000012). 2010. URL: http://www.chambernt.com.au/documents/File/Modernised_Awards/Pharmacy_Industry_Award_2010_MA000012.pdf [accessed 2012-08-09] [WebCite Cache ID 69mcdyi6k]
19. APRIL Face Aging Software. 2012. AgeMe URL: http://www.aprilage.com/ageme.html [accessed 2012-08-09] [WebCite Cache ID 69mcgHvzs]
20. Hurley SF, Matthews JP. The Quit Benefits Model: a Markov model for assessing the health benefits and health care cost savings of quitting smoking. Cost Eff Resour Alloc 2007;5:2 [FREE Full text] [doi: 10.1186/1478-7547-5-2] [Medline: 17241477]
21. Gold MR, Siegel JE, Russell LB. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996.
22. OANDA. 2012. Historical Exchange Rates URL: http://www.oanda.com/currency/historical-rates/ [accessed 2012-08-09] [WebCite Cache ID 69mchwPZf]
23. Hawkins J, Hollingworth W, Campbell R. Long-term smoking relapse: a study using the british household panel survey. Nicotine Tob Res 2010 Dec;12(12):1228-1235. [doi: 10.1093/ntr/ntq175] [Medline: 21036960]
24. Krall EA, Garvey AJ, Garcia RI. Smoking relapse after 2 years of abstinence: findings from the VA Normative Aging Study. Nicotine Tob Res 2002 Feb;4(1):95-100. [doi: 10.1080/14622200110098428] [Medline: 11906685]
25. Drummond MF, Sculpher MJ, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2005.
26. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
27. Aveyard P, Raw M. Improving smoking cessation approaches at the individual level. Tob Control 2012 Mar;21(2):252-257. [doi: 10.1136/tobaccocontrol-2011-050348] [Medline: 22345262]
28. Ronckers ET, Groot W, Ament AJ. Systematic review of economic evaluations of smoking cessation: standardizing the cost-effectiveness. Med Decis Making 2005;25(4):437-448. [doi: 10.1177/0272989X05278431] [Medline: 16061896]
29. Woolacott NF, Jones L, Forbes CA, Mather LC, Sowden AJ, Song FJ, et al. The clinical effectiveness and cost-effectiveness of bupropion and nicotine replacement therapy for smoking cessation: a systematic review and economic evaluation. Health Technol Assess 2002;6(16):1-245 [FREE Full text] [Medline: 12269277]
30. Miller M, Wood L. Effectiveness of smoking cessation interventions: review of evidence and implications for best practice in Australian health care settings. Aust NZ J Public Health 2003;27(3):300-309. [Medline: 14705286]
31. The Pharmacy Guild of Australia. 2010 May. The Roadmap: the strategic direction for community pharmacy URL: http://www.guild.org.au/sites/The_Guild/tab-Pharmacy_Services_and_Programs/The_Roadmap/The%20Roadmap.page [accessed 2012-08-09] [WebCite Cache ID 69mclFUY3]

32. Saba M, Saini B. Smoking cessation in pharmacy - Shedding light on the current concepts. Australian Pharmacist 2010;29:1034-1039.
33. Sinclair HK, Bond CM, Stead LF. Community pharmacy personnel interventions for smoking cessation. Cochrane Database Syst Rev 2004(1):CD003698. [doi: 10.1002/14651858.CD003698.pub2] [Medline: 14974031]
34. 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]
35. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BDDQ: Body Dysmorphic Disorder Questionnaire
CO: carbon monoxide
ICER: incremental cost-effectiveness ratio
NRT: nicotine replacement therapy
OTC: over-the-counter
RCT: randomized controlled trial
WTP: willingness to pay

Edited by G Eysenbach; submitted 03.09.12; peer-reviewed by P Hysert; comments to author 26.01.13; accepted 03.02.13; published 26.03.13.

Please cite as:
Burford O, Jiwa M, Carter O, Parsons R, Hendrie D
Internet-Based Photoaging Within Australian Pharmacies to Promote Smoking Cessation: Randomized Controlled Trial
J Med Internet Res 2013;15(3):e64
URL: http://www.jmir.org/2013/3/e64/
doi:10.2196/jmir.2337
PMID:23531984

Original Paper

How User Characteristics Affect Use Patterns in Web-Based Illness Management Support for Patients with Breast and Prostate Cancer

Elin Børøsund[1], MSc; Milada Cvancarova[1], PhD; Mirjam Ekstedt[1,2], PhD; Shirley M Moore[3], PhD; Cornelia M Ruland[1,4], PhD

[1]Centre for Shared Decision Making and Collaborative Care Research, Oslo University Hospital, Oslo, Norway

[2]Royal Institute of Technology, KTH, School of Technology and Health, Stockholm, Sweden

[3]Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, United States

[4]Department of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Elin Børøsund, MSc
Centre for Shared Decision Making and Collaborative Care Research
Oslo University Hospital
P.O. Box 4950 Nydalen
Oslo, 0424
Norway
Phone: 47 23075452
Fax: 47 23075450
Email: elin.borosund@rr-research.no

Abstract

Background: Frequently eHealth applications are not used as intended and they have high attrition rates; therefore, a better understanding of patients' need for support is warranted. Specifically, more research is needed to identify which system components target different patient groups and under what conditions.

Objective: To explore user characteristics associated with the use of different system components of a Web-based illness management support system for cancer patients (WebChoice).

Methods: For this secondary post hoc analysis of a large randomized controlled trial (RCT), in which WebChoice was tested among 325 breast cancer and prostate cancer patients who were followed with repeated measures for 1 year, usage patterns of 162 cancer patients in the intervention arm with access to WebChoice were extracted from the user log. Logistic regression was performed to identify patterns of associations between system use and patient characteristics. Latent class analyses (LCA) were performed to identify associations among the use of different system components and levels of social support, symptom distress, depression, self-efficacy, and health-related quality of life.

Results: Approximately two-thirds (103/162, 63.6%) of the patients logged on to WebChoice more than once, and were defined as users. A high level of computer experience (odds ratio [OR] 3.77, 95% CI 1.20-11.91) and not having other illnesses in addition to cancer (OR 2.10, 95% CI 1.02-4.34) increased the overall probability of using WebChoice. LCA showed that both men with prostate cancer and women with breast cancer who had low scores on social support accompanied with high levels of symptom distress and high levels of depression were more likely to use the e-message component. For men with prostate cancer, these variables were also associated with high use of the self-management advice component. We found important differences between men with prostate cancer and women with breast cancer when associations between WebChoice use and each user characteristic were analyzed separately. High use of all components was associated with low levels of social support among women with breast cancer, but not among men with prostate cancer. High use of e-messages, advice, and the discussion forum were associated with high levels of depression among women with breast cancer, but not among men with prostate cancer. For men with prostate cancer (but not women with breast cancer), high use of symptom assessments, advice, and the discussion forum were associated with high levels of symptom distress. However, it is unclear whether these findings can be attributed to differences related to diagnosis, gender, or both.

Conclusions: This study provides evidence that different user characteristics are associated with different use patterns. Such information is crucial to target Web-based support systems to different patient groups. LCA is a useful technique to identify

subgroups of users. In our study, e-messages and self-management advice were highly used components for patients who had low levels of social support and high illness burden, suggesting that patients with these characteristics may find such tools particularly useful.

Trial Registration: ClinicalTrials.gov NCT00710658; http://clinicaltrials.gov/ct2/show/NCT00710658 (Archived by WebCite at http://www.webcitation.org/6EmEWZiwz)

(J Med Internet Res 2013;15(3):e34) doi:10.2196/jmir.2285

KEYWORDS

Web-based intervention; Internet; symptom management; self-care; use patterns; user characteristics; targeting

Introduction

Web technologies provide an opportunity to target support to different patient groups and can reach a large number of users. Recent years have seen growth in Web-based interventions and support systems that have been shown to assist a wide range of patients successfully [1-6], including people with cancer [7-9], asthma [10], arthritis [11], and heart disease [12]. They have also been found to assist in health behavior change, such as improving diet and weight loss [13], increasing physical activity [14], or in smoking cessation [15]. According to a Cochrane review of 24 randomized clinical trials that summarized the effects of different Web-based interventions for people with chronic diseases, these interventions had a significant effect on knowledge, perceived social support, health behaviors, clinical outcomes, and a possible positive effect on self-efficacy [2].

Web-based interventions that provide targeted support are more likely to be successful because information that is relevant to specific groups is more likely to be used [16]. However, successful targeting presents challenges because characteristics of the user groups of Web-based interventions are not clear. Recently the Comprehensive Model of Information Seeking (CMIS) was developed to better understand how user characteristics affect the use of Web-based interventions [17]. The model includes antecedent factors (demographics, personal experiences, salience, and beliefs), information carrier factors (characteristics and utility of the information channels), and information-seeking actions. The CMIS has recently been used as a framework to understand use of an interactive cancer communication system [18,19].

Demographic factors are considered important in predicting the use of health information resources in the CMIS framework, and are reported to influence use in several eHealth studies. Older age [20-23], female gender [20,24,25], higher education [20,21,24-28], and higher income [27,29] have all been associated with higher use in some studies, whereas other studies show younger age [26,28,29] to be associated with higher use. A recent systematic review of patients' acceptance of health information technologies revealed no consistent effect of age or gender on acceptance [30]. However, acceptance increased with higher education. On the other hand, level of education did not influence use of a Web-based diabetes program, and neither did age or health literacy [31]. These divergent findings might reflect ongoing change in Internet sociodemographics and dynamics [20,32].

According to CMIS, a person's direct experience with a disease will affect their need for information and predict their health-seeking behavior [17]. Different diagnoses cause different symptoms, need different treatment, and have different illness trajectories. Higher levels of functional well-being [33] and not having a chronic condition [22] are also associated with higher use of eHealth applications.

Psychosocial factors also affect information-seeking behaviors. For example, a person's health beliefs and perceived salience of the information affect information seeking [17], as does the individual's perception of their ability to control events. This could go both ways. A person with high self-efficacy might have higher confidence in seeking and using information in eHealth applications. On the other hand, for a person with low self-efficacy, eHealth applications could be an additional tool improving his or her confidence or capacity to deal with their symptoms and treatment [34]. The level of social support could affect how much information and support is needed. Lower levels of social support and symptoms of depression or negative mood have also been associated with higher use of eHealth applications [18,24,33,34]. In addition, prior Internet experience has been identified as a factor linked with increased use and acceptance of eHealth applications in some studies [23,30], but not others [31].

Although eHealth applications have been shown to be effective and can offer easier communication and cost savings, high rates of dropout and nonusage have been shown in many eHealth studies [35,36]. Moreover, users tend to use these applications differently than intended, indicating a need for a better understanding of patients' varying needs for support and for examining the best way to deliver eHealth applications. Research aiming to identify which components can be the most beneficial for different patient groups is critical in optimizing such systems to patients' preferences [2,35,37]. Because Web-based support systems usually include more than one component, it is not known yet which components may be particularly helpful to patients. Also, patients may prefer different types of support features, and their support needs may vary based on type of illness or user characteristics. Furthermore, although perceptions of a system's perceived usefulness have been investigated in a number of studies, the systems have primarily been evaluated as a whole on a set of general criteria, and the usefulness of specific components that the system offers have not been addressed. An average usefulness score in these types of studies may well result from a user who has evaluated some aspects as high and some as low. To design Internet systems that can better target different user groups, more research is needed examining

which user group characteristics are associated with different types and use of Internet support [2]. Exploring the characteristics of different users is of importance in refining and optimizing Web-based interventions to better fit different patient groups and increase their potential advantages.

The aim of this exploratory study was to describe user characteristics associated with the use of different components of WebChoice [7], a Web-based illness management support system for cancer patients. The following research questions were addressed:

1. What demographic-related, illness-related, and psychosocial variables are associated with the use of WebChoice?

2. Among WebChoice users, what are the associations among levels of patients' symptom distress, social support, depression, self-efficacy, health-related quality of life, and their use of different WebChoice components?

Methods

This exploratory study is a post hoc analysis of a large randomized controlled trial (RCT) in which WebChoice was tested among 325 breast cancer and prostate cancer patients who were followed with repeated measures for 1 year [7]. In the RCT, effects on symptom distress, depression, self-efficacy, health-related quality of life (HRQOL), and social support were measured. Use of WebChoice significantly reduced symptom distress, and patients in the WebChoice group also showed significant within-group improvements in depression during the study period. This was not observed in the control group; the control group also experienced significant deterioration of self-efficacy and HRQOL [7].

Recruitment took place between May 2006 and July 2007. Patients were recruited through advertisements in newspapers, on the Norwegian Cancer Society's website and in their magazine, and through information pamphlets mailed to patients through the Norwegian Cancer Registry. Patients who were interested called the research center to participate. Inclusion criteria were age over 18 years, able to read/speak Norwegian, having Internet access at home, and undergoing active treatment for breast or prostate cancer.

In total, 325 cancer patients took part in this study. Patients in the experimental group (n=162) who had access to WebChoice for 1 year constituted the sample for this paper. Participants received a user manual for WebChoice and a smart card-based public key solution for secure system access. They were instructed that they could use the system as much or as little as they liked. All data were submitted to a secure server using an encrypted connection. The study was approved by the Regional Committee for Medical and Health Research Ethics and the Data Security Inspectorate in Norway. Written informed consent was obtained from all participants. WebChoice could be used with both slow and fast Internet connections. Download times for the different components were the same.

Intervention

WebChoice [38] was developed in close cooperation with users and health care personnel [39]. The modules tested in this study targeted breast and prostate cancer patients and contained the following components (see Figure 1):

1. An assessment component in which patients could monitor their symptoms, problems, and priorities for support in physical, functional, and psychosocial dimensions. Patients choose symptoms and problems they were experiencing from a predefined list, and could rate the burden of these and what they needed help with. This information could be used to monitor improvement/deterioration of the condition, knowing when to alert health care personnel, preparing for a hospital/physician consultation, improving patient-provider communication, or with obtaining immediate access to self-management advice components described subsequently (Figure 2).

2. An advice component provided illness self-management support. Patients' self-reported symptoms/problems triggered the display of appropriate self-management activities that patients could choose from to relieve their symptoms and problems (Figure 2). Each choice contained an explanation of what the activity was; how to perform it; potential risks, side effects, and contraindications; when to contact a physician; levels of evidence; references to the source of the evidence; and links to other reliable websites for related information (Figure 3).

3. An information component in which patients had access to other reliable Web sources in Norwegian and English, such as information about tests, treatments, and potential side effects, lifestyle suggestions, and information about patients' legal rights.

4. A communication component for sharing experiences with other patients or for obtaining help from oncology nurses. Patients could participate in an online forum group discussion that allowed them to exchange messages anonymously with other patients, or use the online messaging system for private e-communication in which they could ask questions, share experiences, and get advice from oncology nurses. The nurses in this study were employed at the research center and were not involved in the direct care of the patients. They logged onto the communication component each weekday and contributed to the discussion group when appropriate.

5. An electronic diary in which patients could keep personal notes.

Figure 1. Screenshot of the WebChoice overview page.

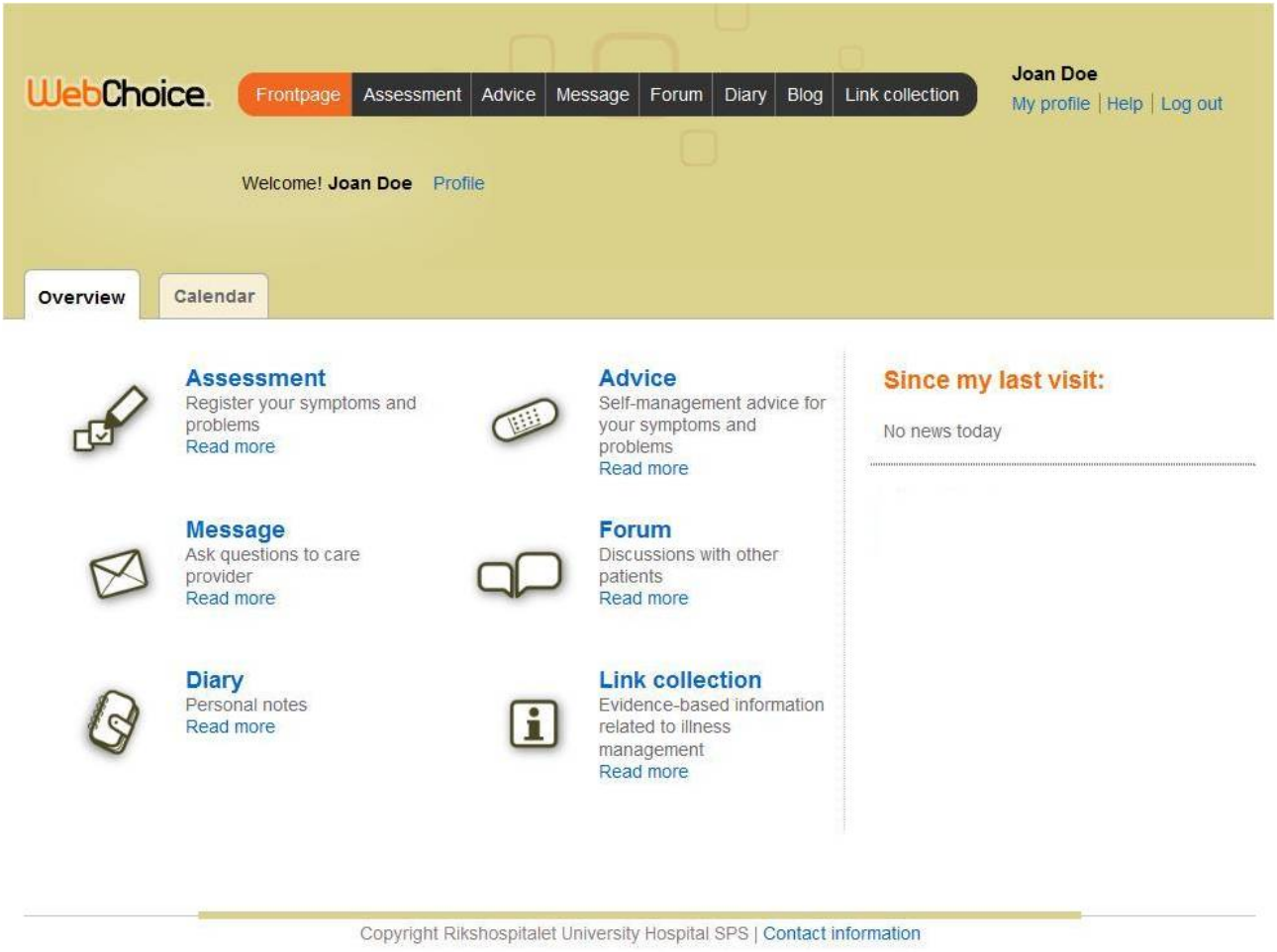


Figure 2. Screenshot of the results of an assessment and the associated advice/interventions.

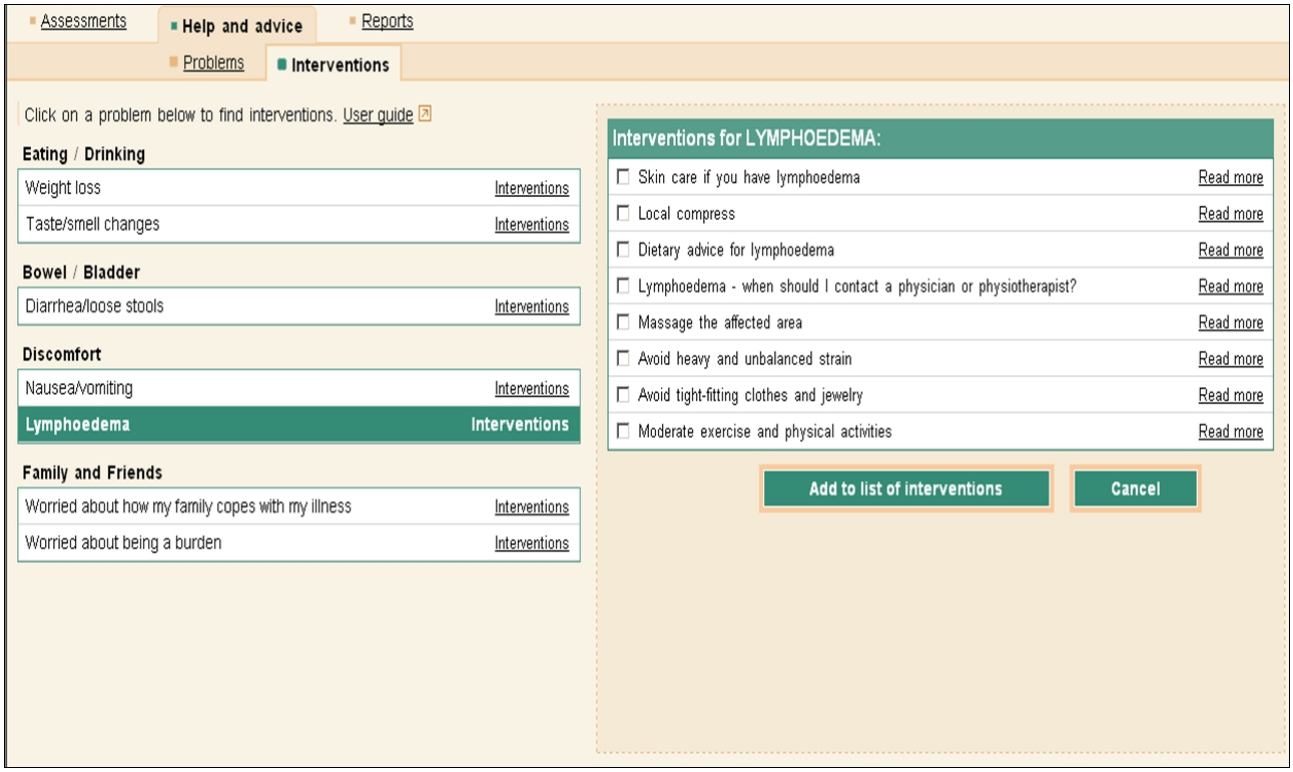
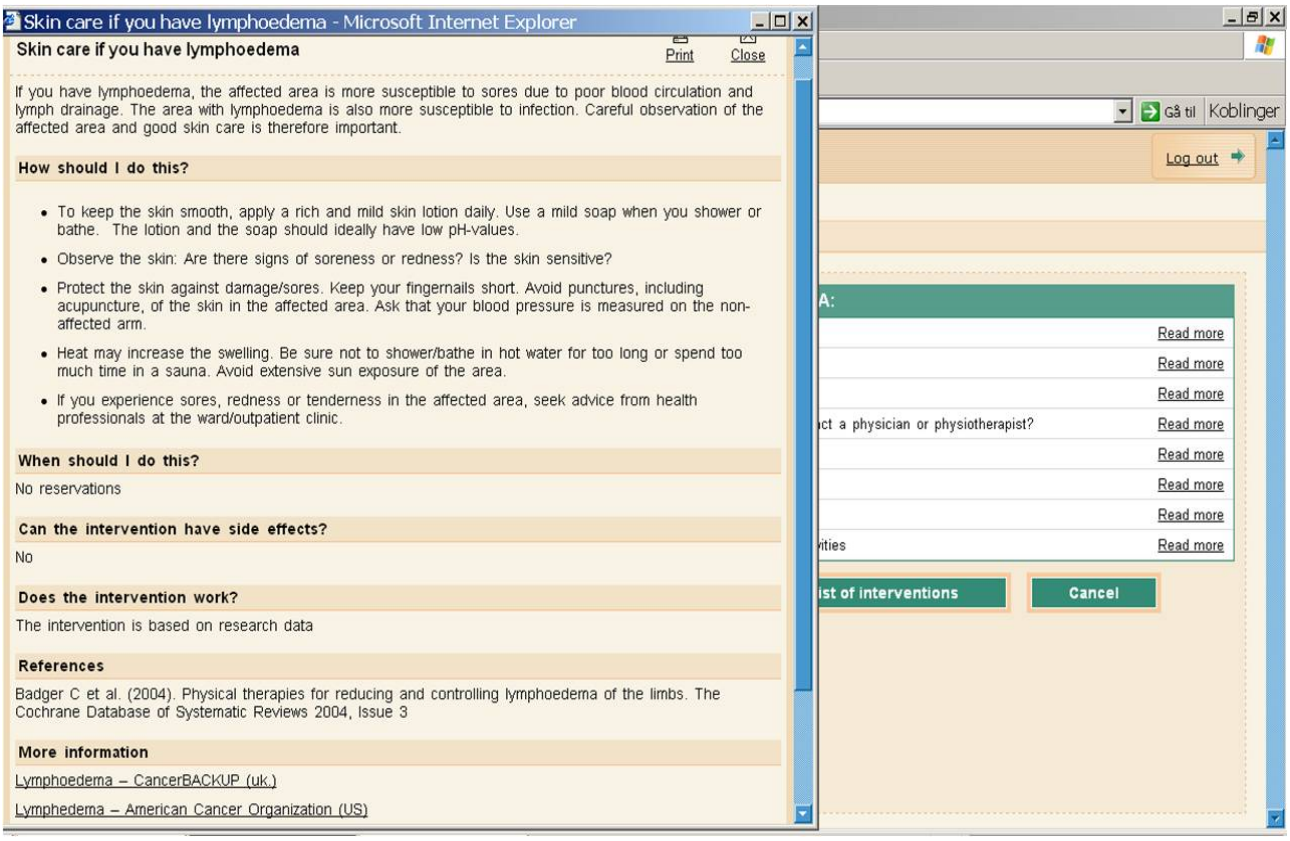


Figure 3. Screenshot showing an example of content and layout in the advice component.



Measures of System Use

Data on system use were extracted from the user logs on the server. Those who had logged on 0 to 1 times were categorized as nonusers; those who had logged on twice or more were categorized as users. We specified a minimum of 2 log-ins because patients who only logged on once may have only read the welcome message and never actually used the system. Information was collected on how many times the users logged on, how much time they spent on the site and on each component, and which components of WebChoice were accessed or used actively.

Table 1. Measures of system use of the WebChoice website.

Variable	Description
Total visits	Total number of times that a user logged on to the system.
Duration	Minutes spent using the system and its different components. If a user, when visiting a component, did not make a Web server request within the 20 minutes, the visit was ended. After the last Web server request during the visit, 10 minutes were added to the duration of this visit to reflect the fact that most usage consisted of reading information (an activity that cannot be logged). Duration is a complicated measure because of a lack of control over what the user actually does while logged on, and because some users do not log out of the system after visiting a component.
Assessments	Number of times a user chose symptoms and problems from a predefined list and generated their own symptom list.
Messages sent	Number of messages that users sent to the oncology nurse.
Posts in forum	Number of postings that users made in the discussion forum.
Diary notes	Number of notes made in the diary.
Visits to the different components	Number of times that a user entered a component. Unlike the previous measures, this measure does not indicate whether or not an action was taken within the component, for example, participants could visit a component to read what they had written previously or what others had written (eg, forum or answers from nurses).

User Characteristics

Demographic variables (age, gender, marital status, level of education, and household income) and diagnosis-related variables (diagnosis of breast or prostate cancer, time since diagnosis, recurrence/metastasis, type of treatment, and other illnesses) were recorded with a study-specific questionnaire.

Symptom distress was measured by using the 32-item Memorial Symptom Assessment Scale-Short Form (MSAS-SF) [40].

Symptom distress was measured with 5-point Likert scales, in which respondents rated the degree from 0 (not at all) to 4 (very much). Higher scores indicated greater symptom distress. Cronbach alpha for our sample at baseline was .92.

Social support was measured with the 20-item Medical Outcomes Study Social Support Survey (MOS-SS), including 5 subscales addressing emotional, instrumental, tangible, and affectionate support, and positive social interaction [41]. Responses on 5-point Likert scales ranged from 0 (none of the time) to 4 (all the time). Higher scores indicated more social support. Cronbach alpha for our sample at baseline was .96.

Depression was measured with the 20-item Center for Epidemiological Studies Depression scale (CES-D) [42] with responses on 4-point Likert scales ranging from 0 (rarely or none of the time) to 3 (most or all the time). Higher scores indicated greater depression. Cronbach alpha for our sample at baseline was .88.

Self-efficacy was measured with the 33-item Cancer Behavior Inventory (CBI) version 2.0 [43] which measured coping self-efficacy with cancer-related stress on 7 dimensions: (1) maintenance of activity and independence, (2) seeking and understanding medical information, (3) stress management, (4) coping with treatment-related side effects, (5) accepting cancer and maintaining a positive attitude, (6) affective regulation, and (7) seeking support. Responses on 9-point Likert scales ranged from 1 (not at all confident) to 9 (totally confident). Higher scores indicated greater self-efficacy. Cronbach alpha for our sample at baseline was .95.

HRQOL was measured using the 15-item 15D preference-based single index [44]. From 5 ordinal levels on each dimension, respondents chose the one that best described their present health status. Higher scores indicated greater HRQOL. Cronbach alpha for our sample at baseline was .77.

Computer experience was assessed with a simple question asking for patients' experience with computer use, ranging from 1 (no computer experience at all) to 5 (a lot of computer experience).

To prepare the data for latent class analysis (LCA), total scores on symptom distress, social support depression, self-efficacy, HRQOL, and the data on system use of the different WebChoice components were divided into tertiles based on the scores for breast and prostate cancer patients combined, using the whole sample of patients with access to WebChoice (Multimedia Appendix 1).

Marital status was dichotomized as married/cohabitating and any other status as single. Education was simplified from 4 to 3 categories. Elementary and high school were merged because there were few participants in the "elementary school only" group. Income was recoded from 5 to 3 categories. Computer experience was recoded from 5 to 3 categories because the 2 lowest ranks were related to no or little experience and the 2 highest described a great deal of computer experience.

Statistical Analyses

Data are presented as medians and interquartile ranges (IQRs) for continuous variables and as proportions for categorical data. Associations between categorical variables among users and nonusers were analyzed using the chi-square test for pairs of categorical variables; the Mann-Whitney test was used for continuous data with skewed distributions. A multiple logistic regression model was fitted to compare users and nonusers on demographic, disease-related, and psychosocial variables. Variables with a P value <.2 in the univariate logistic regression analyses were included in a multiple regression model. Based on the assumption that we could not expect a linear association between age and the outcome, age was first divided into several categories. Later, categories with similar odds ratios (ORs) were combined resulting in a dichotomizing of the age variable.

To identify characteristics associated with use of different components of WebChoice among users, we used LCA, which is a statistical method designed to identify if there are underlying types or subgroups of individuals that share specific characteristics [45]. LCA can best be thought of as an improved cluster analysis. It is a pattern recognition technique based on the statistical concept of likelihood and, thus, based on the same principle as factor analysis. The main difference is that cases are not absolutely assigned to classes, but have a probability of membership for each class. The results are presented with estimated probabilities and formal statistical tests can be performed to evaluate different models. The main goal for this LCA was to identify how a set of user-related variables could be associated with different system use. Utilizing LCA made us not only able to identify these associations, but also to quantify their direction and strength. Because of a limited sample size, we fitted LCA models with 3 classes and 4 explanatory variables at most to avoid overfitting and multicollinearity. As a final step, the best models were chosen based on the Akaike information criterion (AIC) [46] and the Bayesian information criterion (BIC) [47].

We expected different user patterns for breast and prostate cancer patients because these patient groups differed with regard to gender, age, treatment, and presence of other illnesses. Two-thirds (37/56) of the breast cancer patients were treated with chemotherapy compared to 26% (12/47) of the prostate cancer patients. Therefore, all LCA models were stratified by type of diagnosis and adjusted for age at inclusion. The cutoffs used to categorize variables were based on a tertile division of the entire sample. Because the division was based on the actual data and not on predefined cutoffs, a tertile division was chosen to provide categories with sufficient sizes and enable us to make distinctions among the high, medium, and low values. To ensure a sufficient number of breast and prostate cancer patients within each tertile-based category, patient numbers were checked for each variable (Multimedia Appendix 1) and the numbers were found to be satisfactory.

To select variables for inclusion in the final LCA models, we first fitted models in which 1 psychosocial variable at a time was tested together with 3 user variables. The psychosocial variables that revealed clear patterns of use were kept and integrated in the final LCA models, where a cluster of psychosocial variables was tested with single user variables.

The descriptive statistics and logistic regression were carried out using SPSS version 16.0 (SPSS Inc, Chicago, IL, USA).

LCA was performed with SAS version 9.3 (SAS Institute Inc, Cary, NC, USA), by using the PROC LCA procedure for LCA [48]. All tests were 2-sided and *P* values <.05 were considered statistically significant.

Results

Of the 162 participants who had access to WebChoice, 103 (63.6%) logged on twice or more over the 1-year study period and were defined as *users* (Table 2). There were no statistically significant differences in demographic, disease-related, or psychosocial variables between users and nonusers of WebChoice (Tables 2 and 3). Although not statistically significant, there were indications of higher use among men with prostate cancer (*P*=.09), patients without other illnesses (*P*=.08), and patients with more computer experience (*P*=.07) (see Table 2).

User Characteristics Associated with Use of WebChoice

Baseline scores on symptom distress, social support, depression, HRQOL, and self-efficacy did not differ significantly between users and nonusers; therefore, they were not included as covariates in the logistic regression models. Multiple logistic regressions revealed that high levels of computer experience (OR 3.77, 95% CI 1.20-11.91) and not having other illnesses (OR 2.10, 95% CI 1.02-4.34) were significantly associated with increased use of WebChoice after controlling for type of diagnosis and age (Table 4). The other illnesses most frequently reported were heart disease, rheumatic illness, lung disease, muscle and skeletal conditions, diabetes, and neurologic conditions.

Frequencies of Use of Different WebChoice Components

As displayed in Table 5, the WebChoice components visited 1 or more times by most of the users during the year of study participation were the advice component (98/103, 95.1%), the information component (96/103, 93.2%), assessments (95/103, 92.2%), and messages (93/103, 90.3%).

Patients visited the discussion forum far more often than they submitted their own postings, and the discussion forum was the component in which the patients spent most time (median 84 minutes, range 0-5108). Similarly, patients visited the message component far more often than they sent messages; the median time spent using this component was 21 minutes (range 0-701). The diary was among the least-used components in WebChoice. There were large differences between users. When we analyzed patterns of use of different components in WebChoice, it became apparent that different users used components quite differently, eg, a participant with high e-message use did not necessarily use the assessments or advice more often. Therefore, as the next step, we analyzed which user characteristics were associated with utilizing specific components.

Associations Between User Characteristics and Component Use

Overall, we did not find any association among use of different WebChoice components (measured in minutes) and demographics, most disease-related characteristics, or computer experience.

However, use patterns were different for men with prostate cancer and women with breast cancer. Levels of use were related to several psychosocial factors. When fitting separate models for 1 user characteristic at a time together with 3 WebChoice components, some important differences emerged (Table 6). Degree of social support and depression were more important for overall WebChoice use for women with breast cancer, whereas symptom distress was more influential for men with prostate cancer. High use of all WebChoice components was associated with low levels of social support among women with breast cancer. No such pattern was detected for men with prostate cancer. High use of advice, e-messages, and the discussion forum was associated with high levels of depression in women with breast cancer, but not in men with prostate cancer. For men with prostate cancer, high use of symptom assessments, advice, and the discussion forum was associated with high levels of symptom distress. Symptom distress did not appear to impact the use of WebChoice for women with breast cancer. No specific patterns of use were associated with levels of self-efficacy or HRQOL. Different combinations of patients' levels of social support, symptom distress, depression, and use of the components were explored with LCA. Because HRQOL and self-efficacy were not clearly associated with use, these variables were not included in the final LCA models. A summary of associations among patients' characteristics and selected components of WebChoice can be found in Table 6.

Use of the Message Component

The LCA revealed that lower levels of social support, higher levels of symptom distress, and higher depression were associated with higher use of messages to oncology nurses. This applied to both men with prostate cancer and women with breast cancer (Table 7). The estimated probabilities in Table 6 can be interpreted as follows: the model has identified a class of patients with breast cancer, latent class 1. Members of latent class 1 have a .60 probability of high use of messages, a .96 probability of low social support, a .74 probability of high symptom distress, and a .89 probability of depression.

Use of the Advice Component

Lower levels of social support, higher levels of symptom distress, and higher levels of depression were associated with higher use of the advice component among men with prostate cancer, but we did not detect such an association among women with breast cancer (Table 8).

Use of the Discussion Forum

Analyses of the discussion forum section revealed that higher levels of social support and lower level of symptom distress and depression among men with prostate cancer were associated with low use of the forum (Table 9). No clear pattern was detected among women with breast cancer.

No clear patterns of use associated with any user characteristics were found for the level of minutes spent at the assessment and information components.

Table 2. Characteristics of all breast and prostate cancer patients (users and nonusers) with access to WebChoice.

Characteristics	WebChoice access N=162	Users n=103	Nonusers n=59	*P* value
Demographic factors				
Age, median (range)	57 (35-80)	58 (36-79)	56 (35-80)	.15
Marital status, n (%)				.94
Married/cohabitating	135 (83.3)	86 (83.5)	49 (83.1)	
Single/divorced	27 (16.7)	17 (16.5)	10 (16.9)	
Education, n (%)				.89
Elementary/high school	62 (38.3)	38 (36.9)	24 (40.7)	
University/college ≤4 years	69 (42.6)	45 (43.7)	24 (40.7)	
University/college >4 years	31 (19.1)	20 (19.4)	11 (18.6)	
Household annual income (NOK), n (%) [a]				.63
<400,000	48 (29.6)	32 (31.1)	16 (27.1)	
400,000 to 600,000	44 (27.2)	26 (25.2)	18 (30.5)	
>600,000	65 (40.1)	44 (42.7)	21 (35.6)	
Missing data	5 (3.1)	1 (1.0)	4 (6.8)	
Disease-related factors				
Diagnosis, n (%)				.09
Breast cancer	96 (59.3)	56 (54.4)	40 (67.8)	
Prostate cancer	66 (40.7)	47 (45.6)	19 (32.2)	
Months since diagnosis, median (IQR)[b]	11.5 (21)	11.0 (22)	12.0 (20.3)	.52
Metastasis, n (%)	26 (16.0)	18 (17.5)	8 (13.6)	.51
Recurrence, n (%)	13 (8.0)	6 (5.8)	7 (11.9)	.17
Other illnesses, n (%)	63 (38.9)	35 (34.0)	28 (47.5)	.08
Psychosocial factors, median (IQR) [b]				
Symptom distress	28 (27)	29 (28)	25 (23)	.49
Social support	84 (28)	84 (28)	84 (31)	.38
Depression	9.5 (12)	10 (12)	9 (12)	.51
Self-efficacy	219 (67)	216 (71)	227 (64)	.57
Health-related quality of life	0.86 (0.16)	0.86 (0.15)	0.88 (0.16)	.49
Computer experience, n (%)				.07
None/little	16 (9.9)	6 (5.8)	10 (16.9)	
Medium	48 (29.6)	32 (31.1)	16 (27.1)	
High	93 (57.4)	62 (60.2)	31 (52.5)	
Missing data	5 (3.1)	3 (2.9)	2 (3.4)	

[a] NOK = Norwegian kroner (400,000 NOK≈US $67,000; 600,000 NOK≈US $100,000).

[b] IQR = Interquartile range.

Table 3. Baseline treatment characteristics of patients with access to WebChoice by diagnosis.

Diagnosis and treatment	WebChoice access n (%)	Users n (%)	Nonusers n (%)	*P* value
Breast cancer (n=96) [a]				
Radiotherapy	62 (65)	37 (66)	25 (63)	.72
Chemotherapy	71 (74)	40 (71)	31 (78)	.50
Hormone treatment	61 (64)	37 (66)	24 (60)	.54
Prostate cancer (n=66) [a]				
Radiotherapy	17 (26)	12 (26)	5 (26)	.95
Chemotherapy	4 (6)	3 (6)	1 (5)	.99
Hormone treatment	41 (62)	26 (74)	15 (88)	.30

[a] Patients can be given several treatments simultaneously.

Table 4. Binary logistic regression of patient characteristics and use of WebChoice (2 or more log-ins) (N=162).

Sociodemographic and health characteristics	Univariate analysis			Multiple analysis		
	OR[a]	95% CI	*P* value	OR[a]	95% CI	*P* value
Diagnosis						
Breast cancer (ref)	1.00					
Prostate cancer	1.77	0.90-3.45	.10	1.45	0.65-3.23	.36
Age						
<50 years (ref)	1.00					
≥50 years	1.69	0.83-3.42	.15	1.75	0.74-4.13	.20
Other illnesses						
Yes (ref)	1.00					
No	1.79	0.93-3.45	.08	2.10	1.02-4.34	.045
Computer experience						
Low (ref)	1.00					
Medium	3.33	1.02-10.81	.05	3.09	0.91-10.49	.07
High	3.33	1.11-10.02	.03	3.77	1.20-11.91	.02

[a]OR: odds ratio.

Table 5. Usage of different components of WebChoice over the year of accessibility (N=103).

Components in WebChoice	Times accessed			Users who accessed at least once	
	Median	IQR[a]	Range	n	%
Total visits	12	29	2-892	103	100
Total duration (minutes)	250	490	21-11,167	103	100
Assessments	2	5	0-51	77	74.8
Assessment visits	7	17	0-103	95	92.2
Assessment duration (minutes)	13	33	0-254	95	92.2
Advice visits	5	9	0-63	98	95.1
Advice duration (minutes)	15	32	0-372	98	95.1
Information section visits	4	7	0-97	96	93.2
Information duration (minutes)	25	54	0-431	96	93.2
Messages sent	1	5	0-49	62	60.2
Total messages visits	6	13	0-163	93	90.3
Message duration (minutes)	21	68	0-701	93	90.3
Posts in forum	0	4	0-58	50	48.5
Forum visits	8	36	0-536	87	84.5
Forum duration (minutes)	84	309	0-5108	87	84.5
Diary notes	1	4	0-142	54	52.4
Diary visits	2	6	0-94	73	70.9
Diary duration (minutes)	2	32	0-1003	73	70.9

[a] IQR=interquartile range.

Table 6. Summary of associations among single patient characteristics or a cluster of patients' characteristics and use of components in WebChoice stratified by diagnosis.

Characteristics	Associations with use of components in WebChoice	
	Prostate cancer	Breast cancer
Single characteristics [a]		
Low social support	No associations	High use of assessment, advice, information, messages, and forum
High symptom distress	High use of assessments, advice and forum	No associations
High depression	No associations	High use of advice, messages, and forum
Low health-related quality of life	No associations	No associations
Low self-efficacy	No associations	No associations
Cluster of characteristics [b]		
Low social support, high levels of symptom distress and high levels of depression	High use of messages and advice	High use of messages

[a] Tables with exact values for the probability of use of different components based on single patient characteristics can be found in Multimedia Appendix 2.

[b] Tables with exact values for the probability of use of different components based on a cluster of characteristics can be found in Tables 7-9.

Table 7. Use of the message component in WebChoice (in minutes). Latent class model, association with levels of social support, symptom distress, and depression. The numbers represent item probabilities. All models were stratified by diagnosis and adjusted for age at inclusion.

Variables	Prostate cancer latent class[a]			Breast cancer latent class[a]		
	1[b]	2	3	1[b]	2	3
Use of messages						
Low	.33	.37	.44	.20	.23	.45
Medium	.09	.32	.27	.20	*.76*	.07
High	*.57*	.31	.28	*.60*	.01	.48
Social support						
Low	*.63*	.09	.11	*.96*	.21	.26
Medium	.10	*.61*	.26	.03	*.52*	.41
High	.27	.30	*.63*	.02	.27	.32
Symptom distress						
High	*.51*	.36	.11	*.74*	.51	.01
Medium	.24	.38	.26	.24	.13	*.56*
Low	.25	.26	*.63*	.01	.36	.43
Depression						
High	*.87*	.01	.01	*.89*	.26	.07
Medium	.12	*.97*	.02	.10	.47	.36
Low	.02	.02	*.97*	.01	.27	*.57*

[a] Item response probabilities >.5 in italics to facilitate interpretation.

[b] Most prominent class.

Table 8. Use of the advice component in WebChoice (in minutes). Latent class model, association with levels of social support, symptom distress, and depression. The numbers represent item probabilities. All models were stratified by diagnosis and adjusted for age at inclusion.

Variables	Prostate cancer latent class[a]			Breast cancer latent class[a]		
	1[b]	2	3	1	2	3
Use of advice						
Low	.02	*.50*	.15	.06	.18	*.65*
Medium	.19	.25	*.83*	*.51*	.18	.21
High	*.80*	.25	.03	.43	*.64*	.15
Social support						
Low	*.69*	.09	.26	*.92*	.47	.01
Medium	.14	*.55*	.01	.08	.34	*.59*
High	.17	.35	*.73*	.01	.19	.40
Symptom distress						
High	*.79*	.29	.02	*.72*	.01	.41
Medium	.20	.33	.28	.22	.28	.34
Low	.02	.38	*.70*	.06	*.70*	.25
Depression						
High	*.92*	.01	.29	*.81*	.08	.18
Medium	.02	*.63*	.01	.19	.18	*.55*
Low	.06	.37	*.69*	.01	*.74*	.27

[a] Items response probabilities >.5 in italics to facilitate interpretation.

[b] Most prominent class.

Table 9. Use of the forum component in WebChoice (in minutes). Latent class model, association with levels of social support, symptom distress, and depression. The numbers represent item probabilities. All models were stratified by diagnosis and adjusted for age at inclusion.

Variables	Prostate cancer latent class[a]			Breast cancer latent class[a]		
	1	2	3[b]	1	2	3
Use of forum						
Low	*.57*	.27	*.93*	.09	.35	.15
Medium	.01	.44	.06	.49	.32	.41
High	.42	.29	.01	.42	.34	.43
Social support						
Low	*.73*	.11	.01	*.72*	.49	.01
Medium	.11	*.60*	.14	.23	.32	*.54*
High	.16	.29	*.85*	.05	.19	.45
Symptom distress						
High	*.55*	.25	.15	*.69*	.02	.33
Medium	.35	.34	.19	.22	.19	.43
Low	.10	.40	*.66*	.09	*.79*	.23
Depression						
High	*.73*	.01	.14	*.82*	.02	.02
Medium	.17	*.70*	.02	.17	.17	*.66*
Low	.10	.29	*.84*	.01	*.82*	.33

[a] Items response probabilities >.5 in italics to facilitate interpretation.

[b] Most prominent class.

Discussion

In this study, we explored how cancer patients' demographic, disease-related, and psychosocial factors were associated with the use of different components of a Web-based self-management support system for cancer patients. Men with prostate cancer and women with breast cancer who reported low levels of social support and high levels of symptom distress and depression indicating high illness burden made high use of sending e-messages to oncology nurses. Men with prostate cancer with these characteristics also made high use of the advice component. Levels of social support and depression were more important for use patterns among women with breast cancer, and symptom distress was more influential for men with prostate cancer.

Research studies designed to improve understanding of how different subgroups of patients use Web-based support systems can provide insight into how to target such systems and better meet the needs of user groups. This study makes an important contribution to this area. To our knowledge, this is the first study that systematically evaluated how a cluster of factors, such as social support, symptom distress, and depression, were associated with patients' use of different system components. Although use does not necessarily reflect usefulness or patients' preferences for different system components, the study results suggest that there are identifiable subgroups of patients who make different use of Web-based support. It also confirms that there are no "one size fits all" patterns of use or systems of support. Although much more research in this area is needed, our findings hold promise that we may eventually be able to identify specific patient characteristics or preferences through use of appropriate screening tools that will allow us to offer the right set of support components for specific patient groups.

It is acknowledged that behavioral and self-management interventions are more effective when they include more than 1 mechanism for support, but it is challenging to determine which components have the most positive effects. One way to address this question would be to perform RCTs with multiple intervention groups, assigning them different dosages or components of the Web-based support system. For example, Baker and colleagues [49] explored 3 different combinations of an interactive cancer communication system compared with regular Internet access. Results revealed that the information and support services significantly benefited breast cancer patients, but more complex and interactive services did not. Studies with multiple groups need a large sample size to be able to detect clinically relevant group differences, which is costly and not always feasible. Identification of subgroups of users, as in our study, does not predict the relative contribution of different components to achieve positive outcomes, but it can help identify potential candidates for component inclusion in future studies.

User Characteristics Associated with Patterns of Use

Among those who became users of WebChoice, our study suggests that several factors affected the use of different components. Participants with high symptom distress and

depression, indicating high illness burden, and who also had low social support utilized the e-message service and advice component the most. Demographic and other factors did not play a role in their use patterns. High illness burden accompanied with low social support might indicate a higher need for support than for those who do relatively well, and these patients might have potentially more to gain from the e-message component. The analyses of messages presented in a previous paper identified that living with physical symptoms and side effects, living with a fear of relapse, concerns about everyday life, and unmet informational needs from health care providers were important themes in these messages and were used to address both emotional and informational issues by breast and prostate cancer patients alike [50]. Although used by less than two-thirds of users (62/103, 60.2%), the message service was evaluated by patients as a supportive and useful component [51]. This is consistent with findings from a study on a Web-based support system for diabetes patients in which email communication with a nurse was also highlighted as an important reason for using the system [36]. Patients felt they received personal feedback and that the nurse looked after them. The opportunity to communicate directly with a health care provider seems to be an important and highly valued feature across different patient populations. Our study suggests that this component may particularly appeal to people with high illness burden and low levels of social support.

In our study, high use of the advice section was associated with low social support and high illness burden among men with prostate cancer, but less so for women with breast cancer. This is in line with findings on prostate cancer patients' preferences for informational support seen in studies of support groups [52,53]. Interestingly, no cluster of user characteristics were associated with high use of the forum. Although the forum was the component in which patients spent the most time, it seems that a discussion forum is not the place to turn to if one has little social support in addition to a high illness burden.

The finding that men with prostate cancer and women with breast cancer used WebChoice differently might indicate different needs for support and information. As described previously, low levels of social support without any other user characteristics was associated with high use of all components of WebChoice for women with breast cancer, but not for men with prostate cancer. Findings from a study of a computer support group for women with breast cancer were similar, showing a trend toward a higher volume of forum postings among those with lower levels of preexisting social support [33].

In this study, high levels of depression were associated with high use of several components in WebChoice among women with breast cancer whereas high levels of symptom distress were associated with high use among men with prostate cancer. One of the reasons could be difference in symptoms among prostate and breast cancer patients. For example, urinary incontinence and reduced sexual function are highly prevalent and bothersome for prostate cancer patients, potentially indicating different needs for support. It has been reported that men with prostate cancer use online support sites for information and women with breast cancer use them for emotional support [53,54]. The fact that low social support and high illness burden are associated with high use of several components, indicate that these characteristics are not barriers to use but rather function as motivators for use. The same pattern was seen previously in a study of an interactive cancer communication system [18].

Differences Between Users and Nonusers

Consistent with earlier studies, previous computer experience made the patients somewhat more likely to use WebChoice compared with those with no or little former experience. As Internet access and computer literacy have increased in society, it suggests that more patients may be reached with Web-based support in the future.

Individuals without other illnesses in our study were more than twice as likely to use WebChoice compared to those with multiple illnesses. This finding is consistent with previous studies showing that users of Internet interventions are healthier than nonusers [22,55]. Chronically ill people are also reported to be less eHealth literate [32], thus they may not regard Web-based interventions as a suitable alternative for them, or they may be too ill to benefit. This could indicate that, at present, it may be more difficult to reach those with a higher illness burden. Another explanation for our findings might be that some of the WebChoice components (assessment, information, and advice) in this study specifically targeted breast and prostate cancer patients; thus, it could have been considered of limited value for patients with multiple conditions.

Known demographic predictors for use of Web-based interventions, such as education [20,21,24,26-28,56] and income [27,29], were not associated with use in our sample. This might be related to the inclusion criterion of having Internet access, and the fact that the sample was self-recruited. When the study started in 2006, 69% of the Norwegian population had access to the Internet [57]. This increased to 84% in 2008. Those with Internet access at that time were younger and had higher education and income compared to those without access. This is reflected in the study sample, as our participants were higher educated than the general population in Norway at that time. Interestingly, we did not find any statistically significant association between age and gender and frequency of use; other studies have described younger people and women as the most frequent users of Web-based interventions [20,24-26,28,29]. In our study we found a trend, although not statistically significant, that individuals over 50 years were almost twice as likely to be users compared to those under 50 (OR 1.75, 95% CI .74-4.13, P=.20), which corresponds to more recent studies in which older people are reported to be more motivated to use e-consultation than younger people [58], and also to use eHealth applications [22]. Moreover, age and diagnosis/gender were closely correlated in our study because all breast cancer patients were women and tended to be younger than the men with prostate cancer. Gender might also be a factor, but because breast and prostate cancer diagnoses are gender specific, it was not possible to distinguish between the effect of gender and diagnosis. As more people become computer literate, we might see several new groups of users of Web-based interventions.

The finding that only two-thirds of participants actively used WebChoice is consistent with other studies on the use of Web-based interventions, and it raises the question of what may motivate patients to use a system such as WebChoice. Interestingly, baseline levels of factors such as social support, symptom distress, and depression did not predict whether a patient became a user; the only variables that were associated with use were previous computer experience and having additional illnesses. This suggests that other factors may also be at play. Very few studies have investigated patients' reasons for using or not using Web-based interventions, and previous studies on user experiences have primarily included active users only. Grimsbø et al [51] recently interviewed individuals with access to WebChoice to gain more insights into patients' reasons and motivations for their use or nonuse. These interviews suggested that cancer patients had different needs, and that WebChoice was meaningful and suitable for some patients, but not all. Although some described perceived helpfulness as their main reason for using the application, others reported they did not want to assume a "sick" role or be reminded of having cancer and wished to "get on with their lives" as reasons for nonuse [51].

An important question is whether the factors that predict usage might also predict benefit. Similarly, does amount of usage relate to benefit? Higher dosage of an intervention has been connected to better outcomes of behavioral change programs for fruit consumption and maintenance of weight loss [21,59]. However, the dose of use and its relationship with effect is rarely reported in effect studies of Web-based interventions. Volume of use does not necessary lead to increased benefit. As in our study, different users utilized different components. Thus, the "right component" could be the one that also has some benefit, but not necessarily total volume of use, depending on the importance of the information or support provided. For example, reading advice for a very bothersome symptom once may be enough to learn how to relieve it and, thus, be tremendously beneficial, whereas reading postings on the discussion forum many times may help one to feel good, but may not be equally beneficial for managing symptoms. Reading advice once would be registered as low usage, engaging in the forum as high. As noted by some of the users of WebChoice [51], reading information on WebChoice could upset them or calm them down, and extensive use could be based on the fear of missing information about the cancer.

Strengths and Limitations

The data from the server log provided detailed information on overall use of WebChoice and for each component on an individual and group level in this study, allowing us to perform the types of analysis presented here. Another strength is that we had baseline scores for all individuals and a low proportion of missing data. Data on use patterns of Web-based interventions can be challenging to analyze because they are rarely normally distributed, and there are often large variations in use. The use of LCA is a valid and valuable method used to analyze user patterns and identify subgroups of users according to specific characteristics, which adds to the strengths of this study.

However, several limitations also need to be addressed. First, participants in this study were self-selected and recruited through advertisement and pamphlets. In addition, they had to have Internet access. Increased willingness to participate in research studies is associated with higher education and use of the Internet (at least at the time of this study); therefore, these factors may limit generalizability. The men with prostate cancer in our study were younger (median 67 years) compared to patients diagnosed with prostate cancer in Norway in 2004 (median 72 years) [60]. Women with breast cancer in our study were also younger (median 50 years) compared to a study of 337 breast cancer patients in 2004 (median 55 years) [61]. Secondly, the relatively small sample put some constraints on the choice of statistical models. We chose to stratify all our statistical models by diagnosis/gender, thus making the comparison groups even smaller. In addition, the prostate cancer group was smaller than the breast cancer group, which might have further limited statistical power. The small sample size also reduced the number of variables that we were able to include in the LCA models. Given the limited statistical power, we chose to categorize our variables by using tertiles. We are aware that by doing so we might have lost some degree of information on differences in the underlying population. On the other hand, when fitting the models with categorical variables we were able to detect a direction on how low or high scores were associated with different use patterns of WebChoice. Computer experience and not having other illnesses were significantly related to increased probability of using WebChoice. However, given the size of our confidence intervals, reliability of our estimates may be somewhat uncertain. In addition, the measure of computer experience was based on a single question about the participant's own opinion about their experience with computer use on a 5-item scale and those specific results should be interpreted with caution.

Furthermore, because this is an exploratory study with many tests of associations, other studies are warranted to replicate these findings. Finally, we do not know if the observed differences between breast and prostate cancer patients were related to the 2 diagnoses or to gender. Therefore, to clarify, future studies should include cancer diagnoses that affect both women and men.

Conclusion

This study provides evidence about how different user characteristics influence the use of a Web-based illness self-management system among cancer patients. Such knowledge is crucial to target Web-based support systems to different patient groups. In our study, e-messages and advice for self-management support were components highly used by patients with low levels of social support and high levels of symptom distress and depression. Because patients with these characteristics may have a high need for support, these are components that may be particularly important to include in Web-based support systems for illness management support. Low levels of social support and high levels of depression influenced use of the system for women with breast cancer, whereas high levels of symptom distress influenced use for men with prostate cancer. Results highlight the importance of integration of multiple components in Web-based support

systems to address different needs and reasons for use. LCA is a useful technique to identify subgroups of users and can be successfully applied for the analysis of user patterns of Web-based interventions. Our results will be employed in the further development of the WebChoice application, and can be utilized by developers and researchers in creation and evaluation of Web-based interventions optimizing content for different user groups.

Acknowledgments

The study was funded by the South-East Regional Health Authority of Norway (grant number: 2009051) and The Norwegian Cancer Society.

Conflicts of Interest

CMR is the developer of the system, but has no ownership rights to the application.

Multimedia Appendix 1

Data-driven tertile division of variables that is necessary for latent class analysis models.

[PDF File (Adobe PDF File), 38KB - jmir_v15i3e34_app1.pdf]

Multimedia Appendix 2

Supplemental tables depicting levels of single patient characteristics and associations with use of different components in WebChoice.

[PDF File (Adobe PDF File), 42KB - jmir_v15i3e34_app2.pdf]

References

1. 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]
2. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005 Oct(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
3. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010 Jun;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
4. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
5. Lorig KR, Ritter PL, Dost A, Plant K, Laurent DD, McNeil I. The Expert Patients Programme online, a 1-year study of an Internet-based self-management programme for people with long-term conditions. Chronic Illn 2008 Dec;4(4):247-256. [doi: 10.1177/1742395308098886] [Medline: 19091933]
6. Solomon M, Wagner SL, Goes J. Effects of a Web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012;14(1):e32 [FREE Full text] [doi: 10.2196/jmir.1924] [Medline: 22353433]
7. Ruland CM, Andersen T, Jeneson A, Moore S, Grimsbø GH, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]
8. Gustafson DH, Hawkins R, McTavish F, Pingree S, Chen WC, Volrathongchai K, et al. Internet-based interactive support for cancer patients: are integrated systems better? J Commun 2008 Jun;58(2):238-257 [FREE Full text] [doi: 10.1111/j.1460-2466.2008.00383.x] [Medline: 21804645]
9. Yun YH, Lee KS, Kim YW, Park SY, Lee ES, Noh DY, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303. [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
10. van der Meer V, Bakker MJ, van den Hout WB, Rabe KF, Sterk PJ, Kievit J, SMASHING (Self-Management in Asthma Supported by Hospitals, ICT, Nurses and General Practitioners) Study Group. Internet-based self-management plus education compared with usual care in asthma: a randomized trial. Ann Intern Med 2009 Jul 21;151(2):110-120. [Medline: 19620163]
11. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]

12. Moore SM, Primm T. Designing and testing telehealth interventions to improve outcomes for cardiovascular patients. J Cardiovasc Nurs 2007;22(1):43-50. [Medline: 17224697]
13. 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 Dec;10(4):e52 [FREE Full text] [doi: 10.2196/jmir.1114] [Medline: 19073541]
14. Lieber SB, Redberg RF, Blumenthal RS, Gandhi A, Robb KJ, Mora S. A national interactive web-based physical activity intervention in women, evaluation of the american heart association choose to move program 2006-2007. Am J Cardiol 2012 Jun 15;109(12):1754-1760. [doi: 10.1016/j.amjcard.2012.02.017] [Medline: 22494850]
15. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
16. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003 Nov;27 Suppl 3:S227-S232. [Medline: 14672383]
17. Johnson JD. Cancer-Related Information Seeking (Health Communication). Cresskill, NJ: Hampton Press; 1997.
18. Han JY, Wise M, Kim E, Pingree R, Hawkins RP, Pingree S, et al. Factors associated with use of interactive cancer communication system: an application of the Comprehensive Model of Information Seeking. J Comput Mediat Commun 2010 Apr;15(3):367-388 [FREE Full text] [doi: 10.1111/j.1083-6101.2010.01508.x] [Medline: 21760702]
19. Han JY, Kim JH, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. J Health Commun 2012 Mar;17(3):356-371 [FREE Full text] [doi: 10.1080/10810730.2011.585696] [Medline: 22085215]
20. Im EO, Chee W, Liu Y, Lim HJ, Guevara E, Tsai HM, et al. Characteristics of cancer patients in internet cancer support groups. Comput Inform Nurs 2007;25(6):334-343 [FREE Full text] [doi: 10.1097/01.NCN.0000299655.21401.9d] [Medline: 18000430]
21. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res 2010 Jul;12(3):e29 [FREE Full text] [doi: 10.2196/jmir.1504] [Medline: 20663751]
22. 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 Apr;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
23. Kerr C, Murray E, Noble L, Morris R, Bottomley C, Stevenson F, et al. The potential of Web-based interventions for heart disease self-management: a mixed methods investigation. J Med Internet Res 2010 Dec;12(4):e56 [FREE Full text] [doi: 10.2196/jmir.1438] [Medline: 21156471]
24. Owen JE, Goldstein MS, Lee JH, Breen N, Rowland JH. Use of health-related and cancer-specific support groups among adult cancer survivors. Cancer 2007 Jun 15;109(12):2580-2589 [FREE Full text] [doi: 10.1002/cncr.22719] [Medline: 17503435]
25. 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 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
26. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371. [Medline: 15224975]
27. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Use of the Internet by women with breast cancer. J Med Internet Res 2002 Nov;4(2):E9 [FREE Full text] [doi: 10.2196/jmir.4.2.e9] [Medline: 12554556]
28. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007 Apr;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
29. Høybye MT, Dalton SO, Christensen J, Ross L, Kuhn KG, Johansen C. Social and psychological determinants of participation in internet-based cancer support groups. Support Care Cancer 2010 May;18(5):553-560. [doi: 10.1007/s00520-009-0683-6] [Medline: 19579037]
30. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009 Aug;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
31. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011 Jan;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
32. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
33. Shaw BR, Hawkins R, Arora N, McTavish F, Pingree S, Gustafson DH. An exploratory study of predictors of participation in a computer support group for women with breast cancer. Comput Inform Nurs 2006 Feb;24(1):18-27. [Medline: 16436908]
34. Shaw BR, Dubenske LL, Han JY, Cofta-Woerpel L, Bush N, Gustafson DH, et al. Antecedent characteristics of online cancer information seeking among rural breast cancer patients: an application of the Cognitive-Social Health Information

Processing (C-SHIP) model. J Health Commun 2008 Jun;13(4):389-408. [doi: 10.1080/10810730802063546] [Medline: 18569368]
35. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
36. Nijland N, van Gemert-Pijnen JE, 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 [FREE Full text] [doi: 10.2196/jmir.1603] [Medline: 21959968]
37. Williamson K. Where one size does not fit all: understanding the needs of potential users of a portal to breast cancer knowledge online. J Health Commun 2005 Sep;10(6):567-580. [doi: 10.1080/10810730500228961] [Medline: 16203634]
38. CommuniCare Tools. WebChoice URL: http://www.communicaretools.org/webchoice.aspx?lang=en-US [accessed 2012-05-31] [WebCite Cache ID 6849QNlnO]
39. Ruland CM, Jeneson A, Andersen T, Andersen R, Slaughter L, Bente-Schjødt-Osmo, et al. Designing tailored Internet support to assist cancer patients in illness management. AMIA Annu Symp Proc 2007:635-639 [FREE Full text] [Medline: 18693913]
40. Portenoy RK, Thaler HT, Kornblith AB, Lepore JM, Friedlander-Klar H, Kiyasu E, et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer 1994;30A(9):1326-1336. [Medline: 7999421]
41. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [Medline: 2035047]
42. Radloff LS, Teri L. Use of the Center for Epidemiological Studies-Depression Scale with older adults. WCLI 1986 Nov 1986;5(1):119-136 [FREE Full text] [doi: 10.1300/J018v05n01_06]
43. Merluzzi TV, Nairn RC, Hegde K, Martinez Sanchez MA, Dunn L. Self-efficacy for coping with cancer: revision of the Cancer Behavior Inventory (version 2.0). Psychooncology 2001 Jun;10(3):206-217 [FREE Full text] [Medline: 11351373]
44. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. Ann Med 2001 Jul;33(5):328-336. [Medline: 11491191]
45. Collins LM, Lanza ST. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. Hoboken, NJ: Wiley; 2010.
46. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr 1974 Dec 1974;19(6):716-723. [doi: 10.1109/TAC.1974.1100705]
47. Schwarz G. Estimating the dimension of a model. Ann Statist 1978 Mar 1978;6(2):461-464. [doi: 10.1214/aos/1176344136]
48. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. Struct Equ Modeling 2007;14(4):671-694 [FREE Full text] [Medline: 19953201]
49. Baker TB, Hawkins R, Pingree S, Roberts LJ, McDowell HE, Shaw BR, et al. Optimizing eHealth breast cancer interventions: which types of eHealth services are effective? Transl Behav Med 2011 Mar;1(1):134-145 [FREE Full text] [Medline: 21709810]
50. Grimsbø GH, Finset A, Ruland CM. Left hanging in the air: experiences of living with cancer as expressed through e-mail communications with oncology nurses. Cancer Nurs 2011 Apr;34(2):107-116. [doi: 10.1097/NCC.0b013e3181eff008] [Medline: 20921887]
51. Grimsbø GH, Engelsrud GH, Ruland CM, Finset A. Cancer patients' experiences of using an Interactive Health Communication Application (IHCA). Int J Qual Stud Health Well-being 2012;7. [doi: 10.3402/qhw.v7i0.15511] [Medline: 22582085]
52. Mo PK, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. Patient Educ Couns 2009 Apr;75(1):16-24. [doi: 10.1016/j.pec.2008.08.029] [Medline: 19233586]
53. Blank TO, Schmidt SD, Vangsness SA, Monteiro AK, Santagata PV. Differences among breast and prostate cancer online support groups. Computers in Human Behavior 2010 Nov 2010;26(6):1400-1404. [doi: 10.1016/j.chb.2010.04.016]
54. Seale C, Ziebland S, Charteris-Black J. Gender, cancer experience and internet use: a comparative keyword analysis of interviews and online cancer support groups. Soc Sci Med 2006 May;62(10):2577-2590. [doi: 10.1016/j.socscimed.2005.11.016] [Medline: 16361029]
55. 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 Jan;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
56. 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 Jan;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
57. Statistics Norway. The information society URL: http://www.ssb.no/ikt_en/ [accessed 2012-05-31] [WebCite Cache ID 6849tiJ9O]
58. 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]

59. 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]
60. Hernes E, Kyrdalen A, Kvåle R, Hem E, Klepp O, Axcrona K, et al. Initial management of prostate cancer: first year experience with the Norwegian National Prostate Cancer Registry. BJU Int 2010 Mar;105(6):805-11; discussion 811. [doi: 10.1111/j.1464-410X.2009.08834.x] [Medline: 19735258]
61. Dahl AA, Nesvold IL, Reinertsen KV, Fosså SD. Arm/shoulder problems and insomnia symptoms in breast cancer survivors: cross-sectional, controlled and longitudinal observations. Sleep Med 2011 Jun;12(6):584-590. [doi: 10.1016/j.sleep.2011.01.011] [Medline: 21645872]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
CBI: Cancer Behavior Inventory
CES-D: Center for Epidemiological Studies Depression scale
CMIS: Comprehensive Model of Information Seeking
HRQOL: health-related quality of life
IQR: interquartile range
LCA: latent class analysis
MOS-SS: Medical Outcomes Study Social Support Survey
MSAS-SF: Memorial Symptom Assessment Scale-Short Form
OR: odds ratio
RCT: randomized controlled trial

Edited by H Potts; submitted 27.07.12; peer-reviewed by K Pal, S Lepore; comments to author 15.08.12; revised version received 25.09.12; accepted 26.11.12; published 01.03.13.

Please cite as:
Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM
How User Characteristics Affect Use Patterns in Web-Based Illness Management Support for Patients with Breast and Prostate Cancer
J Med Internet Res 2013;15(3):e34
URL: http://www.jmir.org/2013/3/e34/
doi:10.2196/jmir.2285
PMID:23454601

Original Paper

Prevalence and Characteristics of Smokers Interested in Internet-Based Smoking Cessation Interventions: Cross-sectional Findings From a National Household Survey

Jamie Brown[1], Ph.D; Susan Michie[2,3], DPhil; Tobias Raupach[1,4], MD; Robert West[1,3], Ph.D

[1]Cancer Research UK Health Behaviour Research Centre, Department of Public Health and Epidemiology, University College London, London, United Kingdom

[2]Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

[3]National Centre for Smoking Cessation and Training, London, United Kingdom

[4]Department of Cardiology and Pneumology, University Medical Centre Gottingen, Gottingen, Germany

Corresponding Author:
Jamie Brown, Ph.D
Cancer Research UK Health Behaviour Research Centre
Department of Public Health and Epidemiology
University College London
1-19 Torrington Place
London, WC1E 6BT
United Kingdom
Phone: 44 7717886880
Fax: 44 20 7679 8354
Email: jamie.brown@ucl.ac.uk

Abstract

Background: An accurate and up-to-date estimate of the potential reach of Internet-based smoking cessation interventions (ISCIs) would improve calculations of impact while an understanding of the characteristics of potential users would facilitate the design of interventions.

Objective: This study reports the prevalence and the sociodemographic, smoking, and Internet-use characteristics of smokers interested in using ISCIs in a nationally representative sample.

Methods: Data were collected using cross-sectional household surveys of representative samples of adults in England. Interest in trying an Internet site or "app" that was proven to help with stopping smoking was assessed in 1128 adult smokers in addition to sociodemographic characteristics, dependence, motivation to quit, previous attempts to quit smoking, Internet and handheld computer access, and recent types of information searched online.

Results: Of a representative sample of current smokers, 46.6% (95% CI 43.5%-49.6%) were interested in using an Internet-based smoking cessation intervention. In contrast, only 0.3% (95% CI 0%-0.7%) of smokers reported having used such an intervention to support their most recent quit attempt within the past year. After adjusting for all other background characteristics, interested smokers were younger (OR=0.98, 95% CI 0.97-0.99), reported stronger urges (OR=1.29, 95% CI 1.10-1.51), were more motivated to quit within 3 months (OR=2.16, 95% CI 1.54-3.02), and were more likely to have made a quit attempt in the past year (OR=1.76, 95% CI 1.30-2.37), access the Internet at least weekly (OR=2.17, 95% CI 1.40-3.36), have handheld computer access (OR=1.65, 95% CI 1.22-2.24), and have used the Internet to search for online smoking cessation information or support in past 3 months (OR=2.82, 95% CI 1.20-6.62). There was no association with social grade.

Conclusions: Almost half of all smokers in England are interested in using online smoking cessation interventions, yet fewer than 1% have used them to support a quit attempt in the past year. Interest is not associated with social grade but is associated with being younger, more highly motivated, more cigarette dependent, having attempted to quit recently, having regular Internet and handheld computer access, and having recently searched for online smoking cessation information and support.

(J Med Internet Res 2013;15(3):e50) doi:10.2196/jmir.2342

KEYWORDS

smoking cessation intervention; Internet-based; website; prevalence; characteristic

Introduction

The World Health Organization recently attributed 12% of all global deaths among adults aged 30 years and over to tobacco [1]. Almost all these deaths could be avoided if smokers quit before their mid-30s [2]. Yet, in many countries such as the United Kingdom, less than a quarter of smokers quit by this age despite the majority wanting and trying to stop [3,4]. The most effective interventions involve face-to-face behavioral support combined with medication such as nicotine replacement therapy or varenicline [5-7]. However, even in England, where there is a universally available behavioral support program, the vast majority of smokers do not use face-to-face support and almost half attempt to stop unaided [8]. The Internet could be an ideal medium for helping those people who do not wish, or are unable, to engage in face-to-face behavioral support [9,10].

Behavioral support delivered via the Internet has the advantage that it is extremely cost-effective, and some patients prefer the increased convenience and confidentiality and reduced stigma [11,12], while others who are less able to access face-to-face support because of either mobility or geographical barriers may also find it useful. The benefits of Internet support over other low-cost and convenient alternatives to face-to-face support, such as written materials, include the capacity for interactivity and tailoring. Additionally, researchers and practitioners should be attracted by the capability to disseminate evidence-based support faithfully and flexibly update content to reflect new information as it emerges [13].

There is extensive evidence that the Internet can be an effective delivery mode for the behavioral support of a variety of health issues [11,14], and the United Kingdom has issued guidance to use particular programs in routine clinical care (eg, Beating the Blues for mild and moderate depression and FearFighter for phobia, panic, and anxiety) [15]. More importantly, there is also specific evidence from three separate systematic reviews that Internet-based smoking cessation interventions (ISCIs) can help smokers to quit compared with brief written materials or no intervention [16-18]. Current evidence is somewhat limited by the heterogeneity of effect across different interventions, insufficient reporting of content [19-21], and the paucity of data relating to long-term abstinence with biochemical verification of smoking status, yet research work is underway that may be able to address these limitations (eg, StopAdvisor [22,23]). In the context of this modest evidence of efficacy, together with the unique advantages of ISCIs, such as low cost, it is important to identify the prevalence of smokers who would be interested in using such support.

An accurate estimate of the likely reach of ISCIs is necessary for calculations of impact [24]. Previous estimates of potential reach have often been based on either national figures for Internet access or reported interest among nonrepresentative samples [25,26]. One study that did assess a representative sample of smokers estimated that 40% were interested in using an ISCI [27]. However, the study was conducted between 2006-07, and Internet access and usage patterns are relatively fast-moving phenomena [28]. For example, in Britain the percentage of households that have at least one method of using the Internet while at home increased from 58% in 2003 to 70% in 2009 and again to 77% in 2011 [28,29], while the use of wireless Internet hotspots doubled in just 12 months to 4.9 million users in 2011 [29].

Understanding the characteristics of smokers interested in using ISCIs may help the development of new interventions, or modification of existing ones, in several regards including tailoring dimensions, choice of content and features, navigational architecture, and language style and complexity. Similarly, designers would be interested in these associated characteristics for the purpose of dissemination, particularly online advertising, which can often be targeted to reach, or at least focus on, only certain demographic groups. Previous studies have characterized individuals who search for cessation information [30] and who use Internet interventions [31-35], smokers on their use of the Internet [36], and smokers who were either invited to, eligible for, or enrolled in cessation programs according to their subsequent use of the interventions [37-39]. While it is clearly essential to understand these profiles, particularly what determines use among those who are already interested, in order to improve the appeal of these cessation interventions it is also important to establish how interested smokers compare with those who are not in nationally representative samples. To our knowledge, only one other study has characterized a representative sample of smokers on the basis of their interest in ISCIs [27]. In that study, younger and more cigarette dependent smokers who had better Internet access were more likely to express an interest. However, there was no assessment of other important smoking characteristics such as current motivation to stop and past quit attempts, nor was there an assessment of recent online searching behavior.

This study addressed the following research questions:

1. How many smokers in a nationally representative sample are interested in using ISCIs?
2. What smoking, Internet use, and sociodemographic characteristics are associated with interest in the use of these interventions?

Methods

Study Design

The data were taken from the Smoking Toolkit Study [40], which is an ongoing series of cross-sectional household surveys in England designed to provide information about smoking prevalence and behavior. Each month a new sample of approximately 1800 adults aged 16 and over completes a face-to-face computer-assisted survey with a trained interviewer. By conducting a face-to-face rather than online survey, Internet access should not confound the results. Taylor Nelson Sofres-British Market Research Bureau collects the data as part of their monthly omnibus surveys on behalf of researchers at

the Cancer Research UK's Health Behaviour Research Centre, University College London, who conceived of the study and continue to manage it. The surveys use a form of random location sampling. England is split into 165,665 Output Areas, each comprising approximately 300 households. These Output Areas are stratified by A Classification Of Residential Neighbourhoods (ACORN) characteristics (an established geo-demographic analysis of the population provided by CACI International) and then randomly selected to be included in the lists of the interviewers. Interviewers travel to the selected areas and perform interviews with one participant per household until quotas based upon factors influencing the probability of being at home (working status, age, and gender) are fulfilled. Morning interviews are avoided to maximize participant availability. These survey methods have been previously described and have been shown to result in a baseline sample that is nationally representative in its sociodemographic composition and proportion of smokers [40]. Ethical approval was granted by the University College London ethics committee.

Participants

We used data from respondents to the survey between February 2012 and April 2012 who reported smoking cigarettes (including hand-rolled) daily or occasionally at the time of the survey. A total of 5405 adults were surveyed; 1190 reported currently smoking cigarettes regularly of whom 1128 had complete data on all relevant variables.

Measures

Current smokers were asked: "If there were an Internet site that was proven to help with stopping smoking, how likely is it that you would try it?" and also "If there were an application ("app") for your handheld computer (like a "smartphone" [eg, an iPhone, Blackberry, or Android phone], palmtop, PDA, or tablet) that was proven to help with stopping smoking, how likely is it that you would try it?". For the purposes of analysis, smokers' responses on 4-point scales were dichotomized as either being "interested" in using an Internet-based smoking cessation intervention (ie, those responding "very likely" or "quite likely" to either question) or "not interested" (ie, those responding "very unlikely" or "quite unlikely" to both questions).

Additionally, current smokers were asked questions that assessed gender, age, and social grade (AB = higher and intermediate professional/managerial, C1 = supervisory, clerical, junior managerial/administrative/professional, C2 = skilled manual workers, D=semi-skilled and unskilled manual workers, E=on state benefit, unemployed, lowest grade workers), dependence (Heaviness of Smoking Index, HSI [41] and Strength of Urges [42]), motivation to quit (Motivation to Stop Scale [43]), previous attempts to quit smoking, access to the Internet and handheld computers, and recent types of information searched online, that is, "For which of the following activities did you use the Internet in the last 3 months for private use? Please indicate all that apply: (a) using services related to travel and accommodation, (b) reading or downloading online newsnewspapersnews magazines, (c) looking for a job or sending a job application, (d) seeking health related information or support other than stopping smoking (eg, injury, disease, nutrition, improving health, etc), (e) seeking stop-smoking related information or support, (f) looking for information about education, training or courses, (g) doing an online course (in any subject), (h) consulting the Internet with the purpose of learning, (i) finding information about goods or services" [44]. Responses to Items (a) to (c) and (f) to (i) were aggregated to calculate a variable identifying use of the Internet for information other than health related (Item d) or smoking cessation (Item e).

Analysis

Data were analyzed using PASW 18.0.0. We used weighted data only to estimate the prevalence of interest in ISCIs among all smokers regardless of their Internet access. Data were weighted using the rim (marginal) weighting technique to match English census data on age, sex, and socioeconomic group. To assess smoking, Internet use, and sociodemographic characteristics associated with interest in the use of ISCIs, we conducted a series of simple and multiple logistic regressions. Alpha was set at $P<.05$.

Results

Approximately 70% of current smokers had accessed the Internet in the past week while a significant majority also had access to a handheld computer (see Table 1). A minority of users had searched for either smoking or health information support, and more than half had searched for at least one of a variety of "other" types of online information. The sociodemographic and smoking characteristics were typical of a representative sample of smokers [4,40], and by way of comparison, the characteristics of the 4451 current smokers included in the Smoking Toolkit Study for the 12 months before the current study (ie, January 2011 to January 2012) are presented in Table 1.

A total of 42.6% (95% CI 39.6%-45.7%) of current smokers were interested in using Internet sites for smoking cessation, 23.9% (95% CI 21.3%-26.5%) were interested in apps, and 46.6% (95% CI 43.5%-49.6%) were interested in ISCIs (either sites or apps). In contrast, only 0.3% (95% CI 0%-0.7%) of smokers reported having used such an intervention to support their most recent quit attempt within the past year.

Table 2 shows the smoking, Internet use, and sociodemographic characteristics of smokers by their interest in the use of ISCIs. There was evidence that interested smokers were younger, more cigarette dependent (measured by both HSI and Strength of Urges), more motivated to quit within 3 months, more likely to have made a quit attempt in the past year, accessed the Internet at least weekly, had handheld computer access, had used the Internet to search for online smoking cessation information or support in past 3 months, and had used it to search for a variety of "other" online information. After adjusting for all other background characteristics, associations remained between interest and age, cigarette dependence (measured by Strength of Urges), motivation to quit, past year quit attempt, weekly Internet access, handheld computer access, and recent searching for online smoking cessation information. Last, this pattern of results was unchanged during sensitivity analyses in which the associations between interest and the various characteristics

were re-assessed separately when smokers were classified according to whether or not they had expressed interest in (1) an Internet site, or (2) an app (data not shown).

Table 1. Characteristics of current smokers.

	Current sample of Smoking Toolkit Study: Feb. ‘12 to Apr. ‘12 (n=1128)	Population of Smoking Toolkit Study: Jan. ‘11 to Jan. ‘12 (n=4451)
Mean (SD) age	41.7 (16.7)	43.0 (16.9)
% (N) women	47.6 (537)	49.7 (2210)
% (N) social grade C2DE	72.4 (817)	69.9 (3112)
Mean (SD) heaviness of smoking index	2.0 (1.5)	2.1 (1.5)
Mean (SD) strength of urges score	2.1 (1.0)	2.1 (1.1)
% motivated to quit within 3 months	20.3 (229)	22.9 (1020)
% quit attempt in past year	30.9 (348)	30.5 (1358)
% using face-to-face behavioral support in most recent quit attempt within past year	1.3 (15)	1.1 (49)
% access Internet at least weekly	70.1 (791)	—
% handheld computer access	41.6 (469)	—
% seeking online cessation information or support in past 3 months	3.0 (34)	—
% seeking online health (not smoking) information or support in past 3 months	9.5 (107)	—
% seeking “other” online information or support in past 3 months	63.5 (716)	—

Table 2. Factors associated with interest in Internet-based smoking cessation interventions.

	Interested (n=488)	Not interested (n=640)	OR (95% CI)	Adj. OR (95% CI)
Mean (SD) age	36.6 (13.5)	45.6 (17.9)	0.97 (0.96-0.97)[a]	0.98 (0.97-0.99)[a]
% (N) women	47.3 (231)	47.8 (306)	0.98 (0.77-1.24)	1.06 (0.81-1.38)
% (N) social grade C2DE	69.5 (339)	74.7 (478)	0.77 (0.59-1.00)	0.98 (0.73-1.32)
Mean (SD) heaviness of smoking index	2.1 (1.5)	1.9 (1.5)	1.10 (1.02-1.19)[a]	1.10 (1.00-1.22)
Mean (SD) strength of urges score	2.3 (1.0)	2.0 (1.0)	1.33 (1.18-1.50)[a]	1.29 (1.10-1.51)[a]
% motivated to quit within 3 months	29.3 (143)	13.4 (86)	2.67 (1.98-3.60)[a]	2.16 (1.54-3.02)[a]
% quit attempt in past year	41.0 (200)	23.1 (148)	2.31 (1.78-2.99)[a]	1.76 (1.30-2.37)[a]
% using face-to-face behavioral support in most recent quit attempt within past year	1.8 (9)	0.9 (6)	1.99 (0.70-5.62)	0.96 (0.28-3.27)
% access Internet at least weekly	85.9 (419)	58.1 (372)	4.37 (3.24-5.90)[a]	2.17 (1.40-3.36)[a]
% handheld computer access	56.4 (275)	30.3 (194)	2.97 (2.32-3.80)[a]	1.65 (1.22-2.24)[a]
% seeking online cessation information or support in past 3 months	5.3 (26)	1.2 (8)	4.45 (1.99-9.91)[a]	2.82 (1.20-6.62)[a]
% seeking online health (not smoking) information or support in past 3 months	11.3 (55)	8.1 (52)	1.44 (0.96-2.14)	0.82 (0.53-1.29)
% seeking "other" online information or support in past 3 months	77.7 (379)	52.7 (337)	3.13 (2.40-4.07)[a]	1.43 (0.97-2.10)

[a]$P<.05$.

Discussion

Almost half of all current smokers were interested in using an Internet-based smoking cessation intervention, however less than 1% had used one to support their most recent quit attempt in the past year. After adjustment for all background characteristics, smokers who were younger, more dependent, highly motivated to quit, had attempted to quit recently, accessed the Internet regularly, had handheld computer access, and had recently searched for online smoking cessation information or support were all more likely to be interested in using online stop smoking support.

The diffusion of the Internet since its inception over 40 years ago has been phenomenal—recently, the Internet reached a billion users worldwide [28]. In Britain, Internet access has continued to increase with 77% of households connected in 2011 as compared to 58% in 2003 [28,29]. Although smokers tend to have less access than nonsmokers [27], there remains a majority of smokers that would be possible to reach via the Internet—in this study over 70% had used the Internet in the past week—and this number is only likely to increase [28]. As a consequence, ISCIs are often cited as offering a valuable opportunity to deliver low-cost behavioral support to large numbers of smokers [16,21,25,45]. Importantly, this study now adds an up-to-date estimate of the proportion of smokers interested in using these interventions, which provides some indication of their maximum potential reach and should allow more accurate calculations of the likely impact of particular interventions [24]. For example, from the RE-AIM perspective, a public health impact score can be represented as a multiplicative combination of reach, efficacy, adoption, implementation, and maintenance. The accuracy of this calculation for particular ISCIs may be improved by the provision in the current paper of an up-to-date estimate of the denominator necessary to calculate the reach, which is defined as the proportion of the possible target population that participate in a particular intervention.

The current estimate of 47% of smokers who are interested in ISCIs is higher than the 40% previously estimated from a representative sample of smokers [27]. However, that study was conducted in 2006-07, and it is reasonable to assume that interest may have increased as a consequence of improving Internet access [28]. Additionally, interest in ISCIs was operationalized as an expression of interest in either an Internet site or app—the number reporting interest only in Internet sites was 43%. Last, the first study was conducted in Canada, and it is likely there are cultural differences in interest as compared to England.

The finding that interested smokers were likely to be younger and use the Internet more regularly is consistent with previous research [27]. This association with age is particularly important as treatment-seeking smokers tend to be older than those who do not seek treatment [46], and therefore online interventions may be particularly suitable for targeting younger smokers who may otherwise attempt to quit unaided. In contrast, the association between interest and dependence, motivation, and past year quit attempts is characteristic of smokers who are more likely to seek treatment [46-48]. It is important that any future assessments of the real-world effectiveness of ISCIs take these associations into account [48].

The association between interest and recent searching for online smoking cessation information or support is intuitive. However, it is also indicative of the potential demand for these interventions in that smokers do not appear to be deterred nor satisfied by what they are currently finding. The latter point is also suggested by the contrast between the 3% of smokers who recently searched for online smoking cessation information or support as compared with the 0.3% who used online support during a quit attempt in the past year.

The “digital divide” that characterizes Internet use in the wider population is also true of smokers: smokers who use the Internet tend to be more affluent and educated than those who do not [36]. Therefore, the interest of smokers regardless of social grade in the current study is an unexpected finding and suggests the Internet may yet offer a means of equitable treatment delivery, which is particularly important as smokers from more deprived socioeconomic groups typically want, and try, to stop as much as other smokers but find it more difficult [49].

A potential concern with ISCIs is that it might prevent smokers from using face-to-face support, which is currently the most effective delivery mode for behavioral support [5-7]. While relatively few smokers had recently used face-to-face support in the current study, the lack of an association between use of this support and interest tentatively suggests the concern is unwarranted. Instead, it is likely that Internet support would appeal to a different subset of smokers who place more value on convenience and confidentiality, or find other types of support difficult to access [11]. Additionally, in other health areas where Internet support has become routine, the two have emerged as complementary, for example in several areas of mental health online cognitive behavioral therapy is frequently recommended while a patient waits for an appointment or to help with the more automated aspects of the support [11].

Taken together, the pattern of associations provided in the current study present a comprehensive characterization of smokers interested in using ISCIs. This understanding is important as access cannot primarily account for the difference between interest in and use of these interventions; for example, in the current study the majority of smokers had used the Internet in the past week. It is hoped that these associations will inform the development or modification of interventions with regards to targeted dissemination, tailoring dimensions, choice of content and features, navigational architecture, and language style and complexity, and in turn help to actualize the potential of the Internet for delivering smoking cessation interventions [50,51].

An indirect point of interest is the finding that in simple logistic regressions, interest was associated with both measures of dependence: Strength of Urges and HSI. However, the association with only Strength of Urges in multiple logistic regression is consistent with previous research showing that a single-rating measure of urges may be a more useful measure of dependence than those based on consumption [42].

One possible limitation is that an expression of interest during a survey is clearly quite different from actually visiting a program to support a quit attempt and subsequently using the program regularly. However, the purpose of the current study was primarily to provide an estimate of the maximum possible reach if the support was more widely available and promoted. One of the important findings is the huge potential in terms of the difference between the proportion who are interested and the percentage currently using. Additionally, establishing the characteristics associated with interest is arguably critical to realizing this potential and improving uptake by facilitating effective targeting, both in terms of advertising and intervention content, and complements research into understanding the determinants of use among the small proportion currently using, eg, [37-39]. Future research should aim to derive the specific relationship between interest and subsequent uptake of ISCIs following targeted design and promotion. Another limitation is that the questions used to derive interest in ISCIs were framed positively (eg, “If there were an Internet site or ‘app’ that was proven to help with stopping smoking, how likely is it that you would try it?”). However, even if interest is overestimated, it is improbable that it would account for the extent of the discrepancy between interest in and use of ISCIs, which is highlighted by the current study. Another potential limitation may have been to operationalize interest in ISCIs as an expression of interest in either an Internet site or an app. While both require the Internet and a computer for delivery, there are also clear differences, including that only sites require regular Internet access and only apps require ongoing access to a handheld computer. Indeed, fewer smokers were interested in apps as compared to Internet sites; however, similar proportions of smokers who had regular Internet access and those with ongoing access to handheld computers were interested in sites and apps respectively. More importantly, in sensitivity analyses in which smokers were characterized separately according to interest in either sites or apps, the pattern of results was unchanged. Together these results suggest that the current operationalization of ISCIs for the purposes of estimating and characterizing interest is suitable; however, experimental research is needed to establish whether there are important differences in efficacy according to delivery mode of either

Internet site or app. A final limitation is potential error or bias in the measurement of certain smoking characteristics. For example, smokers often forget failed quit attempts, particularly if they only lasted a short time or occurred long ago [52]. However, it is unlikely that forgetting would differ as a function of interest in ISCIs, and quit attempts—as with all other variables of interest in the present study—were assessed with a validated measure.

In conclusion, almost half the smokers in England are interested in using online smoking cessation interventions, and yet only a small proportion of smokers currently use these interventions to support quit attempts. Clearly, ISCIs represent an excellent opportunity to deliver low-cost behavioral support to a large number of smokers, which is currently not being realized. Moreover, as interest is expressed regardless of social grade, the Internet may also offer a means of delivering this support equitably. Last, designers of Internet-based interventions should be aware that potential users are likely to be younger, more cigarette dependent, highly motivated, have attempted to quit recently, have regular Internet and handheld computer access, and have recently searched for smoking cessation information and support.

Acknowledgments

We are grateful to Cancer Research UK and the Department of Health for funding this study. At the outset, data collection for the Smoking Toolkit Study was supported by Pfizer, GlaxoSmithKline, and Johnson and Johnson. The research team is part of the UK Centre for Tobacco Control Studies. We thank John Stapleton and Fabiana Lorencatto for providing comments on a draft of this manuscript.

Conflicts of Interest

TR has received honoraria from Pfizer, Novartis, GlaxoSmithKline, AstraZeneca, and Roche as a speaker in activities related to continuing medical education. RW undertakes research and consultancy and receives fees for speaking from companies that develop and manufacture smoking cessation medications (Pfizer, Johnson & Johnson, McNeil, GlasxoSmithKline, Nabi, Novartis, and Sanofi-Aventis). He also has a share of a patent for a novel nicotine delivery device. JB & SM have no conflicts.

References

1. World Health Organization. WHO Report on the Mortality Attributable to Tobacco. 2012. URL: http://www.who.int/tobacco/publications/surveillance/rep_mortality_attributable/en/index.html [accessed 2013-02-20] [WebCite Cache ID 6Ea33XpP3]
2. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004 Jun 26;328(7455):1519 [FREE Full text] [doi: 10.1136/bmj.38142.554479.AE] [Medline: 15213107]
3. Erikson M, Mackay J, Ross H. The Tobacco Atlas: Fourth Edition. 3rd ed. Atlanta, Georgia: American Cancer Society; 2012.
4. West R, Brown J. Smoking and Smoking Cessation in England. London; 2011. URL: http://www.smokinginengland.info/ [accessed 2012-09-06] [WebCite Cache ID 6ATIwUecr]
5. Fiore MCJC, Jaen CR, Baker TB. Clinical Practice Guideline. Rockville, MD: US Department of Health and Human Services. Public Health Service; 2008. Treating tobacco use and dependence: 2008 update URL: http://bphc.hrsa.gov/buckets/treatingtobacco.pdf [WebCite Cache ID 6EtfEdtDY]
6. Department of Health. NHS Stop Smoking Services - Service and Monitoring Guidance. 2010. URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/@sta/@perf/documents/digitalasset/dh_109889.pdf [accessed 2013-02-20] [WebCite Cache ID 6Ea3Utgru]
7. Brose LS, West R, McDermott MS, Fidler JA, Croghan E, McEwen A. What makes for an effective stop-smoking service? Thorax 2011 Oct;66(10):924-926. [doi: 10.1136/thoraxjnl-2011-200251] [Medline: 21709164]
8. Raupach T, West R, Brown J. The most "successful" method for failing to quit smoking is unassisted cessation. Nicotine Tob Res 2013 Mar;15(3):748-749. [doi: 10.1093/ntr/nts164] [Medline: 22990213]
9. Saul JE, Schillo BA, Evered S, Luxenberg MG, Kavanaugh A, Cobb N, et al. Impact of a statewide Internet-based tobacco cessation intervention. J Med Internet Res 2007;9(3):e28 [FREE Full text] [doi: 10.2196/jmir.9.4.e28] [Medline: 17942390]
10. Graham AL, Cobb NK, Raymond L, Sill S, Young J. Effectiveness of an internet-based worksite smoking cessation intervention at 12 months. J Occup Environ Med 2007 Aug;49(8):821-828. [doi: 10.1097/JOM.0b013e3180d09e6f] [Medline: 17693778]
11. Marks IM, Cavanagh K, Gega L. Computer-aided psychotherapy: revolution or bubble? Br J Psychiatry 2007 Dec;191:471-473 [FREE Full text] [doi: 10.1192/bjp.bp.107.041152] [Medline: 18055948]
12. Murray E. Internet-delivered treatments for long-term conditions: strategies, efficiency and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2008 Jun;8(3):261-272. [doi: 10.1586/14737167.8.3.261] [Medline: 20528378]
13. Andersson G. The promise and pitfalls of the internet for cognitive behavioral therapy. BMC Med 2010;8:82 [FREE Full text] [doi: 10.1186/1741-7015-8-82] [Medline: 21138574]

14. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
15. National Institute for Health and Clinical Excellence. Computerised cognitive behaviour therapy for depression and anxiety. Technology Appraisal 2008;97.
16. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
17. Myung SK, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
18. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
19. Crutzen R. Is it time to start focusing on the content of computer-delivered interventions? Addiction 2011 Oct;106(10):1871. [doi: 10.1111/j.1360-0443.2011.03539.x] [Medline: 21793973]
20. Michie S, Abraham C. Advancing the science of behaviour change: a plea for scientific reporting. Addiction 2008 Sep;103(9):1409-1410. [doi: 10.1111/j.1360-0443.2008.02291.x] [Medline: 18783495]
21. Strecher VJ. The Internet: Just another smoking cessation tool? Addiction 2008;103(3):485-486.
22. Michie S, Brown J, Geraghty A. Development of StopAdvisor: A theory-based interactive internet-based smoking cessation intervention. Translational Behavioral Medicine 2012;2(3):263-275.
23. Brown J, Michie S, Geraghty AW, Miller S, Yardley L, Gardner B, et al. A pilot study of StopAdvisor: a theory-based interactive internet-based smoking cessation intervention aimed across the social spectrum. Addict Behav 2012 Dec;37(12):1365-1370. [doi: 10.1016/j.addbeh.2012.05.016] [Medline: 22795643]
24. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
25. 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]
26. 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 [FREE Full text] [doi: 10.2196/jmir.1666] [Medline: 21873150]
27. Cunningham JA. Access and interest: two important issues in considering the feasibility of web-assisted tobacco interventions. J Med Internet Res 2008;10(5):e37 [FREE Full text] [doi: 10.2196/jmir.1000] [Medline: 18984558]
28. Dutton WH, Helsper EJ, Gerber MM. The Internet in Britain. London: Oxford Internet Institute; 2009.
29. Office for National Statistics. Internet Access - Households and Individuals. 2011. URL: http://www.ons.gov.uk/ons/rel/rdit2/internet-access---households-and-individuals/2011/index.html [accessed 2013-02-20] [WebCite Cache ID 6Ea4Ib9ab]
30. Cobb NK, Graham AL. Characterizing Internet searchers of smoking cessation information. J Med Internet Res 2006;8(3).
31. 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]
32. Christensen H, Griffiths MK, 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]
33. 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]
34. Neil AL, Batterham P, Christensen H, Bennett K, Griffiths KM. Predictors of adherence by adolescents to a cognitive behavior therapy website in school and community-based settings. J Med Internet Res 2009;11(1):e6 [FREE Full text] [doi: 10.2196/jmir.1050] [Medline: 19275982]
35. 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]
36. Stoddard JL, Augustson EM. Smokers who use internet and smokers who don't: data from the Health Information and National Trends Survey (HINTS). Nicotine Tob Res 2006 Dec;8 Suppl 1:S77-S85. [Medline: 17491174]
37. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. J Med Internet Res 2008;10(5):e54 [FREE Full text] [doi: 10.2196/jmir.1016] [Medline: 19097975]
38. Danaher BG, Lichtenstein E, McKay HG, Seeley JR. Use of non-assigned smoking cessation programs among participants of a Web-based randomized controlled trial. J Med Internet Res 2009;11(2):e26 [FREE Full text] [doi: 10.2196/jmir.1172] [Medline: 19632976]
39. 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 [FREE Full text] [Medline: 17491170]

40. Fidler JA, Shahab L, West O, Jarvis MJ, McEwen A, Stapleton JA, et al. 'The smoking toolkit study': a national study of smoking and smoking cessation in England. BMC Public Health 2011;11:479 [FREE Full text] [doi: 10.1186/1471-2458-11-479] [Medline: 21682915]
41. Borland R, Yong HH, O'Connor RJ, Hyland A, Thompson ME. The reliability and predictive validity of the Heaviness of Smoking Index and its two components: findings from the International Tobacco Control Four Country study. Nicotine Tob Res 2010 Oct;12 Suppl:S45-S50 [FREE Full text] [doi: 10.1093/ntr/ntq038] [Medline: 20889480]
42. Fidler JA, Shahab L, West R. Strength of urges to smoke as a measure of severity of cigarette dependence: comparison with the Fagerström Test for Nicotine Dependence and its components. Addiction 2011 Mar;106(3):631-638. [doi: 10.1111/j.1360-0443.2010.03226.x] [Medline: 21134020]
43. Kotz D, Brown J, West R. Predictive validity of the Motivation To Stop Scale (MTSS): a single-item measure of motivation to stop smoking. Drug Alcohol Depend 2013 Feb 1;128(1-2):15-19. [doi: 10.1016/j.drugalcdep.2012.07.012] [Medline: 22943961]
44. Office for National Statistics. UK Data Archive Study Number 6528 - ONS Opinions Survey, Internet Access Module, January, February and March, 2009: Technical Report. 2009. URL: http://www.esds.ac.uk/doc/6529/mrdoc/pdf/6529userguide.pdf [accessed 2013-02-20] [WebCite Cache ID 6Ea55WKkj]
45. 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]
46. Kotz D, Fidler J, West R. Factors associated with the use of aids to cessation in English smokers. Addiction 2009 Aug;104(8):1403-1410. [doi: 10.1111/j.1360-0443.2009.02639.x] [Medline: 19549267]
47. Shiffman S, Di Marino ME, Sweeney CT. Characteristics of selectors of nicotine replacement therapy. Tob Control 2005 Oct;14(5):346-355 [FREE Full text] [doi: 10.1136/tc.2004.009183] [Medline: 16183986]
48. Shiffman S, Brockwell SE, Pillitteri JL, Gitchell JG. Use of smoking-cessation treatments in the United States. Am J Prev Med 2008 Feb;34(2):102-111. [doi: 10.1016/j.amepre.2007.09.033] [Medline: 18201639]
49. Kotz D, West R. Explaining the social gradient in smoking cessation: it's not in the trying, but in the succeeding. Tob Control 2009 Feb;18(1):43-46 [FREE Full text] [doi: 10.1136/tc.2008.025981] [Medline: 18936053]
50. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
51. Crutzen R. From eHealth technologies to interventions. J Med Internet Res 2012;14(3):e93 [FREE Full text] [doi: 10.2196/jmir.2050] [Medline: 22738742]
52. Berg CJ, An LC, Kirch M, Guo H, Thomas JL, Patten CA, et al. Failure to report attempts to quit smoking. Addict Behav 2010 Oct;35(10):900-904. [doi: 10.1016/j.addbeh.2010.06.009] [Medline: 20584571]

Abbreviations

ACORN: A Classification of Residential Neighbourhoods
HSI: heaviness of smoking index
ISCI: Internet-based smoking cessation interventions
RE-AIM: reach, efficacy, adoption, implementation, and maintenance

Edited by G Eysenbach; submitted 06.09.12; peer-reviewed by R Crutzen, E Murray; comments to author 09.11.12; revised version received 16.11.12; accepted 22.11.12; published 18.03.13.

Please cite as:
Brown J, Michie S, Raupach T, West R
Prevalence and Characteristics of Smokers Interested in Internet-Based Smoking Cessation Interventions: Cross-sectional Findings From a National Household Survey
J Med Internet Res 2013;15(3):e50
URL: http://www.jmir.org/2013/3/e50/
doi:10.2196/jmir.2342
PMID:23506944

Original Paper

Cost-Effectiveness and Cost-Utility of Internet-Based Computer Tailoring for Smoking Cessation

Eline Suzanne Smit[1], PhD; Silvia MAA Evers[2], PhD; Hein de Vries[1], PhD; Ciska Hoving[1], PhD

[1]CAPHRI School for Public Health and Primary Care, Department of Health Promotion, Maastricht University, Maastricht, Netherlands

[2]CAPHRI School for Public Health and Primary Care, Department of Health Services Research, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Eline Suzanne Smit, PhD
CAPHRI School for Public Health and Primary Care
Department of Health Promotion
Maastricht University
PO BOX 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882397
Fax: 31 433671032
Email: es.smit@maastrichtuniversity.nl

Abstract

Background: Although effective smoking cessation interventions exist, information is limited about their cost-effectiveness and cost-utility.

Objective: To assess the cost-effectiveness and cost-utility of an Internet-based multiple computer-tailored smoking cessation program and tailored counseling by practice nurses working in Dutch general practices compared with an Internet-based multiple computer-tailored program only and care as usual.

Methods: The economic evaluation was embedded in a randomized controlled trial, for which 91 practice nurses recruited 414 eligible smokers. Smokers were randomized to receive multiple tailoring and counseling (n=163), multiple tailoring only (n=132), or usual care (n=119). Self-reported cost and quality of life were assessed during a 12-month follow-up period. Prolonged abstinence and 24-hour and 7-day point prevalence abstinence were assessed at 12-month follow-up. The trial-based economic evaluation was conducted from a societal perspective. Uncertainty was accounted for by bootstrapping (1000 times) and sensitivity analyses.

Results: No significant differences were found between the intervention arms with regard to baseline characteristics or effects on abstinence, quality of life, and addiction level. However, participants in the multiple tailoring and counseling group reported significantly more annual health care–related costs than participants in the usual care group. Cost-effectiveness analysis, using prolonged abstinence as the outcome measure, showed that the mere multiple computer-tailored program had the highest probability of being cost-effective. Compared with usual care, in this group €5100 had to be paid for each additional abstinent participant. With regard to cost-utility analyses, using quality of life as the outcome measure, usual care was probably most efficient.

Conclusions: To our knowledge, this was the first study to determine the cost-effectiveness and cost-utility of an Internet-based smoking cessation program with and without counseling by a practice nurse. Although the Internet-based multiple computer-tailored program seemed to be the most cost-effective treatment, the cost-utility was probably highest for care as usual. However, to ease the interpretation of cost-effectiveness results, future research should aim at identifying an acceptable cutoff point for the willingness to pay per abstinent participant.

(J Med Internet Res 2013;15(3):e57) doi:10.2196/jmir.2059

KEYWORDS

randomized controlled trial; economic evaluation; smoking cessation; Internet; computer-tailoring; general practice

Introduction

Background

Smoking is the single most preventable cause of illness and premature death in the world and is an important risk factor for 6 of the 8 leading causes of death, including several types of cancer, cardiovascular diseases, and respiratory diseases [1]. Consequently, smoking is related to a reduced quality of life and places a burden of €4 to €7 billion on health care [2-4]. Thus, quitting smoking is important, not only to improve individual and population health, but also to reduce smoking-related health care costs.

Extensive evidence exists on the clinical effectiveness of behavioral interventions for smoking cessation [5-7]. Brief advice from a general practitioner is one of these effective smoking cessation interventions [8]. However, general practitioners and practice nurses often report a lack of time and skills to provide their patients with elaborate smoking cessation advice [9,10]. Another behavioral intervention that has proven to be effective in increasing smoking cessation rates up to 13 months is computer tailoring [5,7,11-16]. Computer tailoring entails the adaption of the content of an intervention to participants' individual characteristics by using computer programs [17]. Most often, a questionnaire is used as a screening instrument [12,18,19]. The answers smokers provided on the questions in this screening instrument are accumulated into a large data file and are subsequently matched with relevant feedback messages that are ultimately combined into a tailored feedback letter. Tailored interventions are more effective in attracting and keeping the smoker's attention [17,20], resulting in better processing of information [21]. A single tailored feedback message is successful in increasing cessation rates [15], but dynamically tailored feedback provided on multiple occasions can be even more effective [11,12,22]. Due to the automatic generation of the tailored feedback and the fact that computer-tailored interventions are increasingly delivered online [23,24], the integration of an Internet-based computer-tailored program in the general practice setting might limit the burden on health professionals and patients, reduce facility and administrative costs, and could potentially be time- and cost-saving. As a combination of effective interventions was expected to achieve higher abstinence rates than either of the 2 alone [25], our research team has developed a smoking cessation intervention consisting of Internet-based multiple computer tailoring and a single tailored counseling session by a practice nurse.

Despite the proven clinical effectiveness of behavioral smoking cessation interventions, information about their relative cost-effectiveness is limited. Previously, several cost-effective smoking cessation interventions have been developed [26-28]. However, the interventions studied all involved the use of smoking cessation medication and investigated the cost-effectiveness of referrals to intensive counseling combined with pharmacotherapy [27], reimbursement of smoking cessation support [28], or the smoking cessation drug varenicline [26]. With regard to behavioral smoking cessation interventions, a computer-based smoking cessation intervention for primary care professionals was successful in increasing abstinence rates and quality-adjusted life years (QALYs) among patients [29]. In addition, in response to a call in the *Journal of Medical Internet Research* for research to economically evaluate eHealth interventions [30], cost-effectiveness and cost-utility studies of Internet-based interventions aimed at alcohol reduction [31] and depressive symptom treatment [32,33] have been initiated. However, to our current knowledge, no such studies have yet been conducted concerning the cost-effectiveness and cost-utility of an Internet-based smoking cessation intervention.

Therefore, the objective of the present study was to compare the cost-effectiveness and cost-utility of (1) an Internet-based multiple computer-tailored smoking cessation program combined with a single tailored counseling session by a practice nurses, (2) only an Internet-based multiple computer-tailored smoking cessation program, and (3) care as usual, defined as practice nurses' standard care regarding smoking cessation.

Methods

Design

Economic evaluation studies aim to determine the costs and effects associated with an intervention and to compare these with the costs and effects of other interventions and/or current practice [34]. They usually consist of 5 steps [35], which are listed in Textbox 1. For a more extensive and detailed description of each of these steps, refer to Multimedia Appendix 1.

Textbox 1. The 5 steps in economic evaluation studies.

Step 1. Identification of relevant costs and effects

Based on a chosen perspective (eg, the health care perspective, health insurer perspective, or societal perspective), relevant costs and effects are identified.

Step 2. Measurement of costs and effects

Costs can be assessed prospectively by means of cost diaries, or retrospectively using cost questionnaires. Effects are usually assessed in terms of quality of life.

Step 3. Valuation of measured costs and effects

Health care and patient costs are usually valued in a monetary currency using manuals for cost analysis in health care research. Effects on quality of life are usually valued in QALYs gained or lost.

Step 4. Calculation of a cost-effectiveness ratio

When comparing two interventions, an incremental cost-effectiveness ratio (ICER) can be calculated: $ICER=(C_i–C_c)/(E_i–E_c)$. When comparing more than 2 interventions, a net monetary benefit (NMB) should be calculated using the willingness to pay (WTP): $NMB=(E_i–E_c)\times WTP–(C_i–C_c)$. A description of both formulas can be found in Multimedia Appendix 1.

Step 5. Uncertainty analysis

To deal with the sampling uncertainty bootstrap analyses can be used, whereas a sensitivity analysis can be conducted to deal with uncertainty due to the assumptions made.

The present economic evaluation study was trial-based, embedded in a randomized controlled trial (RCT) that tested the effectiveness of an Internet-based multiple computer-tailored smoking cessation program and tailoring counseling by practice nurses. This 3-armed RCT was conducted among Dutch adult smokers and had a follow-up period of 12 months. From May 2009 to June 2010, 91 practice nurses working in different Dutch general practices throughout the Netherlands recruited smoking patients for participation in the RCT. To aid recruitment, several recruitment materials were provided (eg, desk displays, posters, and business cards). Smokers interested in participation could sign up for the study on the study website. There, information was provided about the objectives of the study, the randomization procedure, and the incentive provided when respondents completed all questionnaires (ie, a €10 gift voucher). When signing up, participants were able to choose their own username and password and were informed that no one but the research team would be able to retrieve these passwords. After providing informed consent, participants were randomized into 1 of the 2 intervention groups (multiple tailoring and counseling or multiple tailoring only) or into the usual care control group. Randomization took place at the participant level by means of a computer software randomization device.

The trial design was approved by the Medical Ethics Committee of Maastricht University and the University Hospital Maastricht (MEC 08-3-037; NL22692.068.08), and is registered with the Dutch Trial Register (NTR1351). A more detailed description of the study design has been published elsewhere [25].

Participants

Participants were eligible for participation if they smoked, were motivated to quit within 6 months, were 18 years or older, and were able to read and understand Dutch sufficiently to read study materials and participate in the trial. Moreover, they had to have access to the Internet. This resulted in a total of 414 eligible smokers.

The Interventions

Figure 1 presents an overview of the intervention components in each of the study groups.

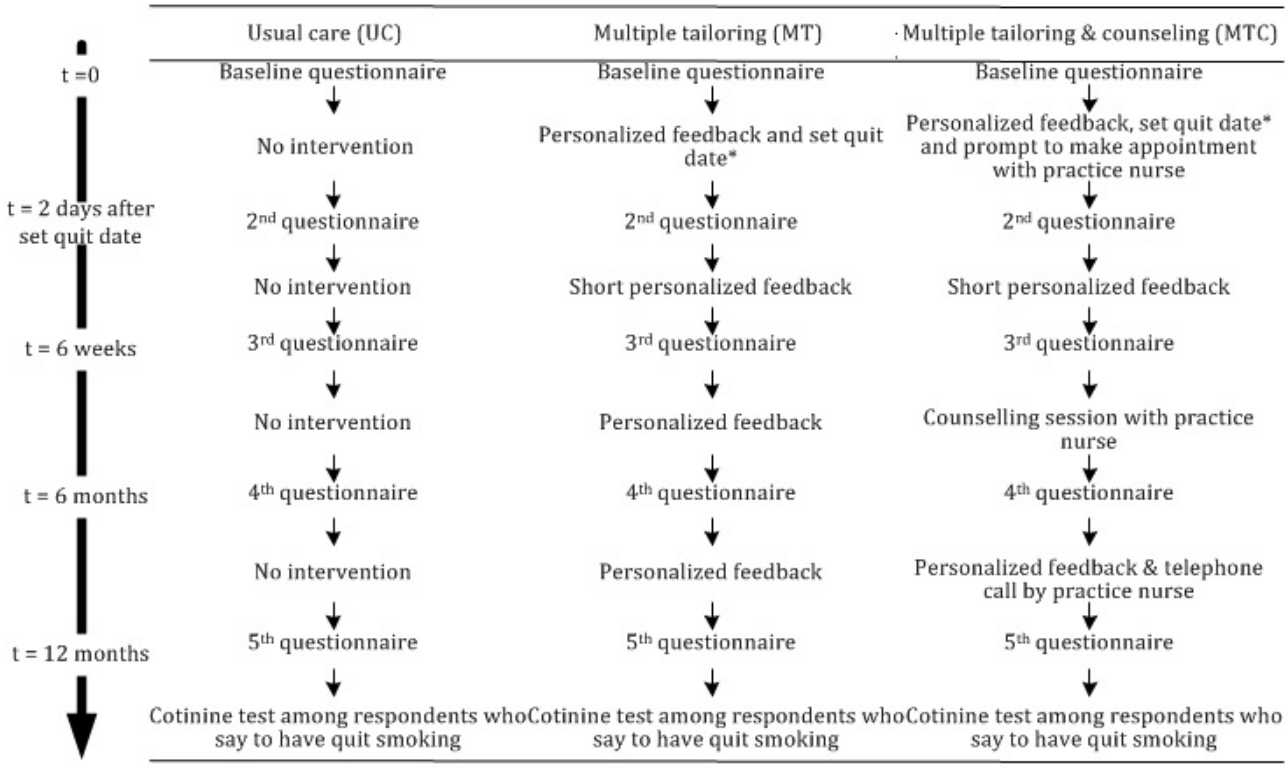
The Internet-based multiple computer-tailored smoking cessation program was based on a previously developed single computer-tailored intervention [12,15] for which the I-Change model (ICM) formed the theoretical framework [36]. As was its predecessor, the attitude-social influence-efficacy model [37], the ICM is a theory of behavioral change which incorporates theoretical concepts from several sociocognitive models, including the transtheoretical model [38], the theory of planned behavior [39], social cognitive theory [40], and the health belief model [41]. According to the ICM, the most proximal predictor of behavior is the intention to perform this behavior. Intention is predicted by 3 motivational constructs, attitude, perceived social influence, and self-efficacy, which in turn can be predicted by several premotivational factors, such as awareness, previous experience with the same and related behaviors, biological factors, and sociocultural factors. To overcome barriers that increase the well-known gap between intention and behavior (eg, [42]), the ICM proposes ability factors, such as an individual's skills to refrain from smoking, and the formation of action plans. The ICM has been used successfully to develop several other effective computer-tailored programs [12,15,18]. While filling out the first online questionnaire (ie, baseline questionnaire), all participants were asked to set a date within the next 4 weeks at which they would attempt to quit smoking. They received a total of 4 feedback letters: at baseline, 2 days after the quit date they had set for themselves at baseline, after 6 weeks, and after 6 months. Feedback was personalized and tailored to several participant characteristics: gender, attitude, social influence and self-efficacy, intention to quit smoking, action planning, and smoking behavior. Feedback letters were iterative: the second, third, and fourth feedback letters did not only concern the participant's present state, but also referred to changes participants had made since they were included in the program.

Most feedback letters consisted of 4 to 5 pages and 7 components: (1) introduction, including specific feedback on the respondent's smoking behavior and on his/her intention to quit smoking and to maintain nonsmoking; (2) feedback on the respondent's attitude, including perceived advantages (pros) and disadvantages (cons) about smoking and quitting smoking; (3) feedback on perceived social influence (not) to smoke; (4) feedback on the respondent's reported self-efficacy to refrain from smoking in specific situations, including suggestions on how to cope with these situations; (5) feedback on the extent to which respondents were planning to undertake specific actions (action plans) while preparing their quit attempt; (6) feedback on how to cope with situations in which it might be difficult not to smoke (coping plans), including the formulation of personal plans in the shape of if-then statements [28]; and (7) ending. Participants could access their feedback letters directly online after questionnaire completion. Additionally, feedback letters were sent to the participant by email. In both cases, feedback letters could be printed. An example of a tailored feedback message is provided in Multimedia Appendix 2.

After receiving the first tailored feedback, participants in the multiple tailoring and counseling group were prompted to schedule a counseling meeting with their practice nurse within 6 to 8 weeks. They received this counseling session instead of the third tailored feedback letter at the 6-week follow-up. A counseling protocol was provided to assist practice nurses in guiding these counseling sessions. This protocol consisted of 3 chapters guiding on 3 different types of participants: smokers who had quit successfully, smokers who had quit but relapsed, and smokers who had not yet quit. The content of the counseling session was developed to be as similar as possible to the content of the computer-tailored feedback and was also tailored to the participant characteristics mentioned previously. After 6 months, practice nurses were instructed to call their patients to ask them about their progress toward permanent cessation and, if needed, to provide them with additional cessation support.

Participants randomized in the usual care group received smoking cessation guidance according to participating practice nurses' standard practice, which can vary from a brief intervention consisting of a single recommendation to stop smoking to more intensive interventions [43,44].

Figure 1. Overview of the intervention elements received by the 3 groups.



Measurements

Self-reported online questionnaires were used to assess both costs and effects. Questionnaires were administered at baseline and at 6-week, 6-month, and 12-month follow-ups. When follow-up questionnaires were not completed by 1 week after the invitation, an email reminder was sent. At 12-month follow-up, this email reminder was followed by a phone call to collect data.

Identification, Measurement, and Valuation of Costs

The present economic evaluation study was conducted from a societal perspective. This implies that intervention costs, health care costs, and patient costs were identified as relevant. Intervention costs consisted of all costs that could be attributed to the delivery of the intervention, such as hosting costs for the Internet-based program and costs associated with counseling sessions. Costs for the development of the intervention and research-specific costs were excluded because these costs are sunk costs, costs that would already be spent before the intervention is implemented. In total, intervention costs were

€57.70 per participant in the multiple tailoring and counseling group and €7.70 per participant in the multiple tailoring group. Interventions costs in the usual care group were considered zero because no intervention materials needed to be developed for this group. Health care costs related to general practitioners' or practice nurses' (telephone) consultations or home visits (other than the counseling session which was part of the multiple tailoring and counseling intervention), inpatient and outpatient specialist care, alternative medicine, mental health care, prescribed and over-the-counter (OTC) smoking cessation medication, hospital admissions, smoking cessation aids, and other care (eg, paramedics consultations or professional home care). Patient costs consisted of traveling and time lost due to participation in the intervention. However, for primary and secondary analyses, patient costs were not valued in monetary costs, but considered as reflected in participants' reported quality of life [45].

Self-reported health care use was assessed during a 12-month follow-up period using a 3-month retrospective costing questionnaire that consisted of open-ended questions. Participants indicated whether they had received each type of care during the past 3 months, and if so, how often. The time participants spent using the online tailoring program was tracked by computer-registered log-in and log-out data. To assess time spent on counseling, we used a mean time of 20 minutes for face-to-face counseling sessions and an average of 10 minutes for telephone consultations. Traveling time was measured based on average travel distances to a general practitioner in the Netherlands [45].

To valuate health care usage and patient costs, the updated Dutch manual for cost analysis in health care research was used [45]. In general, standardized prices were used, but when no standardized prices were available, real costs or tariffs were used to estimate costs. In case of uncertainty, we used the lowest price. Costs of smoking cessation medication were calculated based on daily-defined dosage [46], including 6% Value Added Tax, prescription charges for prescribed medication, and clawback, a lawful discount percentage to be subtracted from medication prices by pharmacists [46]. Prices of informal care were based on shadow prices for unpaid work. The participants' time spent on the program was valued by using the friction cost approach [45]. The index year used was 2011. Because prices in the Dutch manual for cost analysis in health care research [45] were from the year 2009, these prices were indexed to the year 2011. The consumer price indexes used were 105.38 for 2009 and 109.02 for 2011 [47]. A 13-month recruitment period and a 12-month follow-up period can both be considered as relatively short; therefore, it is unlikely that any substantial differences in health care consumption and effects existed between participants who were included at the beginning and toward the end of the recruitment period. As a result, there were no reasons to discount volumes of health care consumption or effects.

Identification, Measurement, and Valuation of Effects

The primary outcome measure used in the cost-effectiveness analysis (CEA) was prolonged abstinence measured at 12-month follow-up. This was assessed by 1 item asking whether the participant had refrained from smoking since the previous measurement at 6-month follow-up (ie, abstinence for at least 6 months; 1=no, 2=yes). Secondary outcome measures were 7-day point prevalence abstinence assessed by 1 item asking whether the participant had refrained from smoking during the past 7 days (1=no, 2=yes) and addiction level measured by the abbreviated Fagerström Test for Nicotine Dependence (FTND; 0=not addicted, 10=highly addicted) [48]. Self-reported abstinence at 12-month follow-up was cotinine validated using a saliva swab test [49]. Smoking abstinence was expressed in a probability score that a smoker would be abstinent (1=not abstinent, 2=abstinent).

The primary outcomes measure for the cost-utility analysis (CUA) was quality of life, measured in terms of QALYs. The valuation of effects on quality of life implies that utility scores need to be computed. In the present study, utilities were measured by the EuroQol EQ-5D [34,50], which is the current recommended measure for assessing quality of life by the National Institute for Health and Clinical Excellence [51] and has been used in other evaluations of smoking cessation programs [52,53] and Internet-based interventions aimed to reduce other health-related problems [32]. The EQ-5D consists of 5 health state dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression) on which participants have to indicate their own health state (1=no complaints, 2=some complaints, 3=many complaints) [50]. Utility scores assessed at different points in time were transformed into an overall QALY score using the area under the curve method. The area under the curve is the duration of the health state (on the x-axis; in our case 12 months/1 year) multiplied by the quality weight for the health state (on the y-axis; utility scores). The resulting QALY score represents the number of QALYs gained or lost during the 12-month follow-up period [34]. For example, gaining 1 QALY means that 1 year is gained in perfect health, whereas gaining 0.8 QALY means that 1 year is gained in less than perfect health (utility score of 0.8) or that 0.8 years are gained in perfect health (utility score of 1).

Analyses

All analyses were conducted according to the intention-to-treat principle. Missing data for costs, EQ-5D items, overall tobacco consumption, and addiction level were replaced by mean imputation by using participants' scores on the previous and next measurement. When mean imputation was not possible because of missing data on multiple measurement points, missing data were replaced using the last observation carried forward (preferred choice) or next observation carried backward method. Missing data for smoking abstinence were replaced using a negative scenario; participants lost to follow-up were considered still smoking.

Baseline Comparability of the Three Study Groups

To investigate the comparability of the 3 groups with regard to demographics, baseline values of outcomes, and health care–related costs over the past 3 months, 1-way analyses of variance (ANOVA) with Tukey post hoc tests and chi-square tests were conducted. To determine whether selective dropout had occurred, a comparison was made between those lost to

follow-up and those who remained in the study after 12 months using 2-sided *t* tests and chi-square tests.

Annual Costs and Effects

The 3 groups were compared with regard to their mean annual costs using nonparametric bootstrapping (5000 times) with 95% confidence intervals in percentiles [34]. To compare the 3 groups with regard to the mean effect assessed 12 months after baseline, 1-way ANOVA with Tukey post hoc tests and chi-square tests were conducted.

Cost-Effectiveness and Cost-Utility Analyses

First, incremental costs and effects were calculated for each of the 3 treatments studied. Subsequently, an incremental cost-effectiveness ratio (ICER) was calculated to compare costs and effects between pairs of study groups according to the following formula: ICER=$(C_i–C_c)/(E_i–E_c)$. In this formula, C_i is the adjusted annual costs of the intervention group, C_c is the adjusted annual costs of the control group, E_i is the adjusted effects for the intervention group, and E_c is the adjusted effects of the control group. With regard to quality of life, ICERs are often called incremental cost-utility ratios (ICURs). However, because ratios such as the ICER or ICUR can compare only 2 groups, to compare the 3 groups in our study a net monetary benefit (NMB) was calculated for each of 3 treatments. The NMB can be calculated by valuing the effectiveness and utility outcomes in monetary values using a threshold for society's willingness to pay (WTP) per abstinent participant and per QALY gained [54] according to the following formula: $(E_i–E_c)×WTP–(C_i–C_c)$. In the present study, we used a WTP of €18,000 because this is an accepted Dutch cutoff point for the WTP per QALY [55], to calculate the likelihood that each treatment would be most likely highest in cost-effectiveness/cost-utility.

Uncertainty Analysis

Sampling uncertainty around the estimates of cost-effectiveness and cost-utility was taken into account by using nonparametric bootstrap resampling techniques [34,56]. Using bootstrapping techniques with replacement n (often 1000) times, a random sample is drawn from the original dataset resulting in 1000 slightly different samples and thus slightly different ICERs. Of these 1000 ICERs, the percentage can be calculated with (1) more effects and lower costs (dominant), (2) with less effects and lower costs, (3) with more effects and higher costs, and (4) with less effects and higher costs (inferior).

To deal with the uncertainty of parameter estimates from the primary analyses, a sensitivity analysis was conducted. As described earlier, in primary analyses patient costs (ie, traveling and time costs) were not valued in monetary costs but considered as reflected in participants' reported quality of life [45]. However, because patient costs can be considered directly related to the treatment received [45], it could be argued that these costs should be included as part of the program cost. Therefore, we tested whether an increase in program costs as a result of patient costs' monetary valuation would lead to a change in results. For the multiple tailoring and counseling group, the inclusion of patient costs meant an increase in program costs from €57.70 to €141.89 per participant; for the multiple tailoring group, this meant an increase in program costs from €7.70 to €82.24 per participant.

Bootstrap analyses were conducted using Microsoft Office Excel 2003. All other analyses were conducted using SPSS 17.0 (SPSS Inc, Chicago, IL, USA).

Results

Sample Characteristics

Of the 414 participants who were eligible for participation, 163 were randomized into the multiple tailoring and counseling group, 132 into the multiple tailoring group, and 119 into the usual care group. No baseline differences were found between the 3 groups (Table 1). After 12 months, 231 (55.8%) of the 414 participants could be followed up. Participants lost to follow-up were significantly younger than those who remained in the study (P=.01). Of the 46 participants who reported prolonged abstinence at the 12-month follow-up, 30 successfully completed a cotinine test and all cases of self-reported abstinence were confirmed. After imputation of missing values, total cost data were available for 370 participants (89.2%) whereas effect data were available for 414 (100%) participants for abstinence measures, 409 (98.8%) for addiction level, and 384 (92.8%) for QALYs.

Annual Costs and Effects

A significant difference was found between the multiple tailoring and counseling and the usual care groups with regard to annual health care–related costs per patient, with significantly higher costs in the multiple tailoring and counseling group (Table 2). No differences were found between the 3 groups with regard to any specific type of health care–related costs (Table 2), nor regarding effects on abstinence, QALYs, or addiction level assessed at 12-month follow-up (Table 3).

Table 1. Comparability of the 3 groups, multiple tailoring and counseling (MTC), multiple tailoring (MT), and usual care (UC), regarding demographics, baseline values of outcomes, and health care–related costs over the past 3 months (N=414).

Variable	MTC (n=163)	MT (n=132)	UC (n=119)	*F* (df)	χ^2 (df)	*P*
Age, mean (SD)	48.1 (12.0)	47.8 (12.5)	48.1 (11.3)	0.03 (2,406)		.97
Male, n (%)	60 (36.8)	41.2 (54)	42.9 (51)		1.2 (2)	.56
Educational level, n (%)					1.9 (2)	.76
High	39 (23.9)	30 (22.9)	24 (20.2)			
Medium	68 (41.7)	63 (48.1)	56 (47.1)			
Low	56 (34.4)	38 (29.0)	39 (32.8)			
Chronic diseases						
Cardiovascular diseases, n (%)	27 (16.6)	17 (13.0)	18 (15.1)		0.7 (2)	.69
Respiratory diseases, n (%)	38 (23.3)	44 (33.6)	36 (30.3)		4.0 (2)	.14
Diabetes, n (%)	8 (4.9)	6 (4.6)	7 (5.9)		0.2 (2)	.89
Cancer, n (%)	10 (6.1)	12 (9.2)	8 (6.7)		1.1 (2)	.59
Cigarettes smoked per day, mean (SD)	20.6 (10.3)	23.5 (23.2)	21.5 (15.5)	1.14 (2,411)		.32
FTND[a] score (range 0-10), mean (SD)	5.3 (2.2)	5.6 (2.0)	5.3 (2.1)	0.94 (2,406)		.39
Utility, mean (SD)[b]	0.8 (0.2)	0.8 (0.2)	0.8 (0.2)	0.60 (2,375)		.55
Health care–related costs (€), [c] mean (SD)	425.9 (1506.9)	286.9 (436.6)	236.9 (474.0)	1.19 (2,369)		.31
General practitioner	53.2 (50.2)	61.0 (73.2)	49.7 (55.1)	1.11 (2,380)		.33
Medical specialist	65.4 (132.1)	78.6 (170.9)	87.6 (202.8)	0.57 (2,373)		.57
Hospital	206.9 (1371.3)	50.9 (205.2)	47.9 (230.8)	1.47 (2,378)		.23
Alternative healer	5.1 (25.1)	9.4 (41.6)	4.9 (24.2)	0.84 (2,379)		.44
Mental health care	30.3 (133.5)	24.3 (100.8)	38.2 (186.5)	0.27 (2,380)		.76
Prescribed and OTC medication	26.7 (78.8)	36.4 (95.5)	13.6 (58.4)	2.37 (2,381)		.10
Medical aids and assistive devices	1.7 (4.1)	3.4 (12.6)	1.9 (4.8)	1.73 (2,380)		.18
Other care	26.0 (268.3)	19.2 (166.4)	12.5 (65.7)	0.15 (2,379)		.86

[a] Fagerström Test for Nicotine Dependence (0=not addicted, 10=highly addicted)

[b] Based on the Dutch algorithm for the EQ-5D scores.

[c] Costs for prior 3 months.

Table 2. Mean annual costs[a] per participant in the MTC, MT, and UC groups.

Cost type	Costs per group (€) mean (SD)[b]			95% CI[b]		
	MTC	MT	UC	MTC–MT	UC–MT	MTC–UC
Fixed costs						
Intervention costs (n=384)	57.70	7.70	0			
Health care–related costs						
General practitioner (n=384)	157 (14)	180 (27)	139 (17)	–86.1 to 32.3	–105.8 to 15.5	–25.4 to 61.4
Medical specialist (n=374)	298 (52)	251 (62)	224 (48)	–115.8 to 198.1	–188.3 to 116.6	–61.2 to 213.8
Hospital (n=380)	610 (288)	267 (106)	172 (84)	–139.7 to 1054.7	–374.0 to 161.1	–17.0 to 1133.4
Alternative healer (n=382)	17 (6)	29 (13)	18 (9)	–42.9 to 13.7	–43.1 to 18.6	–23.4 to 18.8
Mental health care (n=384)	106 (39)	95 (34)	131 (71)	–92.2 to 109.8	–97.1 to 209.4	–200.2 to 111.9
Prescribed and OTC smoking cessation medication (n=384)	148 (24)	144 (30)	90 (23)	–72.7 to 79.1	–129.6 to 18.5	–9.0 to 124.6
Smoking cessation aids (n=384)	20 (10)	15 (10)	19 (14)	–21.6 to 32.5	–27.9 to 41.0	–34.4 to 32.3
Other care (n=382)	122 (87)	21 (12)	41 (22)	–15.9 to 293.6	–24.7 to 72.4	–45.9 to 281.7
Overall health care–related costs (n=370)	1564 (338)	1016 (158)	761 (122)	–95.4 to 1381.4	–642.2 to 139.1	194.3-1611.8

[a] Volumes and price details are available upon request.

[b] Based on 5000 bootstrap replications.

Table 3. Mean annual effect on smoking abstinence, QALY, and addiction level in the MTC, MT, and UC groups.

Effects	MTC	MT	UC	*F* (df)	χ^2 (df)	*P*
Prolonged abstinent (n=414), n (%)	14 (8.6)	20 (15.2)	12 (10.1)		3.4 (2)	.19
QALY (EQ-5D)[a] (n=384), mean (SD)	0.86 (0.15)	0.83 (0.21)	0.84 (0.21)	0.89 (2,381)		.41
7 days abstinent (n=414), n (%)	20 (12.3)	27 (20.5)	15 (12.6)		4.6 (2)	.10
FTND[b] score (n=409), mean (SD)	4.76 (2.41)	5.21 (2.30)	4.81 (2.46)	1.40 (2,406)		.25

[a] Based on the Dutch algorithm for the EQ-5D scores.

[b] Fagerström Test for Nicotine Dependence (0=not addicted, 10=highly addicted); reversed range.

Cost-Effectiveness Analyses

Table 4 shows that for participants in the multiple tailoring and counseling group costs were higher, whereas effects were lower than in the usual care and multiple tailoring groups; thus, multiple tailoring and counseling was dominated by the other 2 treatments. Comparing multiple tailoring with usual care showed that multiple tailoring resulted in more costs, but also in more effects. Compared with usual care, €5100 had to be paid within the multiple tailoring group for each additional abstinent participant (Table 4).

Table 4. Incremental costs and effects per abstinent smoker and per QALY gained for the MTC, MT, and UC groups with a willingness-to-pay threshold of €18,000.

Intervention	Incremental costs (€)	Incremental probability[a]	Incremental costs[b] (€)
Prolonged abstinence [c]			
UC			
MT vs UC	255	.05	5100
MTC vs UC	806	–.02	Dominated[d]
MTC vs MT	551	–.07	Dominated[e]
QALY (EQ-5D) [f]			
UC			
MT vs UC	255	–.01	Dominated[g]
MTC vs UC	806	.02	40,300
MTC vs MT	551	.03	18,367

[a] Probability of being abstinent/gaining 1 QALY.

[b] Per abstinent participant or per QALY; calculated according to the formula ICER/ICUR=$(C_i–C_c)/(E_i–E_c)$; additional information available in Multimedia Appendix 1.

[c] Coded as 1=not abstinent and 2=abstinent.

[d] ICER=–40.300.

[e] ICER=–7.871.

[f] Based on the Dutch algorithm for the EQ-5D scores.

[g] ICUR=–25.500.

The CEA showed that until a threshold value for the WTP of €5100 per abstinent participant, usual care was probably the most efficient treatment. However, from a WTP of €5100 or higher, multiple tailoring was probably most cost-effective (Table 5). With the accepted Dutch cutoff point of €18,000 per QALY for preventive interventions [57], multiple tailoring would be the preferable treatment. These results are visually displayed in the cost-effectiveness acceptability curve (CEAC), showing the probability of each treatment being preferable to the other 2 treatments for varying levels of the WTP per additional abstinent participant (Figure 2). Sensitivity analyses supported these results (Table 5).

Results from secondary analyses showed that with 7-day point prevalence abstinence, a high probability was found (ie, 88%, with a WTP of €18,000 per abstinent participant) that multiple tailoring was the most cost-effective treatment. Regarding the level of addiction, however, it was most probable that multiple tailoring would be least efficient (Table 5).

Cost-Utility Analyses

With regard to QALYs gained, Table 4 shows that multiple tailoring was dominated by usual care because this treatment was both more expensive and less effective. Furthermore, cost-utility analyses showed that multiple tailoring and counseling was more expensive, but also more effective than usual care and multiple tailoring in increasing the number of QALYs gained. This resulted in an incremental cost of €40,300 per QALY gained when comparing multiple tailoring and counseling with usual care, and in an incremental cost of €18,367 per QALY when comparing multiple tailoring and counseling to multiple tailoring.

With a WTP of €18,000 per abstinent participant, the CUA showed that usual care would probably (ie, 64%) be the most efficient treatment (Table 5). Although decreasing this threshold value to €0 led to an increased probability that usual care would be most efficient, increasing this threshold led to a lower probability of usual care being most preferable. With a WTP of almost €40,000, usual care and multiple tailoring and counseling would be equally preferable. These results are further illustrated in the cost-utility acceptability curve (CUAC) (Figure 3). Sensitivity analyses showed similar results (Table 5).

Table 5. Results from cost-effectiveness and cost-utility analyses based on 1000 bootstrap replications.

Type of analysis	Group, n			Probability of highest net monetary benefit[a], %		
	MTC	MT	UC	MTC	MT	UC
Primary analysis						
Prolonged abstinence[f]	145	121	104	0	78	21
QALY (EQ-5D)[b]	145	121	104	18	18	64
Secondary analysis						
7-day ppa[c,f]	145	121	104	1	88	11
FTND score[d]	135	115	96	50	6	45
Sensitivity analysis [e]						
Prolonged abstinence[f]	145	121	104	1	76	24
QALY (EQ-5D)[b]	145	121	104	19	15	66

[a]With a willingness-to-pay threshold of €18,000.

[b]Based on the Dutch algorithm for the EQ-5D scores.

[c]ppa: point prevalence abstinence.

[d]FTND: Fagerström Test for Nicotine Dependence (0=not addicted, 10=highly addicted).

[e]Increase in program costs from €57.70 to €141.89 (MTC group) and from €7.70 to €82.24 (MT group) caused by the inclusion of patient costs.

[f]Coded as 1=not abstinent and 2=abstinent.

Figure 2. Cost-effectiveness acceptability curve for the 3 treatments studied: MTC, MT, and UC.

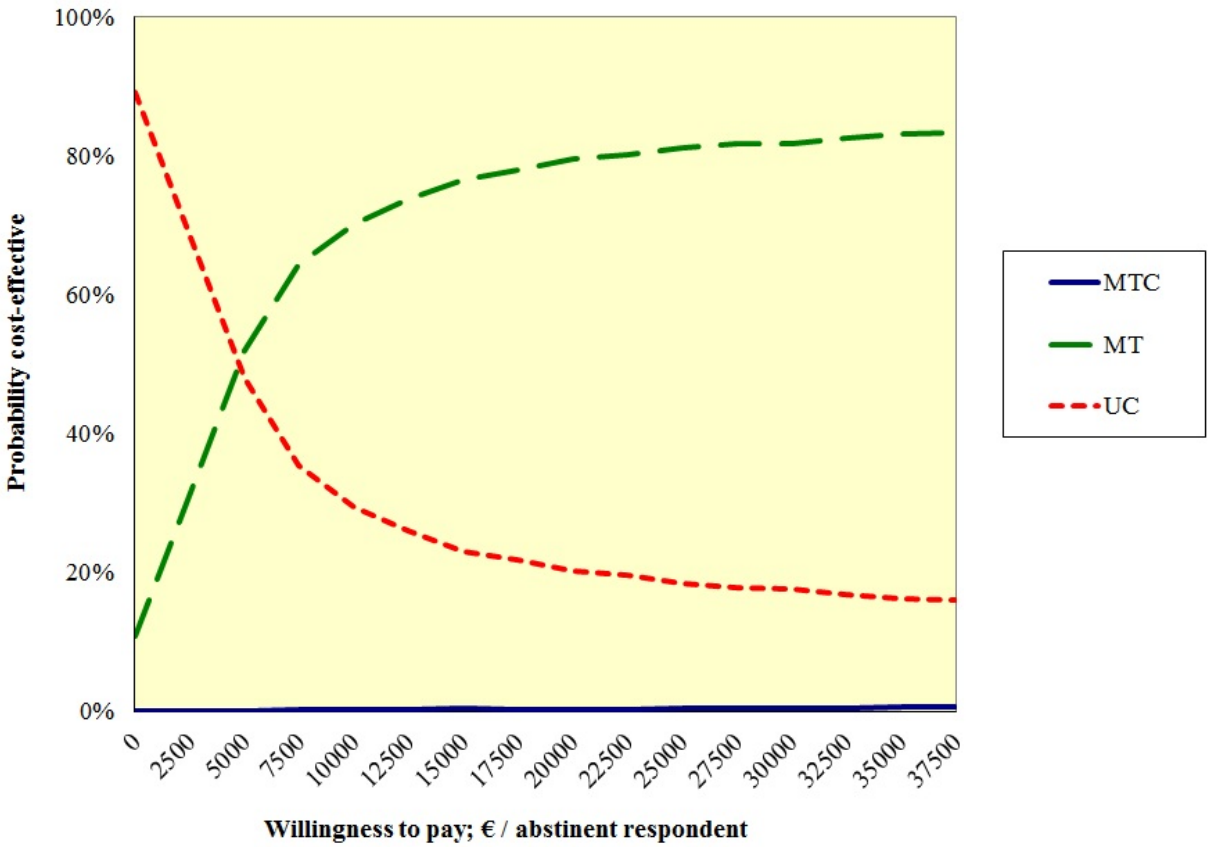
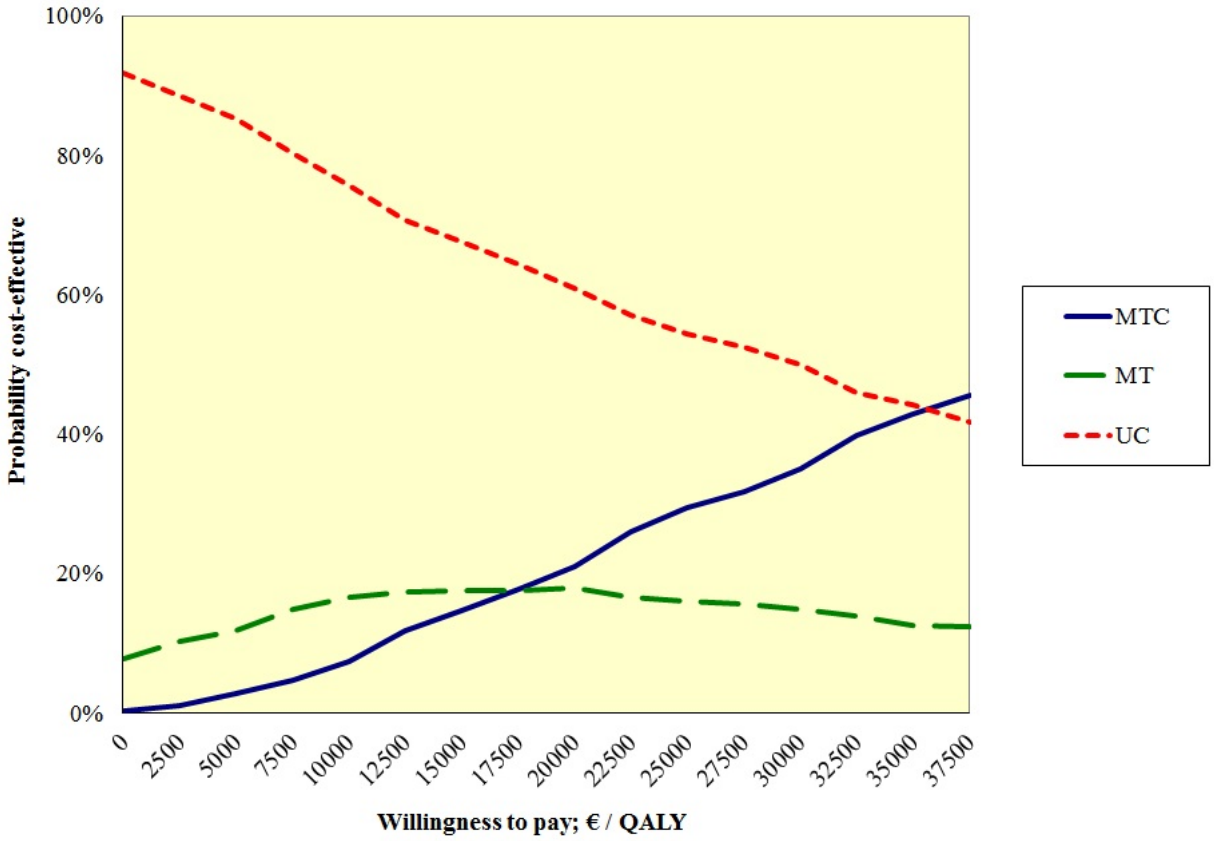


Figure 3. Cost-utility acceptability curve for the 3 treatments studied: MTC, MT, and UC.



Discussion

Main Findings

To our current knowledge, this was the first study to determine the cost-effectiveness and cost-utility of a behavioral smoking cessation intervention consisting of Internet-based computer tailoring with and without counseling by a practice nurse. The results presented suggest that participants who received the Internet-based multiple computer-tailored program and tailored counseling by their practice nurse reported significantly more annual health care-related costs than participants who received care as usual. A potential explanation for this finding might be that smokers were prompted by the tailored feedback they received to ask for more smoking cessation guidance, eg, additional counseling sessions with the practice nurse or a prescription for smoking cessation medication. Although participants who received the Internet-based program only might have had the same tendency, the practice nurses for participants in the multiple tailoring and counseling group might have been prompted by their patients' visit to offer them more smoking cessation help. Because the current smoking cessation guidelines in the Netherlands recommend more than 1 counseling session [44,58], this is not unlikely. Interestingly, though the significant difference in total health care-related costs could not be explained by a difference in any particular type of health care-related costs. For instance, no significant differences were found with regard to costs spent on general practitioner or practice nurse consultations or on the use of smoking cessation aids.

Furthermore, the present study showed that the Internet-based multiple computer-tailored smoking cessation program would probably be the most cost-effective of the 3 treatments under study. Although no similar studies yet exist within the field of smoking cessation, this finding is in-line with findings from recent studies toward the cost-effectiveness of Internet-based interventions aimed at other health-related behaviors or health problems [31,32]. Compared with current practice, the incremental costs per abstinent participant associated with the Internet-based multiple computer-tailored smoking cessation program were €5100. This is slightly more than what was found in previous studies [59,60]. However, one of these studies only included costs directly related to the interventions received [59], whereas we conducted our economic evaluation from a broader societal perspective. Nevertheless, the interpretation of the incremental costs per abstinent participant is difficult because no information exists on the amount of money that society is willing to pay per abstinent participant. Although a WTP of €18,000 per QALY is an accepted Dutch cutoff point [55], no such cutoff point exists with regard to smoking abstinence rates. To enable the interpretation of the incremental costs per abstinent participant, future research should aim at identifying an acceptable cutoff point for the WTP per abstinent participant.

Regarding cost-utilities, the results suggest that care as usual would probably be the most preferable of the treatments studied. A potential explanation for this finding might be that the follow-up period of 12 months was not sufficiently long for the beneficial effects of the intervention on smoking abstinence to be translated into detectable changes in quality of life, as recent

ex-smokers are known to suffer from withdrawal symptoms [61]. A potential solution would be to use short-term trial data as input for a model predicting the effects of smoking cessation interventions on long-term quality of life, a technique used in several recent studies [28,29,60]. Although trial data served as input for these models, several assumptions had to be made to build these models, bringing about additional uncertainty in the results presented [34, 60]. Another solution would be to lengthen the follow-up period of clinical trials to gather longer-term data on quality of life. Although this might imply an increased burden on the participant, it might be needed to establish the intervention's cost-utility in an as certain as possible way. However, as previous studies were able to detect a positive association between quitting smoking and quality of life during a 12-month follow-up period [62,63], additional explanations for this finding need to be sought. One such explanation might be that within Dutch general practices, care as usual for smoking cessation is rather intensive. Anecdotal evidence suggests that practice nurses usually offer 4 to 6 consultations as part of smoking cessation care (unpublished). Although participants in the multiple tailoring and multiple tailoring and counseling groups received 4 and 3 tailored feedback letters, respectively, a primarily Internet-based program might have been perceived as more distant and less intense than care as usual. As a result, participants in the usual care group might have established a better or stronger social bond with their practice nurse than participants in the multiple tailoring or multiple tailoring and counseling groups. This assumed social bond [64] might subsequently have resulted in positive effects on quality of life among usual care participants.

Strengths and Limitations

The present study aimed to contribute to the literature by examining the cost-effectiveness and cost-utility of an Internet-based smoking cessation intervention, something that has not been done before to our knowledge. In the present study, in addition to generic quality of life, disease-specific effects of the intervention (ie, smoking abstinence) were taken into account. To facilitate the comparison of the cost-effectiveness of interventions targeting different diseases, effects are usually assessed in terms of quality of life. However, to compare smoking cessation interventions more specifically, disease-specific effect measures might be more informative.

Nevertheless, the present study also had its limitations. First, it suffered from relatively high dropout rates. High rates of attrition seem to be inherent to many Internet-based interventions and dropout rates of 44% are not uncommon [24,65-67]. As a consequence, however, there was not sufficient power for us to conduct a complete-case analysis as part of the sensitivity analyses. Secondly, because we expected higher attrition rates in the multiple tailoring and counseling group, slightly more participants were randomized into this intervention group at baseline. However, attrition rates appeared to be similar among the groups, resulting in a skewed distribution of participants with 163 participants in the multiple tailoring and counseling group, 132 in the multiple tailoring group, and 119 in the usual care group. Although we do not expect that this has biased our results, this finding of no selective attrition can be considered valuable in the design of future trials.

Conclusions

The Internet-based multiple computer-tailored program seemed to be the most cost-effective treatment when smoking abstinence was the outcome measure. However, the cost-utility, using quality of life as the outcome measure, was probably the highest with care as usual. To enable the interpretation of the incremental costs per abstinent participant found in the cost-effectiveness analyses, future research should aim at identifying an acceptable cutoff point for the WTP per abstinent participant.

Acknowledgments

The study was funded by the Dutch Cancer Society (UM 2007-3834). We would like to thank Thea van Asselt for the additional assistance in conducting the cost-effectiveness and cost-utility analyses.

Conflicts of Interest

Hein de Vries is scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools.

Multimedia Appendix 1

Economic evaluation studies in a nutshell.

[PDF File (Adobe PDF File), 96KB - jmir_v15i3e57_app1.pdf]

Multimedia Appendix 2

A tailored smoking cessation advice for a respondent who reported to still be smoking at six-month follow-up and whose self-efficacy to quit has decreased since baseline.

[PDF File (Adobe PDF File), 33KB - jmir_v15i3e57_app2.pdf]

Multimedia Appendix 3

CONSORT-EHEALTH checklist V1.6.2 [68].

[PDF File (Adobe PDF File), 998KB - jmir_v15i3e57_app3.pdf]

References

1. WHO Report on the Global Tobacco Epidemic, 2008: the MPOWER package. Geneva: World Health Organization; 2008. URL: http://www.who.int/tobacco/mpower/mpower_report_full_2008.pdf [accessed 2013-02-24] [WebCite Cache ID 6Eg4lp7Qh]
2. Allender S, Balakrishnan R, Scarborough P, Webster P, Rayner M. The burden of smoking-related ill health in the UK. Tob Control 2009 Aug;18(4):262-267. [doi: 10.1136/tc.2008.026294] [Medline: 19509003]
3. Muennig P, Fiscella K, Tancredi D, Franks P. The relative health burden of selected social and behavioral risk factors in the United States: implications for policy. Am J Public Health 2010 Sep;100(9):1758-1764. [doi: 10.2105/AJPH.2009.165019] [Medline: 20019300]
4. Scarborough P, Bhatnagar P, Wickramasinghe KK, Allender S, Foster C, Rayner M. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs. J Public Health (Oxf) 2011 Dec;33(4):527-535 [FREE Full text] [doi: 10.1093/pubmed/fdr033] [Medline: 21562029]
5. Lancaster T, Stead LF. Self-help interventions for smoking cessation. Cochrane Database Syst Rev 2005(3):CD001118. [doi: 10.1002/14651858.CD001118.pub2] [Medline: 16034855]
6. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2005(2):CD001292. [doi: 10.1002/14651858.CD001292.pub2] [Medline: 15846616]
7. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
8. Stead LF, Bergson G, Lancaster T. Physician advice for smoking cessation. Cochrane Database Syst Rev 2008(2):CD000165. [doi: 10.1002/14651858.CD000165.pub3] [Medline: 18425860]
9. Hall S, Vogt F, Marteau TM. A short report: survey of practice nurses' attitudes towards giving smoking cessation advice. Fam Pract 2005 Dec;22(6):614-616 [FREE Full text] [doi: 10.1093/fampra/cmi082] [Medline: 16055470]
10. Vogt F, Hall S, Marteau TM. General practitioners' and family physicians' negative beliefs and attitudes towards discussing smoking cessation with patients: a systematic review. Addiction 2005 Oct;100(10):1423-1431. [doi: 10.1111/j.1360-0443.2005.01221.x] [Medline: 16185204]
11. Borland R, Balmford J, Hunt D. The effectiveness of personally tailored computer-generated advice letters for smoking cessation. Addiction 2004 Mar;99(3):369-377. [doi: 10.1111/j.1360-0443.2003.00623.x] [Medline: 14982550]
12. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. Health Educ Res 1998 Jun;13(2):207-214 [FREE Full text] [Medline: 10181019]
13. Hoving C, Mudde AN, Dijk F, Vries HD. Effectiveness of a smoking cessation intervention in Dutch pharmacies and general practices. Health Education 2010 Jun 2010;110(1):17-29 [FREE Full text] [doi: 10.1108/09654281011008726]
14. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. Patient Educ Couns 1999 Feb;36(2):107-117. [Medline: 10223016]
15. Te Poel F, Bolman C, Reubsaet A, de Vries H. Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months. Health Educ Res 2009 Dec;24(6):930-940 [FREE Full text] [doi: 10.1093/her/cyp036] [Medline: 19574405]
16. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. J Med Internet Res 2008;10(5):e45 [FREE Full text] [doi: 10.2196/jmir.1008] [Medline: 19033146]
17. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. Patient Educ Couns 1999 Feb;36(2):99-105. [Medline: 10223015]
18. van Keulen HM, Mesters I, Brug J, Ausems M, Campbell M, Resnicow K, et al. Vitalum study design: RCT evaluating the efficacy of tailored print communication and telephone motivational interviewing on multiple health behaviors. BMC Public Health 2008;8:216 [FREE Full text] [doi: 10.1186/1471-2458-8-216] [Medline: 18565222]
19. van Stralen MM, Kok G, de Vries H, Mudde AN, Bolman C, Lechner L. The Active plus protocol: systematic development of two theory- and evidence-based tailored physical activity interventions for the over-fifties. BMC Public Health 2008;8:399 [FREE Full text] [doi: 10.1186/1471-2458-8-399] [Medline: 19055806]
20. Kreuter M. The Active plus protocol: systematic development of two theory- and evidence-based tailored physical activity interventions for the over-fifties. In: Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
21. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539 [FREE Full text] [doi: 10.1093/her/cyh014] [Medline: 15701665]
22. Dijkstra A, De Vries H, Roijackers J, van Breukelen G. Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions. Health Psychol 1998 Nov;17(6):513-519. [Medline: 9848801]

23. 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]
24. 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]
25. Smit ES, de Vries H, Hoving C. The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses. Contemp Clin Trials 2010 May;31(3):251-258. [doi: 10.1016/j.cct.2010.03.001] [Medline: 20215047]
26. Annemans L, Nackaerts K, Bartsch P, Prignot J, Marbaix S. Cost effectiveness of varenicline in Belgium, compared with bupropion, nicotine replacement therapy, brief counselling and unaided smoking cessation: a BENESCO Markov cost-effectiveness analysis. Clin Drug Investig 2009;29(10):655-665. [doi: 10.2165/11317730-000000000-00000] [Medline: 19715382]
27. Hoogendoorn M, Feenstra TL, Hoogenveen RT, Rutten-van Mölken MP. Long-term effectiveness and cost-effectiveness of smoking cessation interventions in patients with COPD. Thorax 2010 Aug;65(8):711-718. [doi: 10.1136/thx.2009.131631] [Medline: 20685746]
28. Vemer P, Rutten-van Mölken MP, Kaper J, Hoogenveen RT, van Schayck CP, Feenstra TL. If you try to stop smoking, should we pay for it? The cost-utility of reimbursing smoking cessation support in the Netherlands. Addiction 2010 Jun;105(6):1088-1097. [doi: 10.1111/j.1360-0443.2010.02901.x] [Medline: 20659063]
29. Smith MY, Cromwell J, DePue J, Spring B, Redd W, Unrod M. Determining the cost-effectiveness of a computer-based smoking cessation intervention in primary care. Manag Care 2007 Jul;16(7):48-55. [Medline: 17907712]
30. Ahern DK, Kreslake JM, Phalen JM. What is eHealth (6): perspectives on the evolution of eHealth research. J Med Internet Res 2006;8(1):e4 [FREE Full text] [doi: 10.2196/jmir.8.1.e4] [Medline: 16585029]
31. 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 [FREE Full text] [doi: 10.2196/jmir.1694] [Medline: 21840836]
32. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010;12(5):e53 [FREE Full text] [doi: 10.2196/jmir.1436] [Medline: 21169166]
33. Gerhards SA, de Graaf LE, Jacobs LE, Severens JL, Huibers MJ, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. Br J Psychiatry 2010 Apr;196(4):310-318 [FREE Full text] [doi: 10.1192/bjp.bp.109.065748] [Medline: 20357309]
34. Drummond MF, Sculpher MJ, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
35. Evers SM, Wolf C, van Heugten C. Economische evaluatie van neuropsychologische behandeling. In: Neuropsychologische Behandeling. Amsterdam: Boom Uitgeverij; 2010:115-140.
36. de Vries H, Mudde AN, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]
37. de Vries HD, Mudde AN. Predicting stage transitions for smoking cessation applying the attitude-social influence-efficacy model. Psychology & Health 1998 Apr 1998;13(2):369-385. [doi: 10.1080/08870449808406757]
38. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, editor. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco: Jossey-Bass; 1997:99-120.
39. Ajzen I. From intentions to actions: a theory of Planned Behaviour. In: Action-Control: From Cognition to Behavior. Heidelberg: Springer; 1985:11-39.
40. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
41. Janz NK, Champion VL, Strecher VJ. The health belief model. In: Glanz K, editor. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco: Jossey-Bass; 2002:45-66.
42. Armitage CJ, Conner M. Efficacy of the Theory of Planned Behaviour: a meta-analytic review. Br J Soc Psychol 2001 Dec;40(Pt 4):471-499. [Medline: 11795063]
43. Chavannes NH, Kaper J, Frijling BD, Van der Laan JR, Jansen PWM, Guerrouj S, et al. NHG-Standaard Stoppen met roken. In: NHG-standaarden voor de huisarts. Houten: Bohn Stafleu van Loghum; 2009:1434-1449.
44. Partnership Stop met Roken. Richtlijn Behandeling van Tabaksverslaving herziening 2009. Den Haag: Partnership Stop met Roken; 2009. URL: http://www.partnershipstopmetroken.nl/sup/RichtlijnTabaksverslavingActualisatie2009definitief.pdf [accessed 2013-02-26] [WebCite Cache ID 6EioNXp56]
45. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Geactualiseerde versie 2010. 2010. URL: http://www.cvz.nl/binaries/live/cvzinternet/hst_content/nl/documenten/losse-publicaties/handleiding-kostenonderzoek-2010.pdf [accessed 2012-12-20] [WebCite Cache ID 6D3RMfHct]

46. Commissie Farmacotherapautische Hulp CVZ. Farmacotherapeutisch Kompas: medisch farmaceutische voorlichting. Amstelveen: College voor zorgverzekeringen; 2011.
47. Centraal Bureau voor de Statistiek. Consumentenprijzen; prijsindex 2006. 2011. URL: http://www.statline.nl [accessed 2012-06-06] [WebCite Cache ID 68DTtHGoX]
48. 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]
49. Craig Medical Distribution Inc. NicAlert saliva nicotine test URL: http://www.craigmedical.com/Nicotine_Saliva_tests.htm [accessed 2012-09-14] [WebCite Cache ID 6AfWiUXBk]
50. The EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
51. Sculpher M. NICE's 2008 Methods Guide: sensible consolidation or opportunities missed? Pharmacoeconomics 2008;26(9):721-724. [Medline: 18767891]
52. Hoogendoorn M, Welsing P, Rutten-van Mölken MP. Cost-effectiveness of varenicline compared with bupropion, NRT, and nortriptyline for smoking cessation in the Netherlands. Curr Med Res Opin 2008 Jan;24(1):51-61. [doi: 10.1185/030079908X242917] [Medline: 18021492]
53. Hoogwegt MT, Hoeks SE, Pedersen SS, Scholte op Reimer WJ, van Gestel YR, Verhagen HJ, et al. Smoking cessation has no influence on quality of life in patients with peripheral arterial disease 5 years post-vascular surgery. Eur J Vasc Endovasc Surg 2010 Sep;40(3):355-362. [doi: 10.1016/j.ejvs.2010.05.013] [Medline: 20580273]
54. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998;18(2 Suppl):S68-S80. [Medline: 9566468]
55. Casparie AF, van Hout BA, Simoons ML. [Guidelines and costs]. Ned Tijdschr Geneeskd 1998 Sep 19;142(38):2075-2077. [Medline: 9856218]
56. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. Oxford: Oxford University Press; 2007.
57. Raad voor de Volksgezondheid en Zorg. Zoetermeer. 2006. Zinnige en duurzame zorg URL: http://www.rvz.net/uploads/docs/Advies_-_Zinnige_en_duurzame_zorg.pdf [accessed 2012-12-20] [WebCite Cache ID 6D3RQDL1Z]
58. Pieterse ME, Seydel ER, DeVries H, Mudde AN, Kok GJ. Effectiveness of a minimal contact smoking cessation program for Dutch general practitioners: a randomized controlled trial. Prev Med 2001 Feb;32(2):182-190. [doi: 10.1006/pmed.2000.0791] [Medline: 11162345]
59. Salize HJ, Merkel S, Reinhard I, Twardella D, Mann K, Brenner H. Cost-effective primary care-based strategies to improve smoking cessation: more value for money. Arch Intern Med 2009 Feb 9;169(3):230-5; discussion 235. [doi: 10.1001/archinternmed.2008.556] [Medline: 19204212]
60. Kaper J, Wagena EJ, van Schayck CP, Severens JL. Encouraging smokers to quit: the cost effectiveness of reimbursing the costs of smoking cessation treatment. Pharmacoeconomics 2006;24(5):453-464. [Medline: 16706571]
61. Shiffman S, Patten C, Gwaltney C, Paty J, Gnys M, Kassel J, et al. Natural history of nicotine withdrawal. Addiction 2006 Dec;101(12):1822-1832. [doi: 10.1111/j.1360-0443.2006.01635.x] [Medline: 17156182]
62. Hays JT, Croghan IT, Baker CL, Cappelleri JC, Bushmakin AG. Changes in health-related quality of life with smoking cessation treatment. Eur J Public Health 2012 Apr;22(2):224-229. [doi: 10.1093/eurpub/ckq137] [Medline: 20884658]
63. Sales MP, Oliveira MI, Mattos IM, Viana CM, Pereira ED. The impact of smoking cessation on patient quality of life. J Bras Pneumol 2009 May;35(5):436-441 [FREE Full text] [Medline: 19547852]
64. Shum C, Humphreys A, Wheeler D, Cochrane MA, Skoda S, Clement S. Nurse management of patients with minor illnesses in general practice: multicentre, randomised controlled trial. BMJ 2000 Apr 15;320(7241):1038-1043 [FREE Full text] [Medline: 10764365]
65. 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]
66. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. Patient Prefer Adherence 2008;2:57-65 [FREE Full text] [Medline: 19920945]
67. Wangberg SC, Nilsen O, Antypas K, Gram IT. Effect of tailoring in an internet-based intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011;13(4):e121 [FREE Full text] [doi: 10.2196/jmir.1605] [Medline: 22169631]
68. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CEA: cost-effectiveness analysis
CUA: cost-utility analysis

CEAC: cost-effectiveness acceptability curve
CUAC: cost-utility acceptability curve
FTND: Fagerström Test for Nicotine Dependence
ICER: incremental cost-effectiveness ratio
ICM: I-Change Model
ICUR: incremental cost-utility ratio
MTC: multiple tailoring and counseling
MT: multiple tailoring
NMB: net monetary benefit
OCT: over-the-counter
QALY: quality-adjusted life year
RCT: randomized controlled trial
UC: usual care
WTP: willingness to pay

Edited by G Eysenbach; submitted 19.01.12; peer-reviewed by J Banks, J Van Ooteghem; comments to author 29.07.12; revised version received 26.09.12; accepted 19.12.12; published 14.03.13.

Please cite as:
Smit ES, Evers SMAA, de Vries H, Hoving C
Cost-Effectiveness and Cost-Utility of Internet-Based Computer Tailoring for Smoking Cessation
J Med Internet Res 2013;15(3):e57
URL: http://www.jmir.org/2013/3/e57/
doi:10.2196/jmir.2059
PMID:23491820

Original Paper

Recruitment to Online Therapies for Depression: Pilot Cluster Randomized Controlled Trial

Ray B Jones[1], PhD; Lesley Goldsmith[1], PhD; Paul Hewson[2], PhD; Christopher J Williams[3], MD, FRCPsych

[1]Plymouth University, Faculty of Health, Education, and Society, Plymouth, United Kingdom

[2]Plymouth University, Department of Maths and Computing, Plymouth, United Kingdom

[3]University of Glasgow, Institute of Health and Wellbeing, Glasgow, United Kingdom

Corresponding Author:
Ray B Jones, PhD
Plymouth University
Faculty of Health, Education, and Society
3 Portland Villas
Plymouth, PL4 8AA
United Kingdom
Phone: 44 441752586532
Fax: 44 441752586532
Email: ray.jones@plymouth.ac.uk

Abstract

Background: Raising awareness of online cognitive behavioral therapy (CBT) could benefit many people with depression, but we do not know how purchasing online advertising compares to placing free links from relevant local websites in increasing uptake.

Objective: To pilot a cluster randomized controlled trial (RCT) comparing purchase of Google AdWords with placing free website links in raising awareness of online CBT resources for depression in order to better understand research design issues.

Methods: We compared two online interventions with a control without intervention. The pilot RCT had 4 arms, each with 4 British postcode areas: (A) geographically targeted AdWords, (B) adverts placed on local websites by contacting website owners and requesting links be added, (C) both interventions, (D) control. Participants were directed to our research project website linking to two freely available online CBT resource sites (Moodgym and Living Life To The Full (LLTTF)) and two other depression support sites. We used data from (1) AdWords, (2) Google Analytics for our project website and for LLTTF, and (3) research project website. We compared two outcomes: (1) numbers with depression accessing the research project website, and then chose an onward link to one of the two CBT websites, and (2) numbers registering with LLTTF. We documented costs, and explored intervention and assessment methods to make general recommendations to inform researchers aiming to use similar methodologies in future studies.

Results: Trying to place local website links appeared much less cost effective than AdWords and although may prove useful for service delivery, was not worth pursuing in the context of the current study design. Our AdWords intervention was effective in recruiting people to the project website but our location targeting "leaked" and was not as geographically specific as claimed. The impact on online CBT was also diluted by offering participants other choices of destinations. Measuring the impact on LLTTF use was difficult as the total number using LLTTF was less than 5% of all users and record linkage across websites was impossible. Confounding activity may have resulted in some increase in registrations in the control arm.

Conclusions: Practitioners should consider online advertising to increase uptake of online therapy but need to check its additional value. A cluster RCT using location targeted adverts is feasible and this research design provides the best evidence of cost-effectiveness. Although our British pilot study is limited to online CBT for depression, a cluster RCT with similar design would be appropriate for other online treatments and countries and our recommendations may apply. They include ways of dealing with possible contamination (buffer zones and AdWords techniques), confounding factors (large number of clusters), advertising dose (in proportion to total number of users), record linkage (landing within target website), and length of study (4-6 months).

Trial Registration: clinicaltrials.gov (Registration No. NCT01469689); http://clinicaltrials.gov/ct2/show/NCT01469689 (Archived by WebCite at http://www.webcitation.org/6EtTthDOp)

(J Med Internet Res 2013;15(3):e45) doi:10.2196/jmir.2367

KEYWORDS

cluster randomized trial; pilot study; online advertising; depression; MoodGym; Living Life to the Full, LLTTF; cCBT; Google Analytics; AdWords; computerised CBT

Introduction

Less than 60% of people with diagnosable depression or anxiety seek formal help from practitioners; this represents a significant treatment gap [1]. The remainder may access informal care, alternative therapies, make private arrangements such as counselling, use the voluntary sector, the Internet, or use other sources of information. Nearly 1 in 5 British Internet users search for information related to mental health [2], but patients searching for health information online may not find what they are looking for [3], possibly due to sub-optimal search strategies [4].

There is increasing evidence that online interventions can be effective in changing health behaviors or improving health [5]. Online cognitive behavioral therapy (CBT), is an example offering effective self-help treatment for depression. Online CBT is recommended by the National Institute for Health and Clinical Excellence (NICE) in the UK for mild to moderate depression [6]. The range of resources includes licenced (paid for) sites such as Beating the Blues [7], and free access websites providing access to CBT life skills resources (eg, Living Life To The Full (LLTTF) [8] and MoodGYM [9]).

However, many patients do not benefit through lack of awareness. We found a variation in registration to LLTTF of 15-fold between the highest (Kirkwall, Scotland) and lowest (Wigan, England) postcode areas [10]. Variations in Internet use by region in Britain are small; 80.3% of those in the North East had at some time accessed the Internet compared to 87.6% in London by 2012 [11]. Variation in the prevalence of depression between postcode areas is small [12], and although other packages such as MoodGym may be in use, the most likely explanation for variation in LLTTF registration was lack of awareness.

Raising awareness of online CBT could benefit many people with depression by facilitating early rapid access. Ways of addressing this include online advertising through search engines such as Google AdWords (AdWords) [13-19], advertising on social media sites [20], "snowballing" on social media sites [21], getting websites of other organizations to add links (weblinks) [22], using offline via mass media [23], or via practitioners [24].

AdWords have been used by others to recruit participants to studies, for example, for depression screening [25], use of condoms [26], and quit smoking campaigns [27], but the cost effectiveness of their use has not been assessed. Weblinks from a range of existing sites have also been used in research studies, often routinely as part of a "recruitment package" (eg, [22] and LLTTF is linked from sites such as the English National Health Service's (NHS) website NHS Choices [28]), but we were unaware of any study of weblinks restricted to local organization websites in order to target recruitment on a specific local population as required for this design.

Studies that compare recruitment methods before and after interventions have no control group that could be described and compared with adequately. The only rigorous way of comparing methods of raising awareness of online therapies is by geolocated cluster randomised controlled trials (RCT, [10]). Matching intervention and control areas reduces the chances of bias. If we can limit online advertising to one geographic area and compare it with another area where there is no advertising, we can then assume that any difference is due to the advertising. We can therefore estimate its cost effectiveness in terms of cost per new user. Subsequently, we can decide if it is worth using online advertising to raise awareness, or whether other methods would be more cost effective.

We planned to compare purchasing AdWords with the second strategy of using weblinks and carried out a pilot RCT of the two different recruitment interventions to check the methods and outcome measures for a definitive trial later. In particular, our study objectives were to explore: (1) whether or not the two recruitment interventions seemed to work at all, and so be worthy of further study, (2) the ability to target online Google adverts or weblinks from other local websites without contamination, (3) the ability to link data sources to effectively measure the impact of the recruitment interventions, and (4) to learn more about the likely size and impact on target CBT site use and, given the other methodological issues such as confounding factors, to know what sample size and dose of advertising will be needed for future substantive studies.

Methods

Ethics and Registration

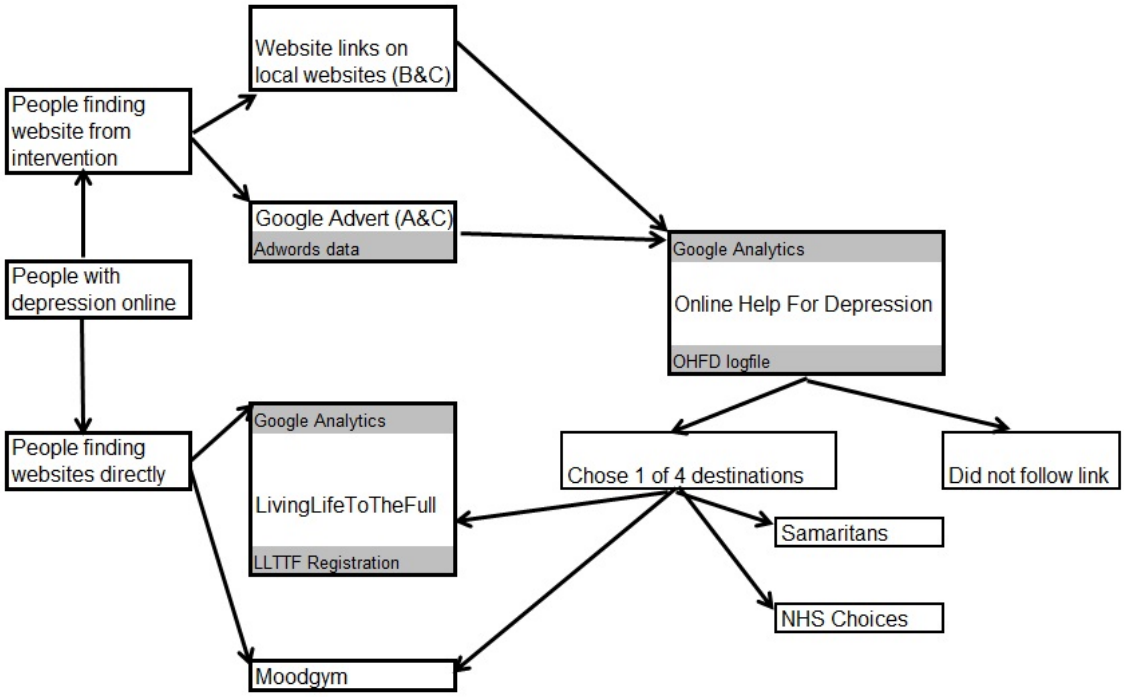
The study was approved by the NHS South West 2 Research Ethics Committee (Reference 11/H0203/8; February 2011) and registered on clinicaltrials.gov (Registration No. NCT01469689).

Design

This was a pilot cluster RCT of recruitment interventions for online CBT for depression. We compared two online recruitment interventions with a control without intervention. The pilot RCT had 4 arms: (A) geographically targeted paid AdWords, (B) free adverts/weblinks placed on local websites by contacting those website owners (weblinks), (C) both interventions combined, (D) control. Participants were directed to a project recruitment website linking to two freely available online CBT sites (Moodgym and LLTTF) and two other sites (Samaritans [29] and NHS Choices [28]). We used data from (1) AdWords, (2) Google Analytics (Analytics) for our project website and for LLTTF, and (3) our Online Help For Depression (OHFD) project website. We examined 2 outcomes in the 4 arms of the

study: (1) numbers accessing the project recruitment website that had depression, and then chose an onward link to one of the two CBT websites, and (2) numbers registering with LLTTF. We documented costs and explored intervention and assessment methods to make recommendations for a definitive trial, for other researchers carrying out similar studies, and tentative recommendations for practitioners and policy makers. Figure 1 shows a schematic of the design.

Figure 1. Schematic of the design showing websites and data sources (in gray).



Sampling and Randomization

All 121 postcode areas in England, Wales, and Scotland were divided into quartiles by rate of registration on LLTTF (based on 36,753 registrations of people between June 2008 and June 2009 who scored 8 or more on the Hospital Anxiety and Depression Scale for either anxiety or depression) and into quartiles by population size (based on the 2001 census). We randomly selected 4 cells from this array, and chose 4 nearly consecutive postcode areas for each cell (to try to achieve similar populations), avoiding adjacent geographical areas. Each set of 4 postcode areas was randomly allocated to the 4 arms of the trial (three interventions and control). Randomization was not blind.

Sample

Each arm of the study had a total population ranging from 1.6 to 2 million people clustered in 4 postcode areas. In total the study included 7 million people in 16 postcode areas across England, Wales, and Scotland. The estimated point prevalence for major depression among 16- to 65-year olds in the UK was 2.1%, rising to 9.8% when the less specific and broader category of “mixed depression & anxiety” was included [30]. We used an intermediate estimate of the prevalence of depression of 5%, which gave an estimated target population of 350,000.

Interventions

Those paying for AdWords campaigns set up one or more adverts, and enter keywords to help determine when the advert is shown. AdWords displays adverts as a sponsored link, either at the top of the list of search results or in the right hand search results panel, depending on the phrase entered, the price offered per advert, bids from competing adverts, and (if requested) by estimated location of the user.

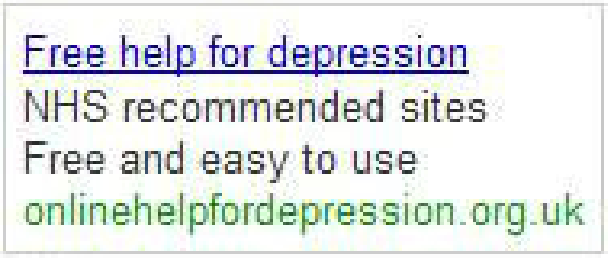
Within the 8 areas in arms A and C, we ran AdWords from April 17th to November 30th, 2011. In the 8 areas in B and C, we aimed to place adverts (weblinks) from local organization websites such as local universities, general practitioner (GP) practices, and local authorities, by contacting these organizations via email and/or phone starting on April 17th. Arm D was a control arm with no recruitment intervention.

Google AdWords

We used a single advert (Figure 2) from April 17th to October 19th as one campaign with a daily budget of £7.50 per day (4.4 pence per 1000 target population per day). Targeting specific postcode areas (eg, Kingston, KT) was not an option offered by AdWords. Options did however include targeting a radius of 1 mile or more around a postcode district (eg, KT2) or to hand draw a polygon to enclose the area of interest. We used a mix of methods: 4 postcode areas were defined using circles of 1 mile radius for all postcode districts within the postcode area and 4 were defined by hand drawn polygons. Preliminary analysis [10] showed leakage from the target areas and an imbalance in the number of presentations of the advert between postcode areas. This prompted a change of strategy with 8 separate adverts for each postcode area and adverts that mentioned the target area. Radius targeting (rather than polygon

targeting) was used for all areas from October 19th to November 30th. The daily budget was increased to £16.10 (9.3 pence per 1000 target population per day) and divided between the 8 areas in proportion to target population.

Figure 2. Google Advert.



We originally asked AdWords to display the advert for the keyword of depression. AdWords suggested other similar keyword combinations and we accepted all suggestions. AdWords gave information on the number of times they presented adverts against Google searches, by day, keyword, and location. AdWords decided when to present the advert based on the price we offered, the price of competing adverts, and other factors such as the search terms used. Users searching on terms such as depression and depression help were presented with our advert, depending on our budget and competing adverts.

Local Weblinks

Google searches were used to identify organizations in 8 postcode areas (arms B and C) with websites to place adverts (weblinks to our project website) on local free access websites. In our search we included websites such as local GP surgeries, local media websites including newspaper, TV, radio, further and higher education institutions, and community-based or local charity websites. In total, 180 emails were sent with 3 weblinks posted free of charge: two in Leeds (university medical practice and a carers' organization), and one in Kirkwall (local online community newspaper).

Project Website

Those who clicked on Google adverts or on weblinks were directed to our project research website OHFD. This gave information about the study and advised that completion of the online questions implied consent for it to be used in the study (see Multimedia Appendix 1 for screenshots). Computer Internet Protocol (IP) addresses, dates, and times were collected via OHFD, and also monitored by Analytics. We specifically did not try to raise the visibility of our website to normal Google, Yahoo, or Bing search engines. Visitors were asked for their postcode area and to complete the Patient Health Questionnaire (PHQ9) [31] assessing depression. Users were then offered 4 links to Moodgym, LLTTF, NHS Choices information on depression, and Samaritans. The order in which the links to Moodgym and LLTTF (top row) and NHS Choices and Samaritans (bottom row) appeared was randomized within row.

Data Sources

We used various sources of data to model patient flow (Figure 1). AdWords and Analytics were used for OHFD. The website log for OHFD recorded stated postcode areas from participants and their website destination choice (if made). Analytics was used for all visitors of LLTTF, identifying those referred by OHFD and those likely to be in study arms (using Google defined locations). Website log was used for LLTTF, including those who registered and stated their postcode area.

Costs

We documented costs for using AdWords and other weblinks and estimated costs per person with depression referred to the online CBT via AdWords, based on time spent, cost £10/hour (based on the hourly rate Plymouth University pays temporary administrative staff) and Google's charges. We included costs needed for routine delivery of these methods and excluded research costs such as time spent in the comparison of methods or in setting up the research project website.

Outcomes

We defined 2 main outcomes and compared them between the 4 arms: (1) the numbers accessing the project research website (OHFD) that completed a PHQ9 depression rating score and had a score of more than 5, indicating at least mild depression, and then chose an onward link to one of the two CBT websites, and (2) the numbers registering with LLTTF who gave their postcode, comparing intervention with control and all areas. In the revised version of LLTTF (issued in January 2011) used in this study, new arrivals at LLTTF do not have to immediately register, instead, registration can be delayed. Registration, by giving more personal details and agreeing to email reminders, signifies a commitment to use the site seriously. In the previous version of LLTTF used to select the sample, most visitors to the site registered as they could not access most content until after registration.

User Panel

We carried out this study at a time when the Improved Access to Psychological Therapies (IAPT) project [32] was investing large amounts of money to improve access to psychological therapies across England. To try to identify other local initiatives and how these may have affected our sample and interventions, we aimed to recruit a panel of informants in our study areas. We emailed five ex-users of LLTTF from each of the 16 areas (ie, total 80), and also tried to contact them via IAPT teams. Panel members were to be paid £30 at the end of the study by e-vouchers.

Results

Did the Two Interventions Work at All?

Examining the number of people clicking through from online adverts to OHFD to online CBT (Figure 3), we saw that our OHFD research site had a total of 8231 visits, the majority (97%, 7980/8231) from Google adverts and only 1% (103/8231) from weblinks. Half of those who visited OHFD (50%, 4118/8231) interacted with the site and 76% (3135/4118) made a choice of destination from the 4 available sites, three-quarters of whom (77%, 2403/3135) chose the online CBT. Although 96% (2306/2403) had depression (according to the PHQ9 score >5), only 515/2306 (22%) were from the study target areas (Table 1).

Table 1 shows that the rates of referral to online CBT for AdWords arms A and C were higher than the weblinks only arm B (450 and 387 vs 93, respectively, per 100,000) and much higher than the control arm (D) (28 per 100,000). These results were confirmed by fitting a Poisson generalized linear model with the R software [33], with predictors for the use of AdWords, weblinks, and an interaction term for both. This produced a 95% confidence interval for rates relative to control of (10.2 and 27.4, respectively) when using Google adverts. The rate relative to control for weblinks was (2.0 and 6.0, respectively) and there was little to be gained from using both, given an interaction between AdWords and weblinks with rate relative to baseline of (0.1, 0.5). The majority of those choosing online CBT chose LLTTF (66%, 1581/2403) and the difference between arms A/C and B/D was still evident (Table 1).

Was it Possible to Link Data to Measure Outcome Two?

It proved impossible to directly track individuals from OHFD to LLTTF using IP address and time data, therefore the impact of the interventions on LLTTF had to be assessed using Analytics and LLTTF website data. Analytics reported 1474 visits landing on LLTTF from OHFD, agreeing approximately with the 1581 referred by OHFD, although the 7% attrition is unexplained (Figure 3). During the period of study, according to Analytics, the majority of the 230,441 visits to LLTTF were from normal search (41%, 93,983/230,441) or direct (37%, 85,000/230,441), with 22% (51,406/230,441) from referring sites. In total, there were 1888 sites from which there were 51,406 visits. Of the referring sites, OHFD was the seventh largest referrer with 1474 but representing just 2.9% (1474/51,406) of referrals and less than 1% (1474/230,441) of all visits. The Royal College of Psychiatrists sent most referrals (5071/51,406, 9.9%), but this still represented only 2.2% (5071/230,441) of all visits.

We know from Analytics that the bounce rate (ie, those who exited from the first page) for visits from OHFD was 47% (no different to the average bounce rate from referred visits of 49%). This suggests that those arriving from OHFD were not more or less likely to continue and to subsequently register. We can see from the difference between arrivals on the site (14,396 from Analytics) and registrations (1143 from LLTTF log data) from the study arms that 10% (1067/10569) of all visitors registered on the current version of LLTTF.

Table 1 shows that the 95% confidence intervals of rate of registration on LLTTF (outcome 2) overlap between all 4 arms, in other words there was no significant difference between the interventions or control. Although none of the differences for individual postcode areas was significant, registration rates for most (11/16, 69%) postcodes tended to decline from January-April to April-November. Of the more populous postcode areas, Liverpool (arm A) showed the greatest increase (360 to 489, 26% increase) but this was still not significant and Nottingham (in the control arm D) had a 6% increase. Two of the small Scottish island postcode areas may have increased but their populations are small, the confidence intervals on estimates are large, and the impact on the whole arm small.

Figure 3. Participant flow diagram showing overall recruitment and different data sources (April 17th - November 30th 2011). Shaded boxes show numbers for the two outcomes.

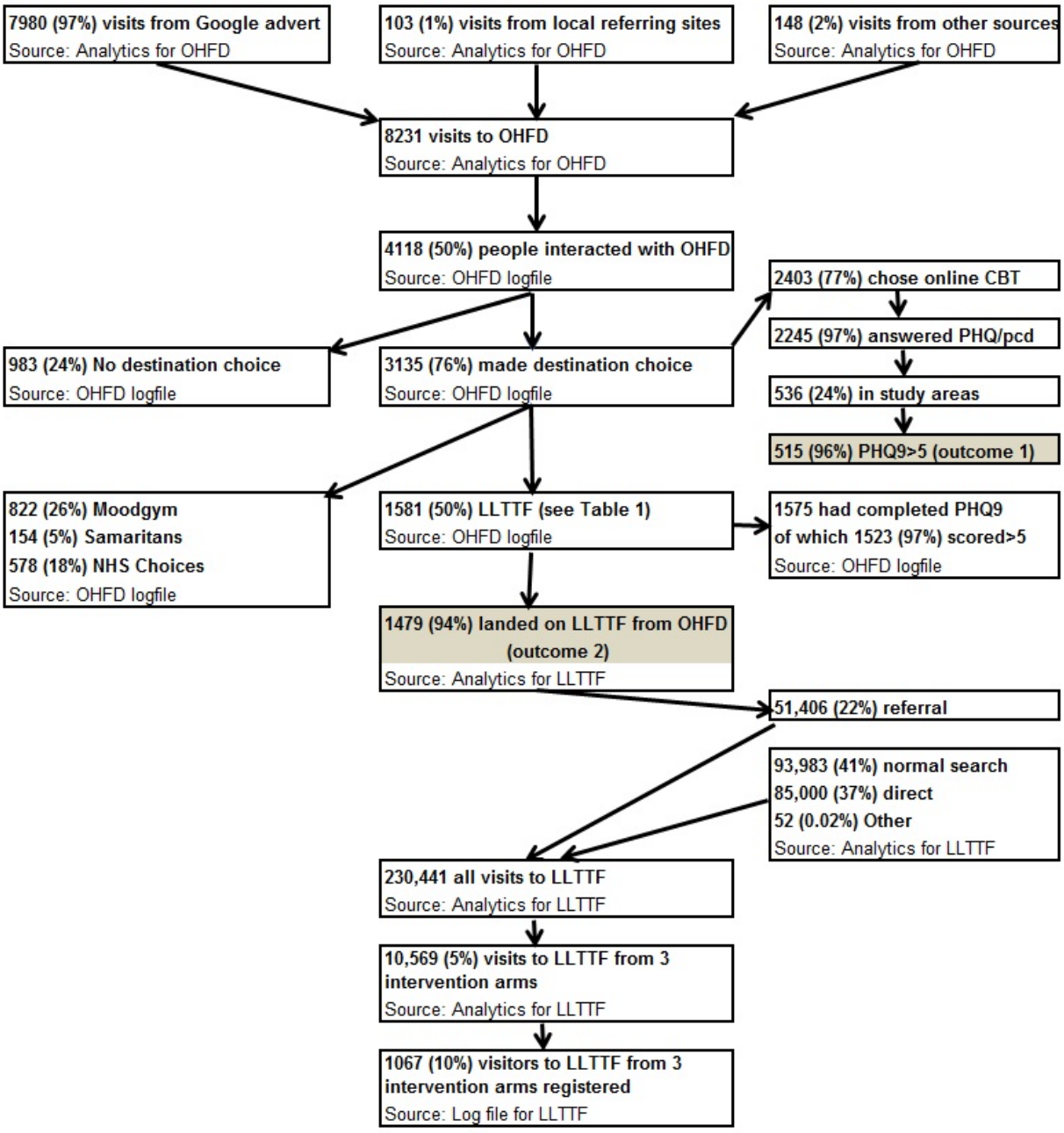


Table 1. Number of people on OHFD with depression (PHQ>5) choosing online CBT and LLTTF, registering on LLTTF, annual rate per 100,000 estimated depressed registering on LLTTF before and during interventions, by postcode area, and trial arm.

		Outcome 1			Outcome 2			
		People on OHFD with depression choosing online CBT and LLTTF			People registering on LLTTF before the intervention (Jan-Apr 2011) and during the intervention (Apr-Nov 2011)			
		Intervention period Apr-Nov 2011			Before intervention Jan-Apr 2011		Intervention period Apr-Nov 2011	
	Estimate of people with depression	People with depression who chose online CBT	Annual rate per 100,000 depressed (95% CI)	People with depression who chose LLTTF	Reg[a] LLTTF	Annual rate per 100,000 depressed (95% CI)	Reg[a] LLTTF	Annual rate per 100,000 depressed (95% CI)
Arm A								
Liverpool (L)	42,173	155	588 (495-681)	111	38	360 (246-475)	129	489 (405-574)
Redhill (RH)	24,721	49	317 (228-406)	34	59	955 (711-1198)	94	608 (485-731)
Lancaster (LA)	16,299	42	412 (288-537)	27	44	1080 (761-1399)	73	717 (552-881)
Harrogate (HG)	6668	7	168 (44-292)	6	19	1140 (627-1652)	40	960 (662-1257)
Total	89,860	253	450 (395-506)	178	160	712 (602-823)	336	598 (534-662)
Arm B								
Leeds (LS)	36,867	24	104 (62-146)	16	69	749 (572-925)	171	742 (631-853)
Southend (SS)	24,660	12	78 (34-122)	8	55	892 (656-1128)	130	843 (698-988)
Slough (SL)	16,882	10	95 (36-154)	6	22	521 (303-739)	59	559 (417-702)
Kirkwall (KW)	2505	1	64 (-61-189)	0	17	2715 (1424-4005)	30	1916 (1230-2602)
Total	80,914	47	93 (66-120)	30	163	806 (682-929)	390	771 (695-848)
Arm C								
London (SW)	39,167	104	425 (343-506)	77	86	878 (693-1064)	170	694 (590-799)
Kingston (KT)	24,505	72	470 (362-579)	50	36	588 (396-780)	75	490 (379-601)
Darlington (DL)	17,074	21	197 (113-281)	15	37	867 (588-1146)	78	731 (569-893)
Shetland (ZE)	1099	1	146 (-140-131)	1	5	1819 (225-3414)	18	2620 (1409-3830)
Total	81,846	198	387 (333-441)	143	164	802 (679-924)	341	667 (596-737)
Arm D								
Nottingham (NG)	54,012	10	30 (11-48)	8	102	755 (609-902)	272	806 (710-902)

		Outcome 1			Outcome 2			
		People on OHFD with depression choosing online CBT and LLTTF			People registering on LLTTF before the intervention (Jan-Apr 2011) and during the intervention (Apr-Nov 2011)			
		Intervention period Apr-Nov 2011			Before intervention Jan-Apr 2011		Intervention period Apr-Nov 2011	
	Estimate of people with depression	People with depression who chose online CBT	Annual rate per 100,000 depressed (95% CI)	People with depression who chose LLTTF	Reg[a] LLTTF	Annual rate per 100,000 depressed (95% CI)	Reg[a] LLTTF	Annual rate per 100,000 depressed (95% CI)
Oldham (OL)	22,190	3	22 (-3-46)	1	42	757 (528-986)	69	498 (380-615)
Dudley (DY)	19,882	4	32 (1-64)	4	30	604 (388-820)	56	451 (333-567)
Hebrides (HS)	1325	0	0 (0-0)	0	4	1207 (24-2391)	15	1811 (895-2728)
Total	97,409	17	28 (15-41)	13	178	731 (624-838)	412	677 (611-742)
Study total	35,0028	515	235 (215-256)	364	665	760 (702-818)	1479	676 (642-711)
Other areas	25,04345	1791	114 (119-120)	1130	5613	897 (873-920)	11894	760 (746-773)
E, W, and S	28,54373	2306	129 (124-135)	1581[b]	6278	880 (858-902)	13,373	750 (737-762)

[a]registrations

[b]1581 includes 87 who gave no postcode on OHFD, but all other indications (eg, IP address) show that they were England (E), Wales (W), or Scotland (S).

Validity Check on Location

As a validity check on location, we carried out an alternative analysis on those registering on LLTTF using the Analytics location data instead of user stated postcodes. The two approaches showed agreement.

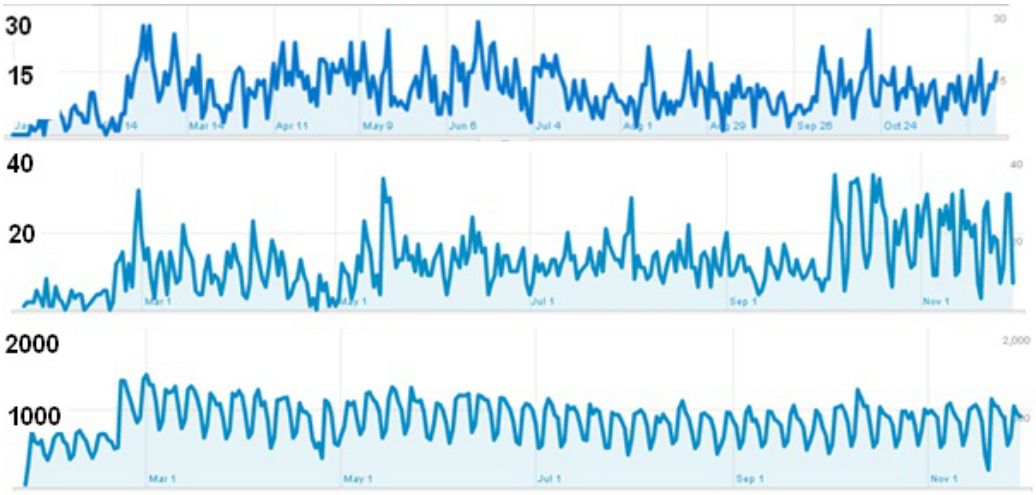
Possible Confounding Factors

According to Analytics, the biggest increase in "landings" on LLTTF was seen in one of the control areas, Nottingham. Figure 4 (middle panel) suggests that there may have been a "step change" for Nottingham in October 2011. Overall, the numbers visiting LLTTF were largely unchanged over the study period (apart from the very regular weekly cycle of visits with fewer during weekends, Figure 4 bottom panel).

Size of the Impact on LLTTF

Ignoring the stated locations of those clicking through to LLTTF, we see that 1581 clicked through to LLTTF in the study period (Figure 3). In the study period, the 3 intervention arms had 10,569 visits (estimated from Analytics) but only 1067 people registered, that is, with the current method of counting registration, only 10% of visitors to LLTTF registered. So if we estimated that, overall,10% of those referred from OHFD registered, we can see that the numbers referred from OHFD (for example, 17 in Arm A, Figure 5) were likely too small to greatly influence the total numbers registering in each of the trial arms (Figure 5, outcome 2-people registering on LLTTF). As we were unable to link individual participants from the project website (OHFD) to LLTTF, we could only tell if there had been an impact on the number of people registering on LLTTF if it was of sufficient size. If it had been possible to follow through individual participants from clicking on an advert to registering on LLTTF, having an impact on the total number registering from each area would have been less important.

Figure 4. Number of visits to LLTTF according to Google Analytics. Top graph shows Liverpool, middle Nottingham, and the bottom graph shows ALL UK visits between January and November 2011.



Leakage

Just under 8000 (7980/8231, 97%, Figure 3) visits to OHFD came from Google adverts, of which 4118/7980 (52%) interacted with OHFD, 3135/4118 made a destination choice, but only 515/3135 (16%) were from study arms with PHQ9 >5 who chose either Moodgym or LLTTF. Of these, only 387/4118 (9.3% of those who interacted with OHFD) chose LLTTF. Of those choosing LLTTF, probably only 10% registered, ie, less than 0.5% of those who clicked on the advert were likely to have registered on LLTTF. The biggest sources of "leakage" were from the location targeting [10] and from participants not engaging fully on LLTTF. The leakage however did not lead to much contamination of control areas. Failing to continue with the OHFD website and being lost to other or no Web destination also caused substantial leakage.

Cost

All 3 weblinks were for locations in arm B. Arm C therefore was effectively another AdWords arm. The total cost of the AdWords campaign in payments to Google was £1841. From April 17th to October 19th, we set a daily budget of £7.50 (4.4 pence per 1000 population) and had one campaign in which all 8 areas were included. As we described elsewhere [10], there was a disproportionate spending on AdWords for London SW. To counter this and to better understand the responses in each area, on October 19th, a new campaign in which each postcode area had its own budget was started. For the remaining 6 weeks, we increased the daily budget to £16.10 (9.4 pence per 1000 population), and divided this in proportion to the target population in each area (but giving a minimum budget of 60 pence for Shetland). Table 2 shows that the average number of clicks per day in the last 60 days was 29.3 (1759/60) compared to 34.2 (6291/184) in the first 184 days. This method (restricting the spending on London SW) also resulted in a higher cost/click overall, although we had noted in the first period that it takes AdWords 3-4 weeks to gain the optimum return on advertising spending after starting a new campaign.

We estimated that the time spent simply adjusting the AdWords campaign, as opposed to time spent on Analytics trying to match with other data, was 25 hours. At £10/hour, the total cost of the AdWords campaign was £2091 (£1841+£250). If we assume that both arms A and C were AdWords, this represents £4.64 per person for the 451 people who chose online CBT and had a PHQ9 >5 in arms A and C (the target group, Figure 5). Seventy hours were spent trying to contact owners of local websites to set weblinks, which if set at £10/hour, gives a total cost of £700 for arm B. This represents £14.89 per person for the 47 people in arm B (Figure 5).

Figure 5. Participant flow diagram showing randomization to the 4 arms of the study.

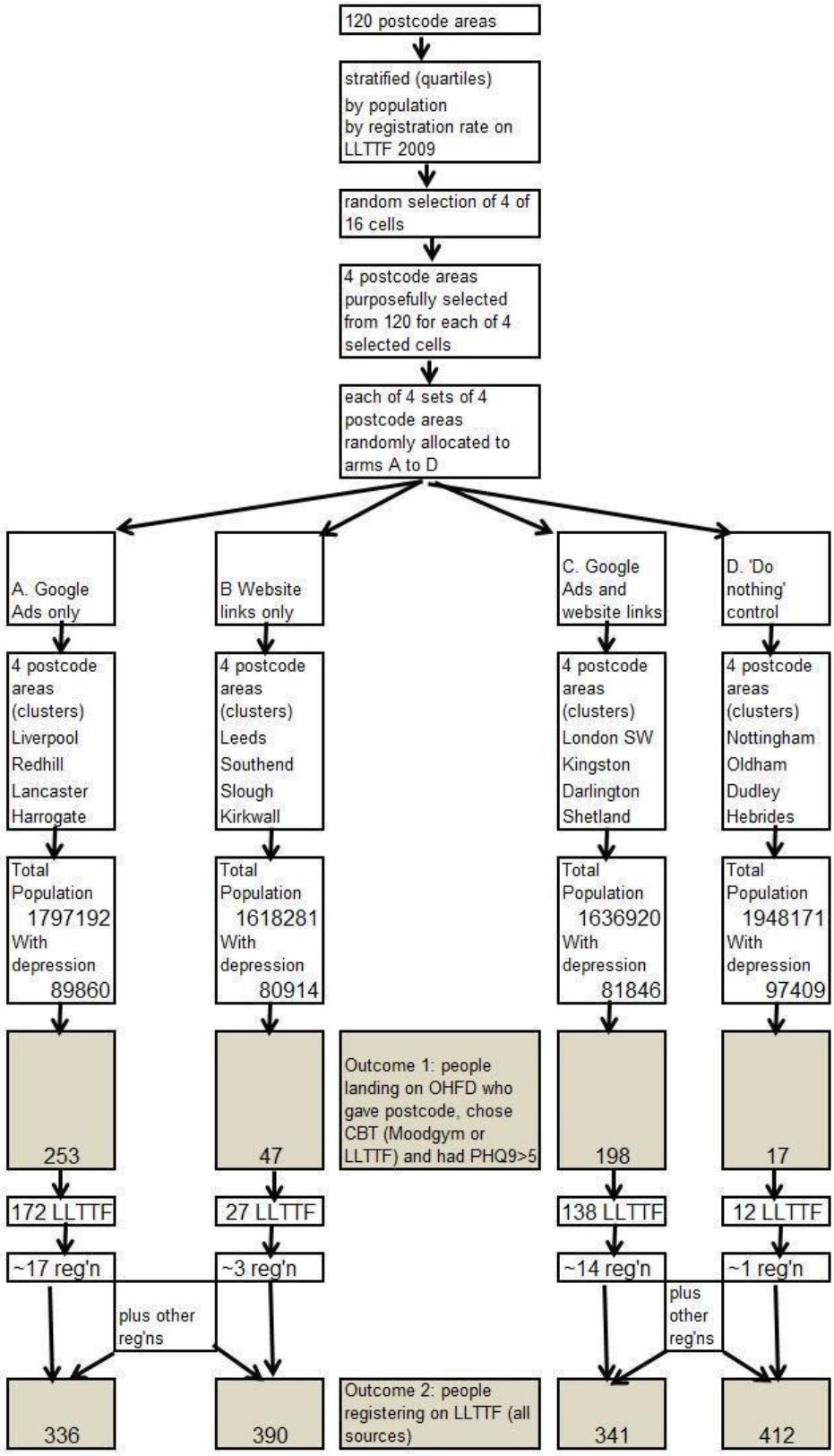


Table 2. Payments to Google for AdWords campaign.

Area	Maximum daily budget	Average daily spend	Clicks	Cost	Cost per click
April 17th -October 19th (184 days) 2011 single campaign					
All	£7.50	£7.07	6291	£1,301.04	£0.21
8 individual campaigns by area October 20th-November 30th (60 days)					
Liverpool	£3.84	£2.37	365	£142.45	£0.39
London SW	£3.56	£2.48	886	£148.83	£0.17
Redhill	£2.24	£1.19	182	£71.65	£0.39
Kingston	£2.22	£1.43	174	£85.99	£0.49
Darlington	£1.56	£0.76	66	£45.72	£0.69
Lancaster	£1.48	£0.68	64	£40.61	£0.63
Harrogate	£0.60	£0.08	21	£4.60	£0.22
Shetland	£0.60	£0.00	1	£0.29	£0.29
Total	£16.10	£9.00	1759	£540.14	£0.31

User Panel

In total, we were able to recruit 12 panel members and all arms, but not all postcode areas, of the study were represented by either an ex-user of LLTTF or a member of an IAPT team. No other local interventions were identified via the user panel but the panel was able to help in examining the search environment across Britain [34].

Discussion

Need for the Study

McCrone estimated the number of people with depression in England as 1.24 million (2.3%) with total cost of services at £1.7 billion while lost employment increased this total to £7.5 billion [35]. Others estimated that a third of people with depression are not in contact with services [36]. Online CBT has a significantly small-to-medium effect size compared to non-active controls for patients with a range of severity of depressive symptoms and was recommended by NICE for mild to moderate depression [6]. There is virtually no extra cost in additional people using online CBT websites such as LLTTF but there was evidence that many people with depression who might benefit from using online CBT were missing the opportunity of effective treatment simply through lack of awareness. By raising awareness of its availability directly to depressed people, we could reduce inequalities in access to this treatment.

Various ways of raising awareness of websites are available including online advertising and weblinks. We did not know if online adverts such as AdWords would be more or less effective than weblinks. These methods may simply attract the same people that would in any case have found online CBT. As a pilot study, we were looking at the feasibility of being able to answer these questions and what study design would enable us to do so.

Which Interventions Seemed to Work?

We have found that using AdWords is possibly going to be cost effective, whereas trying to get weblinks is not worth pursuing further in this locality based research study design. We only managed to put 3 weblinks in place. Scaling up the AdWords campaign would incur further advertising but no further labour costs (so the unit price gradually decreases), whereas scaling up the weblinks campaign would be proportional to the labour costs. Establishing weblinks therefore appears much less cost effective than AdWords and is not worth pursuing as a sole intervention in this study design.

However, the ability to establish weblinks may be significantly easier for local organizations trying to offer local services. In our study, we chose areas remote from the research team, and there were no prior local relationships on which to build. Being locally situated and delivering local services to local people might significantly affect the ease of establishing local websites. Others have used weblinks to successfully recruit as part of their recruitment package. For example, 91% (174/191) of participants in a trial of online treatment for chronic headache learned of the study from weblinks (both mutual and paid for) on websites, registration with major search engines, and notices posted to headache-related news groups, but the authors did not compare these methods and the study was global [22]. Overall 1 in 5 come to LLTTF from a weblink and various national website including NHS Choices and Royal College of Psychiatrists already have links to online CBT [34] and other nationwide charity websites may be willing to include weblinks. However, pursuing local websites such as primary care health centres to fit with this geo-targeted research design did not seem cost-effective. Nevertheless, this should not be discounted as an option for practitioners where it is known that weblinks can be put in place with little effort.

Design Issues to be Addressed in a Definitive Trial

Overview

AdWords were effective in recruiting people to the project website but our pilot study identified a number of issues for our, and similar, cluster RCTs. The main problems were leakage so that its effect was greatly diluted, problems in linking data across websites so that the impact of advertising was lost in the large numbers using LLTTF, and possible confounding factors. The dose of advertising (particularly given the "record linkage" problem) and length of study also need to be considered. The changes that would be needed to the design of a definitive trial are discussed below. These findings will be relevant to others seeking to improve the uptake of online interventions or designing online cluster RCTs.

Leakage

There was leakage at 3 points in the process. First, we have shown elsewhere [10] that there was leakage from the target areas, particularly into neighbouring areas. This was partly due to the methods used and could be improved by using an area that is within the radius of set distance around a point rather than hand-drawn polygons, use of more appropriate radius, and avoidance of the area "edge". Second, there was leakage caused by offering too many choices of destination to participants. We had included both MoodGym and LLTTF as the two main free online CBT sites as it was thought that there might be geographic variation in choice of online CBT. For ethical reasons we added Samaritans as another option, and added NHS Choices to create a balanced (even) number of options. As a result, only 50% of those making a choice chose LLTTF. As we did not have access to Moodgym log data, we were only able to follow participants to the point of choosing online CBT, and the impact on one of those sites (LLTTF). Third, with the current system of registration on LLTTF, only 10% of those accessing LLTTF decided to register and enter demographic data including their postcode (LLTTF allowed people to use the website without registering; registering offered further features such as email reminders).

To decide whether a definitive trial will be feasible, we modelled the effect of reduced leakage. By taking 50% of the number lost on the project website before making a destination choice, by routing all participants to LLTTF, and by reducing losses from location targeting from 76% to 20%, we would reduce the leakage from 94% to 47%.

Record Linkage and Insufficient Dose

We were not able to track individuals from OHFD to LLTTF and therefore could only use the overall data to try to estimate outcome 2, the number of people registering with LLTTF. Although outcome 1 showed significant differences between intervention and controls, to be able to measure a difference, the number referred from AdWords needed to be sufficiently larger compared to the total number that registered on LLTTF. The number of people with depression referred from OHFD to LLTTF was 4.7% (498/10569) of all users of LLTTF in study arms. If, as above, leakage was reduced to 47% the number referred by OHFD would still only be 10% of all those landing on LLTTF. To be certain that an impact of AdWords can be seen, the dose of advertising also needs to be increased. Although we doubled the daily budget to 9.3 pence per 1000 target population for the last two months of this pilot study, the number of clicks was slightly less than in the first period, probably due to breaking the advertising into 8 separate campaigns and AdWords not having had time "to settle". The effectiveness of the AdWords advertising budget depends on competing adverts, and is likely to be less cost effective as the daily budget is increased. As shown in Table 2, the average daily spend was short of the maximum budget so it is not clear if an increase in daily budget would be spent or would increase the number of clicks, but a further doubling of the advertising budget would be worth trying.

Confounding Factors

A major contributing factor to the lack of impact on outcome 2 was the increase in uptake of LLTTF in one of the control areas (Nottingham), possibly as a result of a presentation given to practitioners by one of the authors (CW) to a national conference that overlapped with a surge of use. Although it was not a blinded region selection, day-to-day management of the study was conducted by RJ. We recognized in designing the pilot study that random interventions or effects from other influences, such as local campaigns, could impact on intervention or control areas. For that reason, we had attempted to monitor all areas via the user panel. This pilot study suggested that, to avoid an overdue influence of one postcode area (cluster), each arm needs many more than 4 clusters. As we had designed the pilot study with 2 types of interventions and 4 arms, we only used 16 out of 120 postcode areas for our study.

A more robust approach would be to use all postcode areas in Britain, excluding London postcodes, ordered by population size only, randomized in pairs to two arms, excluding any adjacent postcode areas as buffer zones around study areas. Using this method produces a two-arm study with 32 postcode areas (see Multimedia Appendix 2) and population of approximately 10 million in each arm. One arm would then be randomized to the intervention and one to the control group. The largest postcode area in one arm (Birmingham) still represents 17% of the total arm, so there is still some danger that other activity might lead to confounding factors, however this is much reduced from the design we used in this pilot where Nottingham represented 55% of the control arm population. This sample (Multimedia Appendix 2) of postcode areas could be adopted for other studies in Britain.

Other Changes to Improve the Design of the Definitive Trial

Length of the Study

We paid for over 8 months of advertising but the recruitment numbers stabilized after a few months and we did not change the advert for 6 of the 8 months. On the other hand, it takes 2-3 months for AdWords to reach peak efficiency. With reduced leakage, a bigger sample, and an increased daily budget on advertising, the cost of a definitive study could probably be reduced by examining changes over 6 months.

Need for Better Methods of Assessing Location

Assessing location is subject to error. Previously [10], we have given a detailed comparison of user-reported versus methods based on IP address including using IP lookup tables and Analytics. The method we used to ask participants for their postcode area on our project website (OHFD) was probably the most accurate. We had hoped to be able to use IP addresses and the time of referral from OHFD to LLTTF to track individuals and so to estimate rates of completion of these "additional" registrants with other registrants. However, the changes to LLTTF registration procedures and the difficulties of trying to match IPs from one website to another using time of day made this impractical. LLTTF currently ask participants for their full postcode which probably deters some from giving any information and may encourage falsification or error. We have recommended changes to data collection methods in LLTTF, asking for postcode area from a drop down list. Analytics location information would be more useful if the area associated with a town name was transparent to users, and if it could be aligned with population figures. We have tried to suggest to Google that their geography should be changed. In the meantime, the geography used by Google for London and immediate surroundings suggest that, to have more accurate location data in a study, London should be excluded.

Finding a Period of Website Stability

Studies that aim to change Internet use will be limited by frequent changes to websites and other technical advances. In our case, even though CW was the author of LLTTF, changes underway to the LLTTF website could not be postponed for this study. Our sample was selected based on data extracted from LLTTF between the years 2008 and 2009 when most users of the site registered. In January 2011, after our study had been designed and ethical approval had been sought, LLTTF was reconfigured such that registration was optional and could be done at a later time. As a result, the number of people that registered greatly reduced. This meant that registration figures were no longer directly comparable with those collected earlier. Although we were able to compare the relative differences between regions in 2009 with 2011, by creating an index of use based on the lowest use region, we restricted direct comparison (ie, the "before" period) to January- April 2011.

Demographics

Graham [18] found that online advertising could be an effective and cost-efficient strategy to reach and engage Spanish-speaking Latino smokers in an evidence-based Internet cessation program. She concluded that cultural targeting and smoking-relevant images might be important factors for banner advertisement design. In this pilot we did not collect demographic details. These were collected by the target LLTTF website as part of registration and we hope to achieve better record linkage in the future.

Sample Selection

In this pilot study we selected our sample trying to match for the populations of postcode areas and the baseline registration rates on LLTTF. Although it would be important not to have a grossly imbalanced sample, the need for a greater number of postcode areas (as described above) while ensuring geographical dispersion to avoid contamination would seem to be more important in the design.

Comparison Against Other Methods

Although various online methods exist to raise awareness, the problem of being able to select sample areas allowing for the competing demands of contamination and confounding factors suggest that having a factorial design with more than one recruitment intervention would be difficult. On the other hand, for practitioners seeking to increase recruitment using a mix of all possible methods, more than one recruitment intervention would seem to be sensible.

Conclusions for Practitioners and Policy Makers

Our pilot study confirms other research that many people search online for help with mental health issues [2] such as depression; our advert was displayed 673,074 times in just over 7 months for a total targeted population of 3.5 million. Anecdotally, in discussing this research, many practitioners responded by saying "I never click on adverts" but our pilot has shown that nearly 8000 people clicked on our short advert for NHS recommended sites and many of these were depressed, according to their answers on a self-completed questionnaire.

In discussing this research with NHS policy makers, the response to advertising, perhaps in the light of previous criticisms of their expenditure, was that NHS websites have been optimized and therefore appear high in search results, so there is no need for online advertising. This may be true but our exploration of this [34] suggested that even though online CBT sites can eventually be found via NHS and Royal College of Psychiatrists websites that tended to appear high in search results, the probability was low, and was significantly increased for a naïve user by the addition of an advert. That evidence however was theoretical and the only way of knowing for sure about the cost-effectiveness of online advertising was to study it in a location targeted RCT.

Although the results of the weblinks in this pilot study was not as expected and will be excluded from this research design (unless a better method of placing local weblinks can be found), it would be wrong to conclude that weblinks are not relevant for practitioners and policy makers. The problem may be that the type of site containing relevant links tends to be either national (inappropriate for a location targeted clusters RCT) or very local (researchers likely do not have prior knowledge about existence of these sites). Practitioners and policy makers may therefore have to rely on weaker evidence from before/after studies of cost and impact to decide on how much effort they put into using weblinks.

Part of the reason why LLTTF had greater use in Scotland compared to other parts of Britain is that professionals in those areas were more aware of the existence of these sites and recommended use of these sites to patients. The role of professionals in recommending online CBT may also have explained the confounding results seen in one area in the control group. So it may be that continuing professional development and raising awareness among professionals about resources for depression may be as, or more effective, than direct-to-patient

online interventions. However only one third of patients were in contact with health services and other studies have shown that trying to reach patients to tell them about online resources may be labour intensive, time consuming, and very expensive [24]. Further research is needed to compare the cost-effectiveness of improving access via professionals versus direct to patient methods, but in the meantime, practitioners and policy makers should keep online advertising as an option.

Conclusions for Researchers

Overall Conclusions

A cluster geo-located RCT to test the cost-effectiveness of online advertising seems feasible in Britain. Geolocated adverts are offered by Google in other countries based on a radius around a point, so the general design of a cluster RCT to test the cost effectiveness of online advertising in raising awareness of an online therapy, as piloted in this study, would seem to hold true for other countries and for other online therapies or websites. However, this pilot study has demonstrated 4 general messages concerning contamination, confounding, dose of advertising, and length of study that will be useful for other researchers.

Contamination

Provided it is possible to have a sufficiently large buffer zone between intervention and control regions, it should be possible to deal with potential contamination. The definition of sufficiently large is vague and needs to be piloted in each country, but in Britain, we think a 2-arm trial with 16 postcode areas in each should be possible without too much contamination. The problems of including London in Britain were described in more detail elsewhere [10]. It seems likely that similar studies in other countries may need to exclude the capital or major centres for Internet providers. We have also previously described other design issues with using AdWords including the need for separate campaigns for each postcode area [10].

Confounding Factors

Ideally, such a trial would have a large number of clusters in each arm such that any one region is kept to approximately 5% of the total population in the arm. That would mean that any confounding activity such as local media campaigns would not influence the arm greatly. However, dealing with confounding factors conflicts, to some degree, with the relatively small and densely populated country like Britain. In our proposed best design for Britain, trying to keep contamination to a minimum at the postcode area level, our design of 2-arms of 16 postcode areas each, still included a city representing 17% of that arm. This was slightly risky for introducing confounding factors but was the best we could do. In a bigger country such as the United States, it should be possible to have a stronger design. In a smaller, densely populated country such as the Netherlands, this study design may be impossible.

Dose of Advertising

We recommend that in such a design, adverts are linked to a special “landing page” within the target site. This allows those following the link to track their use of the target site. We had used an intermediate project website in this pilot because we were offering 2 online CBT sites and 2 other sites. This gave us considerable problems in trying to track participants from the advert to the project website to the target website, and in trying to match up the different data sets available from AdWords, Analytics, and website logs. Having a landing page within the target website should remove most problems but it would still help in the design if the dose of advertising, and so the expected “footfall” from advertising was substantial compared to the number arriving from other sources. In this pilot, with a limited advertising budget, although we had significant numbers arriving at our project website, the problem of leakage meant that the number of people finding the target website compared to the numbers finding it from a normal search, national weblinks, or direct entry were small. Any study with this type of design should pay greater attention to the total footfall on the target website and try to calculate the dose of advertising needed to have a demonstrable impact.

Length of Study

The cost of advertising in such a trial will be partly determined by how long the campaign is run. We ran our pilot for 7.5 months. Although it did take several weeks for our campaign to stabilize and for AdWords to get the best return and cost per click, a campaign of 4-6 months should be sufficient for this type of study where the total number clicking is in the range of 300-500 people. The best design (as described above) is to have a larger dose of advertising over a shorter period rather than a small dose over a longer period.

Limitations

We cannot be sure how the findings of this study would translate to other countries and there is no guarantee that location targeting of online advertising will continue to be available in this form. Increasing use of mobile phones may change the way location targeted adverts work. AdWords is of course not the only way of local advertising online and a definitive trial might consider use of advertising solutions from Microsoft, Facebook, LinkedIn, and others. Facebook, for example, offers location targeting and would also be worth exploring in this way.

Our study was limited by the difficulties of trying to match different data sources. Different sources from Google (AdWords and Analytics) do not exactly match due to different ways of collecting the data. Other issues, because of the anonymity of the data, include whether visitors are unique individuals. For example, although Analytics may claim to report unique visitors, we cannot verify that claim and it would be impossible for Analytics to differentiate between two individuals using the same computer (IP address) and one person with two emails using the same computer. LLTTF only collects emails of those who register. It is likely, therefore, that all sources overestimate the number of unique individuals. However, although numbers from different sources do not match exactly, the overall picture seems consistent and reasonably robust.

Overall Conclusion

This pilot study has shown that a definitive cluster trial of AdWords is worthwhile, and that this type of design could be used to assess other online recruitment interventions.

Acknowledgments

We would like to thank Ian Mayer of Mirata Ltd for help in extracting anonymised data from LLTTF, Robert Stillwell for developing the project website. RJ developed the idea for the study, was the principal investigator and grant holder of the study. RJ was responsible for day-to-day management of the project, carried out most of the analysis, wrote and edited the paper. LG contributed to the research proposal, was a co-grant holder, undertook day-to-day management, did some of the analysis, helped write and edit the paper. PH reviewed the whole study, in particular the statistical methods in the paper, suggested an alternative method to model the data, carried out that modelling, and edited the paper. CW contributed to the research proposal, was co-grant holder, and edited the paper. The project was funded by a grant from the Bupa Foundation Philip Poole-Wilson Seedcorn Fund. Ethics committee approval reference: 11-H0203-8.

Conflicts of Interest

Chris Williams is the designer and author of the LLTTF site, and is Director of Five Areas Ltd which licences the LLTTF website. LLTTF is run under licence by the charity Jubilee Trust.

Multimedia Appendix 1

Screenshots from Online Help For Depression.

[PDF File (Adobe PDF File), 314KB - jmir_v15i3e45_app1.pdf]

Multimedia Appendix 2

Two arms for definitive cluster randomized trial, showing the 16 postcode areas in each arm, cluster populations, and total population for each arm.

[PDF File (Adobe PDF File), 11KB - jmir_v15i3e45_app2.pdf]

References

1. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
2. Powell J, Clarke A. Internet information-seeking in mental health: population survey. Br J Psychiatry 2006 Sep;189:273-277 [FREE Full text] [doi: 10.1192/bjp.bp.105.017319] [Medline: 16946364]
3. Fox S. Washington, DC: Pew Internet and American Life Project Online Health Search 2006 URL: http://www.pewinternet.org/Reports/2006/Online-Health-Search-2006.aspx [accessed 2012-11-20] [WebCite Cache ID 6CJwch7eb]
4. 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]
5. Jones R, Goldsmith L. Plymouth: University of Plymouth; November. 2009. What is the evidence for the benefits and outcomes for digital health services? Final Report to NHS Choices URL: http://www.plymouth.ac.uk/researchcover/rcp.asp?page=256&pagetype=G [accessed 2011-11-04] [WebCite Cache ID 62wTB0400]
6. Depression: the treatment and management of depression in adults (update). London: National Institute for Clinical Excellence; 2009. URL: http://publications.nice.org.uk/depression-in-adults-cg90 [accessed 2013-02-22] [WebCite Cache ID 6Ed6DsGqN]
7. Beating the blues. URL: http://www.beatingtheblues.co.uk/patients/ [accessed 2012-11-20] [WebCite Cache ID 6CJwwXbMU]
8. Williams C. www.llttf.com. 2012. Living life to the full URL: http://www.llttf.com/ [accessed 2011-10-31] [WebCite Cache ID 62qfl9MlT]
9. MoodGYM. 2012. URL: https://moodgym.anu.edu.au:443/welcome [accessed 2012-11-20] [WebCite Cache ID 6CJx1zVDp]
10. Jones RB, Goldsmith L, Williams CJ, Kamel Boulos MN. Accuracy of geographically targeted internet advertisements on Google AdWords for recruitment in a randomized trial. J Med Internet Res 2012;14(3):e84 [FREE Full text] [doi: 10.2196/jmir.1991] [Medline: 22718043]
11. Office for National Statistics. 2012. Internet access quarterly update, 2012 Q2 URL: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-268719 [accessed 2012-11-20] [WebCite Cache ID 6CJx930Kp]
12. Weich S, Holt G, Twigg L, Jones K, Lewis G. Geographic variation in the prevalence of common mental disorders in Britain: a multilevel investigation. Am J Epidemiol 2003 Apr 15;157(8):730-737 [FREE Full text] [Medline: 12697577]
13. Estrada C, Krishnamoorthy P, Smith A, Staton L, Houston T. 2009 Southern Regional Meeting Abstracts. Web-analytics: changes of participation measurement URL: http://www.afmr.org/abstracts/2009/SR2009_abstracts/431.cgi [accessed 2012-11-20] [WebCite Cache ID 6CJxHuFMT]

14. Lee K, Otgonsuren M, Younoszai Z, Mir HM, Younossi ZM. Association of chronic liver disease with depression: a population-based study. Psychosomatics 2013 Jan;54(1):52-59. [doi: 10.1016/j.psym.2012.09.005] [Medline: 23295007]
15. Estrada CA, Krishnamoorthy P, Smith A, Staton L, Korf MJ, Allison JJ, et al. Marketing to increase participation in a Web-based continuing medical education cultural competence curriculum. J Contin Educ Health Prof 2011;31(1):21-27 [FREE Full text] [doi: 10.1002/chp.20097] [Medline: 21425356]
16. Ramo DE, Hall SM, Prochaska JJ. Reaching young adult smokers through the internet: comparison of three recruitment mechanisms. Nicotine Tob Res 2010 Jul;12(7):768-775 [FREE Full text] [doi: 10.1093/ntr/ntq086] [Medline: 20530194]
17. 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 [FREE Full text] [Medline: 17491170]
18. Graham AL, Fang Y, Moreno JL, Streiff SL, Villegas J, Muñoz RF, et al. Online advertising to reach and recruit Latino smokers to an internet cessation program: impact and costs. J Med Internet Res 2012;14(4):e116 [FREE Full text] [doi: 10.2196/jmir.2162] [Medline: 22954502]
19. 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]
20. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
21. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, Respondent-Driven Sampling (webRDS) strategy. J Stud Alcohol Drugs 2012 Sep;73(5):834-838. [Medline: 22846248]
22. Devineni T, Blanchard EB. A randomized controlled trial of an internet-based treatment for chronic headache. Behav Res Ther 2005 Mar;43(3):277-292. [doi: 10.1016/j.brat.2004.01.008] [Medline: 15680926]
23. 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]
24. Jones RB, O'Connor A, Brelsford J, Parsons N, Skirton H. Costs and difficulties of recruiting patients to provide e-health support: pilot study in one primary care trust. BMC Med Inform Decis Mak 2012;12:25 [FREE Full text] [doi: 10.1186/1472-6947-12-25] [Medline: 22458706]
25. Leykin Y, Muñoz RF, Contreras O. Are consumers of Internet health information "cyberchondriacs"? Characteristics of 24,965 users of a depression screening site. Depress Anxiety 2012 Jan;29(1):71-77 [FREE Full text] [doi: 10.1002/da.20848] [Medline: 21681872]
26. Purdy CH. Using the Internet and social media to promote condom use in Turkey. Reprod Health Matters 2011 May;19(37):157-165. [doi: 10.1016/S0968-8080(11)37549-0] [Medline: 21555096]
27. Gordon JS, Akers L, Severson HH, Danaher BG, Boles SM. Successful participant recruitment strategies for an online smokeless tobacco cessation program. Nicotine Tob Res 2006 Dec;8 Suppl 1:S35-S41. [Medline: 17491169]
28. Depression. NHS choices, your health, your choices URL: http://www.nhs.uk/conditions/depression/Pages/Introduction.aspx?WT.mc_id=61006 [accessed 2011-10-31] [WebCite Cache ID 62qgIBeRw]
29. Samaritans. 2011. Samaritans URL: http://www.samaritans.org/ [accessed 2011-10-31] [WebCite Cache ID 62qgVnRjR]
30. National Collaborating Centre for Mental Health.: National Institute for Clinical Excellence; 2004. Depression: The management of depression in primary and secondary care URL: http://www.nice.org.uk/nicemedia/pdf/cg023fullguideline.pdf [accessed 2012-11-20] [WebCite Cache ID 6CJxdwMuz]
31. Kroenke K, Spitzer RL. The PHQ9: a new depression diagnostic and severity measure. Psychiatric Annals 2002;32(9 Online):0048-5713.
32. National Health Service. Improving Access to Psychological Therapies URL: http://www.iapt.nhs.uk/ [accessed 2013-02-22] [WebCite Cache ID 6Ed7rOTUY]
33. R Development Core Team (2011). R: A language and environment for statistical computing URL: http://www.r-project.org/ [accessed 2012-11-20] [WebCite Cache ID 6CJxspdtE]
34. Jones RB, Goldsmith L, Hewson P, Kamel Boulos MN, Williams CJ. Do adverts increase the probability of finding online cognitive behavioural therapy for depression? Cross-sectional study. BMJ Open 2012;2(2):e000800 [FREE Full text] [doi: 10.1136/bmjopen-2011-000800] [Medline: 22508957]
35. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the Price: The cost of mental health care in England to 2026. London: King's Fund; 2008. URL: http://www.kingsfund.org.uk/sites/files/kf/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008_0.pdf [accessed 2013-02-22] [WebCite Cache ID 6Ed7JbAri]
36. Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. Psychiatric morbidity among adults living in private households, 2000. Int Rev Psychiatry 2003;15(1-2):65-73. [doi: 10.1080/0954026021000045967] [Medline: 12745312]

Abbreviations

CBT: cognitive behavioral therapy
GP: general practitioner
IAPT: Improved Access to Psychological Therapies
IP: Internet protocol
KT: Kingston
LLTTF: Living Life To The Full
NHS: National Health Service
NICE: National Institute for Health and Clinical Excellence
OHFD: Online Help For Depression
PHQ9: Patient Health Questionnaire
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 25.09.12; peer-reviewed by J Powell, J Woodford, W van Ballegooijen; comments to author 09.11.12; revised version received 21.11.12; accepted 20.01.13; published 05.03.13.

Please cite as:
Jones RB, Goldsmith L, Hewson P, Williams CJ
Recruitment to Online Therapies for Depression: Pilot Cluster Randomized Controlled Trial
J Med Internet Res 2013;15(3):e45
URL: http://www.jmir.org/2013/3/e45/
doi:10.2196/jmir.2367
PMID:23462072

Original Paper

The Effect of Program Design on Engagement With an Internet-Based Smoking Intervention: Randomized Factorial Trial

Jennifer B McClure[1], PhD; Susan M Shortreed[1], PhD; Andy Bogart[1], MS; Holly Derry[2], MPH; Karin Riggs[1], MSW; Jackie St John[1], BS; Vijay Nair[3], PhD; Larry An[2], MD

[1]Group Health Research Institute, Seattle, WA, United States

[2]University of Michigan, Center for Health Communications Research, Ann Arbor, MI, United States

[3]University of Michigan, Department of Statistics, Ann Arbor, MI, United States

Corresponding Author:
Jennifer B McClure, PhD
Group Health Research Institute
1730 Minor Ave.
Suite 1600
Seattle, WA, 98101
United States
Phone: 1 206 287 2737
Fax: 1 206 287 2871
Email: McClure.J@ghc.org

Abstract

Background: Participant engagement influences treatment effectiveness, but it is unknown which intervention design features increase treatment engagement for online smoking cessation programs.

Objective: We explored the effects of 4 design features (ie, factors) on early engagement with an Internet-based, motivational smoking cessation program.

Methods: Smokers (N=1865) were recruited from a large health care organization to participate in an online intervention study, regardless of their interest in quitting smoking. The program was intended to answer smokers' questions about quitting in an effort to motivate and support cessation. Consistent with the screening phase in the multiphase optimization strategy (MOST), we used a 2-level, full-factorial design. Each person was randomized to 1 of 2 levels of each factor, including message tone (prescriptive vs motivational), navigation autonomy (dictated vs not), proactive email reminders (yes vs no), and inclusion of personally tailored testimonials (yes vs no). The effects of each factor level on program engagement during the first 2 months of enrollment were compared, including number of visits to the website resulting in intervention content views (as opposed to supplemental content views), number of intervention content areas viewed, number of intervention content pages viewed, and duration of time spent viewing this content, as applicable to each factor.

Results: Adjusting for baseline readiness to quit, persons who received content written in a prescriptive tone made the same number of visits to the website as persons receiving content in a motivational tone, but viewed 1.17 times as many content areas (95% CI 1.08-1.28; *P*<.001) and 1.15 times as many pages (95% CI 1.04-1.28; *P*=.009). Time spent viewing materials did not differ among groups (*P*=.06). Persons required to view content in a dictated order based on their initial readiness to quit made the same number of visits as people able to freely navigate the site, but viewed fewer content areas (ratio of means 0.80, 95% CI 0.74-0.87; *P*<.001), 1.17 times as many pages (95% CI 1.06-1.31; *P*=.003), and spent 1.37 times more minutes online (95% CI 1.17-1.59; *P*<.001). Persons receiving proactive email reminders made 1.20 times as many visits (95% CI 1.09-1.33; *P*<.001), viewed a similar number of content areas as persons receiving no reminders, viewed 1.58 times as many pages (95% CI 1.48-1.68; *P*<.001), and spent 1.51 times as many minutes online (95% CI 1.29-1.77; *P*<.001) as those who did not receive proactive emails. Tailored testimonials did not significantly affect engagement.

Conclusions: Using a prescriptive message tone, dictating content viewing order, and sending reminder emails each resulted in greater program engagement relative to the contrasting level of each experimental factor. The results require replication, but suggest that a more directive interaction style may be preferable for online cessation programs.

Trial Registration: clinicaltrials.gov NCT00992264; http://clinicaltrials.gov/ct2/show/NCT00992264 (Archived by WebCite at http://www.webcitation.org/6F7H7lr3P)

(J Med Internet Res 2013;15(3):e69) doi:10.2196/jmir.2508

KEYWORDS

smoking cessation; utilization; Internet; behavioral research; electronic mail; motivation

Introduction

Smoking remains a leading cause of death and disability, accounting for approximately 1 in 5 deaths each year in the United States [1]. Effective population-based interventions are critically needed to reduce smoking prevalence and to lessen the detrimental impact of nicotine dependence. The Internet offers many advantages for this, including broad reach, low-cost treatment dissemination, and the ability to highly personalize content to be most appealing and best meet the needs of individual smokers while at the same time standardizing the content delivery across individuals (ie, to deliver personally tailored content). Based on simulation models, effective Internet-based interventions have extraordinary potential to decrease population-level smoking rates [2]. However, recent empirical reviews point out that there is only moderate evidence for the effectiveness of Internet-based cessation programs at this time [3,4]. Differences in effectiveness could be related to differences in the content or the design of existing interventions, both of which interact to dictate participants' level of engagement with the program. Engagement has been defined as the number of site visits, number and type of pages viewed, or duration of time spent viewing the content [5-7].

Although greater program engagement does not automatically mean a program is more effective (in fact, people may not return to the program because it was effective in helping them change their behavior), some level of intervention exposure is clearly important for an intervention to have its intended effect. Research has consistently shown a dose-response effect for smoking cessation interventions, including Internet-based programs [3-5,8-12], and engagement with specific components of online programs can predict cessation [10,12,13]. But it is unclear how best to promote engagement in online nicotine dependence treatment programs in which intervention exposure is left up to the self-direction and motivation of the individual user. Evidence supports the importance of message source and the level of personal tailoring on the number of intervention pages viewed in online smoking cessation interventions [5]. Additional insight can be gleaned from studies evaluating online lifestyle modification programs. For example, supplemental email prompts can increase return website visits [14] and promote greater online self-monitoring of behavioral risk factors [15]. Others have suggested that limiting users' control over their navigation of a website can increase time spent on the website and the number of pages visited [16]. In general, however, little is known about how to best design an Internet-based smoking cessation program to maximize participant engagement, particularly when the program is designed for use on a population level, among all smokers, regardless of their current interest in quitting. The current study addresses this issue.

Consistent with the initial screening phase of the multiphase optimization strategy (MOST) for treatment development [17,18], we implemented a 2-level full-factorial experiment to examine the effects of 4 independent design factors (message tone, navigation autonomy, proactive email outreach, and inclusion of personally tailored testimonials) on participant engagement during the first 2 months of program enrollment, with each factor explored on 2 contrasting levels. We chose to focus on the first 2 months after program enrollment because we hypothesized this to be a critical time for treatment engagement. That is, participants may be more likely to interact with the intervention shortly after joining the program, reflecting their initial motivation to participate. Future analyses will report on the long-term effects of each design factor on smoking abstinence and treatment utilization, the main outcomes for this randomized trial.

For this study, *engagement* was defined as the number of times people visited the website to view the intervention content, the number of content areas viewed, the number of content pages viewed, and the duration of time spent viewing the content. This definition is consistent with the literature [5-7] and reflects the fact that engagement is multidimensional. For instance, increased content exposure (in terms of total page views or content areas viewed) should increase one's duration of exposure, but could also reduce the absolute number of visits if people feel they have maximized their interaction with the website. Thus, it is important to examine each of these measures separately.

Each design factor was chosen based on empirical or theoretical evidence for its effects on smoking cessation or because its treatment effects are unclear. For example, research suggests that interventions grounded in the principles of motivational interviewing can be effective across a range of health risk behaviors, including smoking abstinence [19-24]. Dictating content order based on readiness to quit may also increase treatment effectiveness by making treatment information more salient to smokers. Narrative testimonials can transport readers [25] and may result in greater behavior change [26]. In fact, personally tailored testimonials were associated with higher 6-month abstinence rates in prior research [27]. Finally, periodic email reminders may encourage greater program utilization [14,15] and, therefore, enhance treatment outcome (for further discussion of the rationale for the selection of these factors, see McClure et al [28]).

In the current study, we hypothesized that each of the experimental factor levels would also have a differential effect on our 4 measures of engagement. Participants randomized to receive online intervention content written in a prescriptive tone (as opposed to a motivational tone) would find the content less acceptable; therefore, they would view fewer content areas and Web pages, spend less time reviewing the content, and may return to the site less often. Similar hypotheses were made for people who were required to view content in a prespecified (dictated) order based on their stage of change, as opposed to being allowed to navigate the site freely, based on their interests.

These assumptions are consistent with people's desire for autonomy as described in self-determination theory [29-31]. We also believed that people who received periodic email prompts encouraging a return to the site would visit the website more often and spend more time viewing content as a result. It was unknown if they would view more treatment content areas or Web pages since exposure to the content could be maxed out during the initial visit. Finally, we explored the impact of providing smokers with personally tailored testimonials from other smokers as part of their intervention. This type of narrative is a common technique in persuasive messaging and can facilitate information processing, provide surrogate social connections, overcome resistance, and address emotional issues—all potentially important to behavior change [26]. Because the addition of the testimonials confounded our ability to examine its effects on the total number of content page views or duration of exposure (because these participants had additional content pages to view), we were only able to examine its effects on total content areas viewed and visits to the website. Findings from this study add to the nascent literature informing the optimal design of Internet-based behavior change programs to encourage program engagement.

Methods

The study design and methods, including an extensive overview of each of the experimental factor choices, their theoretical rationale, and how each was operationalized in the Questions about Quitting (Q2) intervention is available elsewhere [28]. Details and information about the trial specific to the current hypotheses are summarized subsequently.

Setting and Population

This study was a collaboration between Group Health Research Institute in Seattle, Washington and the University of Michigan Center for Health Communications Research in Ann Arbor, Michigan. Participants were recruited from Group Health, a large, regional nonprofit health plan in Washington State. All research materials (intervention materials, surveys, and protocols) were approved by the institutional review boards at Group Health Research Institute and the University of Michigan. The study is registered with clinicaltrials.gov (NCT00992264). Data reported in this paper were collected between May 2010 and December 2011.

Factorial Design and Screening Experiment

Consistent with the initial phase of the MOST framework [18,32,33], we conducted a 2-level full-factorial experiment to screen for optimum intervention characteristics. Half of the participants were exposed to each contrasting level of the 4 experimental factors: message tone (prescriptive vs motivational), navigation autonomy (dictated vs not), proactive email reminders (yes vs no), and inclusion of personally tailored testimonials (yes vs no). Randomization to each factor was balanced across the trial arms to control for their effects on each factor of interest and stratified by baseline readiness to quit smoking. Interested readers are referred to McClure et al [28] for a more detailed discussion of the factorial design. Additional discussion of the MOST methodology can be found in the literature [17,18,32,33]. The long-term goal of this programmatic research will be to combine the most effective factors to create an optimized intervention and compare it to an empirically validated control in a future randomized trial.

Recruitment, Screening, Randomization, and Enrollment

A study invitation letter was sent to adult likely smokers identified from automated health plan records. The study program was described as providing information and guidance to help people decide if quitting was right for them and how to quit if and when they decided to do so. The goal was to recruit smokers interested in quitting, as well as those with no interest in quitting.

Individuals interested in learning more about the program were provided a unique log-in access code in the invitation letter and were directed to the study website where they were screened for eligibility, provided consent, and were enrolled online. People were eligible if they were aged 18 years or older, a current member of Group Health, smoked 100 cigarettes in their lifetime, smoked even a puff in the past 7 days, smoked an average of at least 5 cigarettes per day, were not currently enrolled in a smoking cessation program or taking medication to stop smoking, had access to the Internet for personal use, were willing to check their email at least once a week, were comfortable reading and writing in English, had no visual impairments that prevented reading text on a computer screen, and were comfortable using a computer and the Internet.

After providing online consent, participants completed a baseline assessment online and then were randomized to an intervention arm using an automated algorithm. Half of all participants were randomized to each contrasting factor level and assignment to each intervention group was stratified by participants' readiness to quit smoking at baseline (no interest in quitting in the next 6 months, interested in quitting in the next 6 months but not the next month, or interested in quitting in the next month). Following randomization, participants could immediately access their personalized intervention program following the baseline assessment and were encouraged to return to the site as often as they wanted. Because enrollment required log-in using a preassigned log-in code, it was not possible for participants to enroll in the study more than once. Participants were blinded to their group assignment.

Program Development

The program was developed through an iterative and interactive design process. The final design and layout was informed by focus group testing with smokers. Intervention content was written by experts in behavioral science at the Group Health Research Institute and University of Michigan's Center for Health Communications Research (CHCR). The personalized intervention content was tailored using the nonproprietary Michigan Tailoring System, developed by researchers in the CHCR. Additional detail on the program design and content are available in McClure et al [28]. There were no major changes to the intervention design or content after study launch.

Program Design and Core Intervention Content

The intervention was delivered via the Internet. Participants were told they would receive an individually tailored program designed to answer their questions about quitting smoking and to help them make a decision about whether and how to stop smoking, but they were not told any specifics about the treatment arms before or after accessing the intervention. The intervention included a combination of core intervention content and additional special feature content. The core content was accessible from the main page and organized in 3 main drop-down headers or content areas, each targeting smokers at different stages of readiness to quit smoking, specifically those not ready to quit, those ready to quit, and those who already quit (see sample screenshot in Figure 1). Each of the 3 core content areas contained 3 to 5 subsections set up as individual Web pages. Section subheadings reflected questions smokers commonly have (eg, Is quitting right for me? What are my treatment options?). The special features content was also linked to the main page, but was kept distinct from the core content section. This supplemental material included topics other than smoking cessation, such as stress management, time management, and physical activity—topics thought to have a broad appeal to smokers regardless of their interest in quitting smoking and which would, therefore, encourage return visits to the website.

Participants could view the Q2 program as often as they liked and they were encouraged to return to the website in the future. Upon return, if more than 24 hours had elapsed since their last visit, participants were asked to restate their readiness to quit smoking and the content was retailored to reflect their current smoking status and interest in quitting. The basic intervention layout, number of pages, and substantive core content remained unchanged, but the text was refreshed to reflect the change in participants' current motivation for quitting or smoking status. The intent was to ensure that the program content remained responsive to individuals' current needs.

Figure 1. Example screenshot of Questions about Quitting (Q2) layout.

Experimental Factors

Message Tone

Participants were randomized to receive content written in either a prescriptive or motivational tone. Prescriptive messaging was written in a didactic tone and clearly advised smokers to quit smoking and how to achieve this goal. Motivational messaging was written in a tone consistent with the key principles of motivational interviewing (express empathy, develop discrepancy, roll with resistance, support autonomy and self-efficacy) [34]. Messages written in this tone recognized smokers' potential ambivalence about quitting and their autonomy in making decisions about if, how, and when they would quit smoking.

Navigation Autonomy

Half of the participants could freely view content on the website in any order they wished. The other half of the participants, in the dictated navigation arm, were required to first view content matched to their baseline readiness to quit and to view the content in a prespecified order. After this content was seen, they were then free to navigate the site.

Proactive Emails

Participants were also randomized to receive weekly proactive email reminders or not. Email messages were standardized across all individuals and encouraged participants to return to the Q2 website to view the optional special feature content. However, we did not track special feature page views because it was not part of the core intervention. Additionally, not all participants had access to this content at the same time. For those whose navigation of the site was dictated based on their initial readiness to quit, access to this optional content was only available after they viewed all Web pages in their initial stage-appropriate content area.

Testimonials

Participants were randomized to receive 3 highly tailored testimonials designed to promote their self-efficacy for quitting or to not receive these testimonials. Testimonials were tailored on each individual's stage of change, level of nicotine dependence, prior use of pharmacotherapy for nicotine dependence, depression history, perceived risks and benefits of quitting smoking, and their self-efficacy for quitting. One testimonial was included at the end of each of the 3 core content sections. Testimonials were designed to support self-efficacy for quitting by providing personally tailored information and modeling appropriate quitting behaviors. Information was presented in an interview format with a smoker or former smoker. Because not all participants received the extra testimonial content, tracking data from these pages, including time spent viewing this content, were excluded from the analyses to normalize the engagement metrics across treatment arms.

Assessment and Measures

Self-report data were collected at baseline using an online survey. This data included demographics; current smoking status; number of cigarettes smoked per day; stage of readiness to quit smoking; nicotine dependence assessed with the Fagerstrom Test of Nicotine Dependence [35]; motivation for quitting, self-efficacy for quitting, and perceived importance of quitting smoking (assessed on a 10-point Likert scale ranging from not at all to very); and use of alternative treatments for quitting smoking (eg, pharmacotherapy, counseling).

Automated tracking data were collected each time participants visited the website. This data included the date and time each participant visited the website and individual date/time stamps each time a content page was accessed or left.

Intervention exposure was defined as any exposure to the core intervention content. Visits to the main page and special features were excluded. Engagement was defined by: (1) the number of unique visits to the website during which the core intervention content was viewed, (2) the number of unique treatment core content sections viewed (out of a possible 3), (3) the number of times individual pages (core content subsections) were viewed, and (4) the cumulative duration of minutes spent viewing the core intervention content. Sessions automatically timed out after 30 minutes of inactivity or ended when individuals left the website (eg, logged out, closed their browser, or visited a different website).

Data Integrity

Data were monitored over the course of the study to ensure participants were appropriately randomized, baseline data were collected, and automated user statistics on program use were being appropriately captured.

Analyses

The analytic sample included all individuals, regardless of exposure to the core intervention content, to take advantage of the balancing effect of randomization on all covariates, measured and unmeasured. As a result, any observed differences between the randomized treatment groups are because of differences in the effect of the interventions. If we had limited the analyses to only those individuals who observed some content, this would restrict the sample based on a posttreatment outcome. That is, differences in observed program engagement levels between groups could be due to treatment effects or imbalances in covariates between the treatment groups. To further complicate matters, 1 of the factor levels (receipt of proactive emails) directly affected the probability that an individual viewed any content; limiting the analytic sample only to those who saw some content when comparing the 2 levels of this factor would bias the results.

Descriptive statistics were used to characterize the study sample based on data collected during the baseline survey. To assess engagement with the website content, we examined website-tracking data for each participant. We calculated means, standard deviations, medians, and interquartile ranges for each count-based outcome measure. We compared the number of visits to the website in which an individual viewed core intervention content between the 2 levels of each of the 4 factors using Poisson regression models that adjusted for initial readiness to quit smoking because this was a stratification variable in the randomization process. Similar Poisson models were used to estimate the effect of random factor level assignment on the number of content areas visited, ranging from zero to 3. Estimates obtained from Poisson models are generally

interpretable as incidence rate ratios, but in the context of an experiment like ours in which all subjects shared a common period of exposure, estimates can equivalently be interpreted as the ratio of mean event counts among the exposed to that of the unexposed group.

The distributions of the number of individual page views and of the cumulative number of minutes spent viewing intervention content each had a larger proportion of zeros than expected from a Poisson distribution. Due to the inflated number of zeros we used zero-inflated Poisson (ZIP) models to estimate the effects of the factors on these 2 measures [36,37]. A ZIP model is made up of 2 parts: a logistic model that is used to model the excess zeros in the population and a Poisson model used to model the mean of the outcome. In this analysis, the logistic portion of each ZIP model used only an intercept to model the excess zeros for 3 of the factors. The model for the fourth factor, receiving proactive emails, included a parameter to estimate the effect of email receipt on the odds of an excess zero. The estimates reported are the effect of the factor level on the mean of the outcome (accounting for excess zeros in the corresponding Poisson distribution) in the whole population (ie, not just those who viewed the core intervention content) as described in Preisser et al [37]. No other covariate adjustments were made in the logistic portion of our ZIP models.

A total of 683 page views timed out automatically after 30 minutes of inactivity. Among page views that did not time out, most views were significantly shorter than 30 minutes, suggesting it was unlikely that all timed-out sessions truly reflected 30 minutes of time spent viewing these pages. Thus, we treated the true viewing time for these page views as missing values and imputed the viewing time for these page views using a chained equation, multiple imputation procedure [38,39]. Model predictors included baseline data (participant demographics, smoking history, beliefs about smoking, and readiness to quit), randomized level for each of the 4 factors, and the number of minutes spent on the first core content page viewed. We estimated and tested the effects of the experimental factors on the cumulative duration of intervention time by combining results from 5 imputed datasets, accounting for both within- and between-imputation variance components [40,41].

To investigate whether the effects of the random factor assignments may have differed by initial readiness to quit, we refitted each of the regression models described previously with the inclusion of interaction terms between the factors of interest and a categorical variable indicating initial readiness to quit smoking. Joint tests of the set of interaction terms within each model fit were conducted using a Wald test statistic with 2 degrees of freedom calculated to assess the significance of interactions.

Tracking data management was conducted using SAS software version 9.2 (SAS Institute, Inc, Cary, NC, USA) and all analyses, including multiple imputations, were conducted using Stata version 12 (StataCorp LP, College Station, TX, USA).

Results

Participants

Demographic characteristics of the enrolled sample (N=1865) are presented in Table 1. The characteristics of participants within each of the 4 factors' levels were similar to one another and to the overall distribution, so only the overall distribution is shown. Recruitment flow is presented in Figure 2. Reasons for ineligibility were not mutually exclusive. Participants could report more than 1 reason for ineligibility.

Intervention Exposure and Engagement

Intervention Exposure

A total of 690 of 1865 enrolled participants (37.00%) failed to view any of the core intervention content within 2 months after joining the study, whereas 1175 participants (63.00%) viewed at least some core content during this period. Participants who failed to view any core content differed significantly with regard to their baseline readiness to quit (P<.001). More of these individuals had no interest in quitting smoking (15.22% vs 10.98%) or were interested in quitting in the next 6 months (46.52% vs 42.13%), but fewer were interested in quitting within the next month (38.26% vs 46.89%) indicating fewer were ready to quit smoking at baseline. Among those individuals who chose to view the core intervention, the proportion of people viewing content was similar across each factor level: message tone (64.91% prescriptive vs 61.09% motivational), navigation autonomy (64.45% dictated vs 61.55% nondictated), email reminders (63.88% yes vs 62.12% no), and testimonials (61.52% yes vs 64.48% no).

Program Engagement

Participants viewed the core content on a total of 1691 separate visits, resulting in 6592 unique content page views. On average, participants who accessed the core intervention made 1.4 visits (median 1, range 1-11) to view this content, viewed an average of 1.4 of the 3 core content areas (median 1, range 1-3), and viewed on average 5.6 total core content pages (median 4, range 1-53). After imputing duration of timed-out visits, the average cumulative time accrued viewing the core intervention content was 12.3 minutes (median 7.0, range 0.10-180). Of the 3 core content areas, the pages designed for people ready to quit were viewed most often. Content designed for people not yet ready to quit was viewed second most often, followed by the content designed for people who have already quit.

Engagement outcomes by factor level are presented in detail in Table 2. Effect estimates shown represent the ratio of means for each outcome measure, comparing those randomized to the stated factor level to those randomized to the contrasting factor level. For example, after adjustment for baseline readiness to quit, the average number of website visits among those who received content written in a prescriptive tone was approximately the same as the average number of visits among those whose content was written in a motivational tone, yielding a ratio of means of 1.00 (95% CI 0.90-1.10; P=.93). However, those viewing content in a prescriptive tone viewed an average of 1.17 times more content areas (95% CI 1.08-1.28; P<.001), and 1.15 times more content pages (95% CI 1.04-1.28; P=.009)

than those whose content was written in a motivational tone. Duration of time spent viewing materials did not differ statistically between the 2 levels of the tone factor (ratio of means 0.87, 95% CI 0.75-1.01; *P*=.06). Persons receiving proactive email reminders had an average of 1.20 times as many website visits resulting in content views (95% CI 1.09-1.33; *P*<.001), but visited a similar number of content areas as persons receiving no reminders. Individuals with proactive email reminders viewed 1.58 times as many content pages (95% CI 1.48-1.68; *P*<.001), and spent 1.51 times as many minutes online (95% CI 1.29-1.77; *P*<.001). Persons required to view content in a dictated order based on their initial readiness to quit made approximately the same average number of visits as people able to freely navigate the site, but viewed fewer content areas on average (ratio of means 0.80, 95% CI 0.74-0.87; *P*<.001), viewed 1.17 times as many pages (95% CI 1.06-1.31; *P*=.003), and spent 1.37 times as many minutes online (95% CI 1.17-1.59; *P*<.001). There were no significant differences in the average number of visits to the website or content areas viewed between participants who did and did not receive the personally tailored testimonials.

Table 1. Baseline characteristics of enrolled participants[a] (N=1865).

Characteristics	Participants
Sex, n (%)	
Female	1178 (63.16)
Race/ethnicity, n (%)	
White, non-Hispanic	1534 (82.25)
Education level, n (%)	
High school or less	524 (28.10)
Some college	944 (50.62)
College degree or higher	396 (21.23)
Employment status, n (%)	
Employed	1287 (69.00)
Marital status, n (%)	
Married/partnered	1052 (56.41)
Readiness to quit, n (%)	
In next 30 days	815 (43.70)
In next 6 months, but not in next 30 days	816 (43.75)
Not thinking of quitting	234 (12.55)
Years smoked, mean (SD)	24.9 (14.2)
Age (years), mean (SD)	44.2 (14.7)
Nicotine dependence (FTND)[b], mean (SD)	4.2 (2.2)
Psychosocial factors (range 1-10), mean (SD)	
Motivation for quitting	7.4 (2.5)
Self-efficacy for quitting	5.4 (2.6)
Importance of quitting	7.6 (2.6)

[a] Complete data were available on all baseline outcomes, with 1 missing value each for race and education.

[b] FTND: Fagerstrom Test of Nicotine Dependence.

Figure 2. Recruitment flow and allocation of participants.

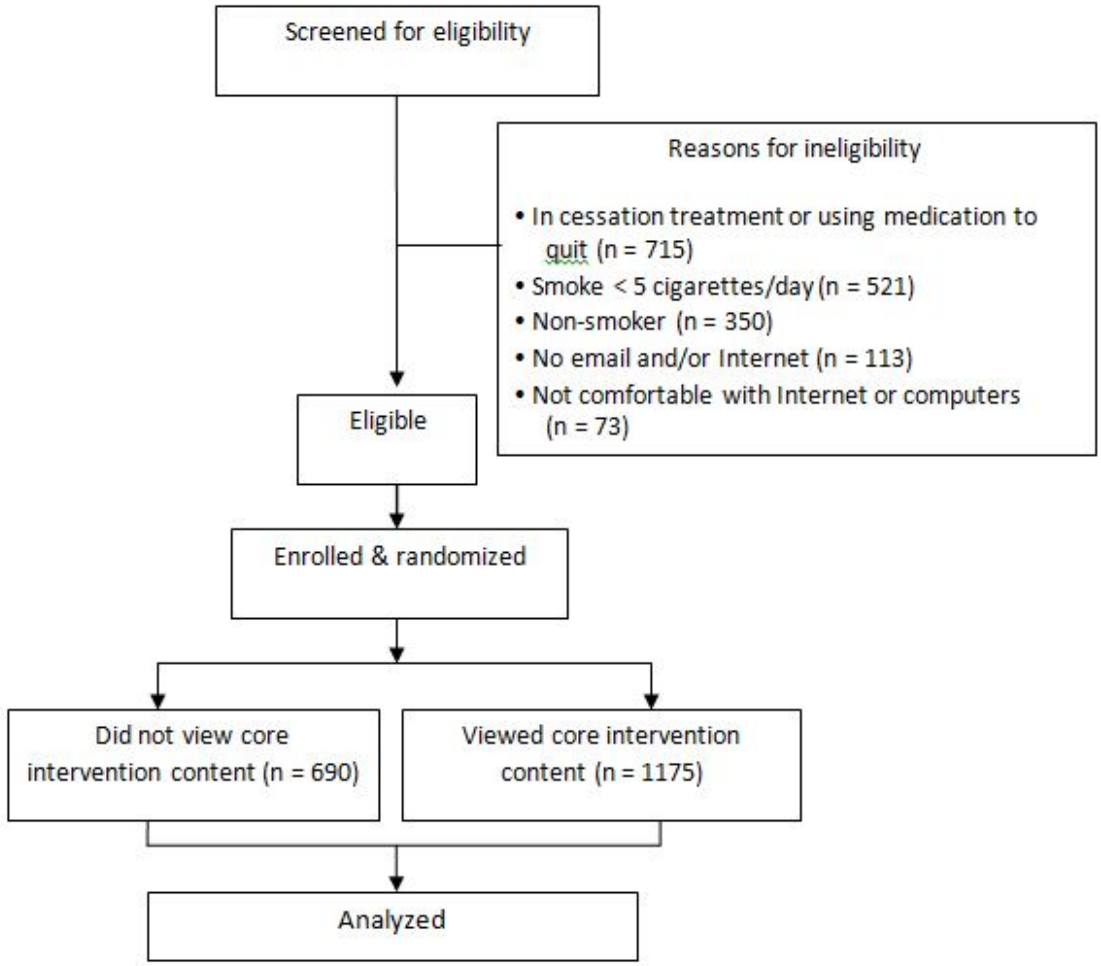


Table 2. Comparison of engagement metrics by factor level.[a]

Factor level	Prescriptive message tone		Proactive emails		Dictated navigation		Tailored testimonial[b]	
	Ratio of means (95% CI)	*P*	Ratio of means (95% CI)	*P*	Ratio of means (95% CI)	*P*	Ratio of means (95% CI)	*P*
Website visits	1.00 (0.90-1.10)	.93	1.20 (1.09-1.33)	<.001	1.02 (0.92-1.13)	.70	0.92 (0.83-1.02)	.11
Content areas viewed	1.17 (1.08-1.28)	<.001	1.08 (0.99-1.17)	.10	0.80 (0.74-0.87)	<.001	0.96 (0.88-1.04)	.30
Content page views	1.15 (1.04-1.28)	.009	1.58 (1.48-1.68)	<.001	1.17 (1.06-1.31)	.003		
Cumulative duration	0.87 (0.75-1.01)	.06	1.51 (1.29-1.77)	<.001	1.37 (1.17-1.59)	<.001		

[a] Point estimates represent ratio of means between each contrasting factor level and are adjusted for baseline stage of change. Results reflect the effect of each factor level (prescriptive tone, email reminders, dictated navigation, and testimonials) relative to those who did not receive the stated factor. Effects of randomized factors on website visits and content areas viewed were estimated with Poisson regression models, and effects on page views and duration were estimated using zero-inflated Poisson regression models.

[b] Content page views and duration of time spent viewing content were not examined for those in the testimonial factor because these individuals had more content pages to view containing more material.

Secondary analyses investigated the interaction between baseline readiness to quit (a measure of motivation) and each of the factor levels, to determine if participants with different levels of readiness to quit at enrollment engaged differently with the core Q2 program (results not shown). Out of 17 tests for interaction, only 1 was statistically significant; the prescriptive tone resulted in significantly less cumulative viewing time (ratio of means 0.54, 95% CI 0.36-0.83) among those with no interest in quitting in the next 6 months, but had no significant effect on the viewing times of those in interested in quitting in the next month (ratio of means 0.83, 95% CI 0.66-1.04) or next 6 months (ratio of means 1.03, 95% CI 0.82-1.29). These 3 estimates, each specific to a level of readiness to quit smoking, differed significantly from one another (P=.02). No adjustments were made for multiple comparisons.

Discussion

Program engagement is critical for any intervention to be effective, but promoting program engagement is a particularly important issue in Web-based interventions because treatment exposure is dependent on the motivation and self-direction of the individual user. In order to maximize the effectiveness of future Internet-based smoking treatment programs, we need a better understanding of how to engage smokers in these programs and, in particular, how to promote engagement with the most critical core program elements designed to motivate and promote behavior change. The current study provides insight into these issues by comparing the relative effects of 2 contrasting levels of each experimental design feature (factor): message tone (prescriptive vs motivational), navigation autonomy (dictated vs not), proactive email reminders (yes vs no), and inclusion of personally tailored testimonials (yes vs no). We sought to determine if one level promoted greater treatment engagement than the other within each factor.

We found that using a prescriptive message tone, dictating the order content was viewed, and sending email prompts had the greatest effects on early program engagement among a population-based sample of smokers at varying stages of readiness to quit smoking. Each of these increased the total number of core page views. Cumulative exposure to the core content was also increased by dictating navigation order and sending emails. Although the prescriptive tone was not statistically significant at the .05 level, the effect estimate was greater than 1 (P=.06). Using a prescriptive tone also increased the total number of core content areas viewed and, as expected, email prompts increased the number of visits to the website. No other factors increased the number of content areas viewed or number of visits made to view the core intervention content.

The inclusion of tailored testimonials did not have an effect on Web visits or the number of content areas viewed, although we did not expect it would. The primary goal of this factor was to promote smoking cessation through enhancing self-efficacy and modeling appropriate cessation-relevant behaviors, so its real impact is expected to be observed in future analyses examining long-term cessation and treatment utilization (the primary study outcomes). Also, persons whose navigation autonomy was dictated saw fewer total content areas, which was expected because they had to view all content in their stage-matched content area before gaining access to the other 2 content areas. This barrier likely deterred exposure to more content areas.

The findings suggest that using a prescriptive message tone, dictating the order in which content is viewed to match smokers' initial level of interest in quitting, and sending weekly email prompts may increase online program engagement. This directive approach is somewhat counterintuitive for an intervention intended to motivate persons to quit smoking. Motivational interviewing suggests that people who are not ready to take action may respond better to counseling which is less directive and recognizes their ambivalence for change and autonomy to make their own decisions [34,42]. We cannot yet comment on the impact of each factor level on abstinence (the true measure of how well people respond to a cessation intervention), but in terms of program engagement, the more directive factor levels were preferable. This was unexpected for the dictated navigation, but is consistent with recently published research demonstrating that limiting user control over navigation increased time spent online and page visits within a website designed to promote hepatitis knowledge [16]. It is unclear why the prescriptive tone had a differential effect on engagement than the motivational tone—counter to what would be predicted based on self-determination theory. To gain insights into this finding we looked for differences in users' acceptability ratings, literacy, or self-reported desire to be "told what to do" by a clinician (data not presented). These data did not reflect differences among the randomization arms that might explain our findings. The most likely explanation at this time is that smokers seeking information about whether and how to quit smoking simply prefer more directive advice. Whether this finding will generalize to other audiences or topics should be explored further.

It is noteworthy that one-third of participants failed to view any of the core intervention content during the first 2 months of the study. The reason for this is not evident, but motivation for quitting could play a role. Overall, people who failed to view core content were less likely to be ready to quit smoking in the next 30 days compared to those who viewed the content. Future planned analyses will explore how those who viewed the intervention differed from those who did not and whether these individuals failed to ever view the core intervention content or simply delayed their viewing. All participants have access to the Q2 program for a full year after enrollment.

Several caveats should be considered when interpreting these results. First, the findings might look different if we had examined tracking data for the testimonial pages and special features. However, including these would inappropriately skew the effects of the factor levels on engagement because exposure to these program elements varied by treatment arm. Also, the special feature content was not considered part of the core smoking cessation intervention. Thus, our more conservative approach is justified. The findings might also look different if engagement was observed over a longer period of time. We chose to measure engagement over the first 2 months of enrollment because this would seem to be a critical time. If one fails to engage with the program within 2 months after making an effort to enroll, it may be that they will not engage at all. We

will be able to address this in future analyses when 1-year follow-up data are available. Next, the average cumulative exposure duration would be higher if we had not imputed missing values for each page view that timed out after 30 minutes of inactivity. However, we believe it preferable to treat this information as missing and use multiple imputations to accommodate this missing data in the analyses than to potentially overestimate this important outcome and artificially inflate exposure. Additionally, although it is tempting to interpret the interaction results as evidence that the prescriptive tone was less effective among people with no interest in quitting smoking, caution must be used in drawing this conclusion since this was the only significant interaction out of 17 and we did not adjust for multiple comparisons. Finally, we should caution readers not to interpret the results as an evaluation of motivational interviewing per se, which is a specific counseling technique. We can only comment on the application of several key principles of motivational interviewing when applied in a Web-based program not the full complement of motivational interviewing skills, which would be difficult to simulate outside an actual counseling session. Thus, we consider this an evaluation of a motivational message tone grounded in motivational interviewing principles.

There are limitations with this study. For one, it is not clear if the results will generalize to other Internet-based treatment programs since engagement is associated with the specific content of an intervention. But because we focused on design principles such as message tone, navigation autonomy, and use of proactive emails, it will be possible for others to apply these same strategies to future programs and test their effects. Similarly, we do not know if the results will generalize to other smokers, particularly uninsured minority males. All smokers in the current study had medical insurance (at least at the time of enrollment), most were female (63%), and most were white (82%). However, enrolling a higher proportion of female and white smokers is consistent with findings from other population-based, online cessation trials [43-45].

The study has several distinct strengths. Chief among these, the study systematically explores how the design of a public health smoking intervention influences smokers' interaction with the program. Other strengths include the large study sample (N=1865) which included a broad spectrum of smokers with differing levels of motivation to quit, use of a rigorous study design grounded in the MOST methodological framework, use of automated tracking data to confirm individual exposure to the website at the level of each individual Web page and time spent viewing specific pages, and use of sophisticated imputation methods to account for time spent online without overinterpreting cumulative exposure time based on time-out parameters.

The results of the current study provide important insight about how to design a population-based, online smoking cessation intervention. Ultimately, it will be important to see what effect each of the experimental factors has on long-term cessation outcomes, but the current study suggests that taking a directive intervention approach, including a prescriptive message tone, dictated site navigation, and proactive email outreach may be useful for increasing program engagement particularly in population-based interventions targeting smokers with varying levels of motivation for quitting. Future research should seek to replicate these findings. Moreover, more methodologically rigorous science should seek to systematically elucidate the optimal strategies for maximizing the effectiveness of online behavioral intervention programs.

Acknowledgments

This research was funded by the National Cancer Institute (R01 CA138598, J McClure, PI). We are grateful to the contributions of the many study team members at Group Health Research Institute and the University of Michigan. This includes the Group Health Research Institute Survey Research Program, Melanie Mamiya, Roy Pardee, and Rebecca Hubbard for their assistance designing and fielding the study, and Annie Shaffer for her assistance with the manuscript preparation. We also thank the Center for Health Communications Research at the University of Michigan, including Kate Barr, Janine Konkel, Mike Nowak, Ian Moore, Sarah Pennewell, and Ed Saunders for their contributions to the development and conduct of the Q2 trial. The intervention evaluated in this study was developed by researchers at the Group Health Research Institute and University of Michigan.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [46].

[PDF File (Adobe PDF File), 1007KB - jmir_v15i3e69_app1.pdf]

References

1. Centers for Disease Control and Prevention (CDC). Annual smoking-attributable mortality, years of potential life lost, and economic costs--United States, 1995-1999. MMWR Morb Mortal Wkly Rep 2002 Apr 12;51(14):300-303 [FREE Full text] [Medline: 12002168]

2. Levy DT, Graham AL, Mabry PL, Abrams DB, Orleans CT. Modeling the impact of smoking-cessation treatment policies on quit rates. Am J Prev Med 2010 Mar;38(3 Suppl):S364-S372 [FREE Full text] [doi: 10.1016/j.amepre.2009.11.016] [Medline: 20176309]
3. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: Web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]
4. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
5. 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]
6. 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]
7. 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]
8. Fiore MC, Bailey WC, Cohen SJ, Dorfman SF, Goldstein MG, Gritz ER. Clinical practice guideline: Treating tobacco use and dependence. Rockville, MD: US Department of Health and Human Services. Public Health Service; 2000 Jun. URL: http://www.treatobacco.net/en/uploads/documents/Treatment%20Guidelines/USA%20treatment%20guidelines%20in%20English%202000.pdf [accessed 2012-12-18] [WebCite Cache ID 6D0aNLCHf]
9. Zbikowski SM, Jack LM, McClure JB, Deprey M, Javitz HS, McAfee TA, et al. Utilization of services in a randomized trial testing phone- and web-based interventions for smoking cessation. Nicotine Tob Res 2011 May;13(5):319-327 [FREE Full text] [doi: 10.1093/ntr/ntq257] [Medline: 21330267]
10. 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]
11. Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]
12. Richardson A, Graham AL, Cobb N, Xiao H, Mushro A, Abrams D, et al. Engagement promotes abstinence in a Web-based cessation intervention: cohort study. J Med Internet Res 2013;15(1):e14 [FREE Full text] [doi: 10.2196/jmir.2277] [Medline: 23353649]
13. Schwarzer R, Satow L. Online intervention engagement predicts smoking cessation. Prev Med 2012 Sep;55(3):233-236. [doi: 10.1016/j.ypmed.2012.07.006] [Medline: 22813919]
14. Schneider F, van Osch L, Schulz DN, Kremers SP, de Vries H. The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program. J Med Internet Res 2012;14(2):e40 [FREE Full text] [doi: 10.2196/jmir.1939] [Medline: 22382037]
15. Greaney ML, Sprunck-Harrild K, Bennett GG, Puleo E, Haines J, Viswanath KV, et al. Use of email and telephone prompts to increase self-monitoring in a Web-based intervention: randomized controlled trial. J Med Internet Res 2012;14(4):e96 [FREE Full text] [doi: 10.2196/jmir.1981] [Medline: 22842775]
16. Crutzen R, Cyr D, de Vries NK. The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. J Med Internet Res 2012;14(2):e45 [FREE Full text] [doi: 10.2196/jmir.1922] [Medline: 22532074]
17. 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]
18. 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 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
19. Burke BL, Arkowitz H, Menchola M. The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. J Consult Clin Psychol 2003 Oct;71(5):843-861. [doi: 10.1037/0022-006X.71.5.843] [Medline: 14516234]
20. Heckman CJ, Egleston BL, Hofmann MT. Efficacy of motivational interviewing for smoking cessation: a systematic review and meta-analysis. Tob Control 2010 Oct;19(5):410-416 [FREE Full text] [doi: 10.1136/tc.2009.033175] [Medline: 20675688]
21. Hettema JE, Hendricks PS. Motivational interviewing for smoking cessation: a meta-analytic review. J Consult Clin Psychol 2010 Dec;78(6):868-884. [doi: 10.1037/a0021498] [Medline: 21114344]
22. Lundahl B, Burke BL. The effectiveness and applicability of motivational interviewing: a practice-friendly review of four meta-analyses. J Clin Psychol 2009 Nov;65(11):1232-1245. [doi: 10.1002/jclp.20638] [Medline: 19739205]

23. Vasilaki EI, Hosier SG, Cox WM. The efficacy of motivational interviewing as a brief intervention for excessive drinking: a meta-analytic review. Alcohol Alcohol 2006;41(3):328-335 [FREE Full text] [doi: 10.1093/alcalc/agl016] [Medline: 16547122]
24. Hettema J, Steele J, Miller WR. Motivational interviewing. Annu Rev Clin Psychol 2005;1:91-111. [doi: 10.1146/annurev.clinpsy.1.102803.143833] [Medline: 17716083]
25. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000;79:701-721. [Medline: 11079236]
26. Kreuter MW, Green MC, Cappella JN, Slater MD, Wise ME, Storey D, et al. Narrative communication in cancer prevention and control: a framework to guide research and application. Ann Behav Med 2007 Jun;33(3):221-235. [doi: 10.1080/08836610701357922] [Medline: 17600449]
27. 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 [FREE Full text] [doi: 10.1016/j.amepre.2007.12.024] [Medline: 18407003]
28. McClure JB, Derry H, Riggs KR, Westbrook EW, St John J, Shortreed SM, et al. Questions about quitting (Q2): design and methods of a Multiphase Optimization Strategy (MOST) randomized screening experiment for an online, motivational smoking cessation intervention. Contemp Clin Trials 2012 Sep;33(5):1094-1102. [doi: 10.1016/j.cct.2012.06.009] [Medline: 22771577]
29. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inquiry 2000;11:227-268 [FREE Full text]
30. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
31. Ryan RM, Deci EL. Self-regulation and the problem of human autonomy: does psychology need choice, self-determination, and will? J Pers 2006 Dec;74(6):1557-1585. [doi: 10.1111/j.1467-6494.2006.00420.x] [Medline: 17083658]
32. Collins LM, Baker TB, Mermelstein RJ, Piper ME, Jorenby DE, Smith SS, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. Ann Behav Med 2011 Apr;41(2):208-226 [FREE Full text] [doi: 10.1007/s12160-010-9253-x] [Medline: 21132416]
33. Collins LM, Dziak JJ, Li R. Design of experiments with multiple independent variables: a resource management perspective on complete and reduced factorial designs. Psychol Methods 2009 Sep;14(3):202-224 [FREE Full text] [doi: 10.1037/a0015826] [Medline: 19719358]
34. Miller WR, Rollnick S. Motivational Interviewing, Second Edition: Preparing People for Change. New York: The Guilford Press; 2002.
35. 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]
36. Böhning D, Dietz E, Schlattmann P, Mendonca L, Kirchner U. The zero-inflated Poisson model the decayed, missing filled teeth index in dental epidemiology. J R Statist Soc A 1999;162:195-209.
37. Preisser JS, Stamm JW, Long DL, Kincade ME. Review and recommendations for zero-inflated count regression modeling of dental caries indices in epidemiological studies. Caries Res 2012;46(4):413-423 [FREE Full text] [doi: 10.1159/000338992] [Medline: 22710271]
38. Raghunathan TE, Lepkowski JM, Hoewyk JV, Solenberger PA. A multivariate technique for multiply imputing missing values using a sequence of regression models. Surv Methodol 2001;27(1):85-95 [FREE Full text]
39. Van Buuren S, Brand JPL, Groothuis-Oudshoorn CGM, Rubin DB. Fully conditional specification in multivariate imputation. J Stat Comput Simul 2006;76(12):1049-1064 [FREE Full text]
40. Little RJA, Rubin DB. Statistical Analysis with Missing Data. Hoboken, NJ: Wiley; 2002.
41. Raghunathan TE, Rubin DB. Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution. J Am Stat Assoc 1991;86(416):1065-1073 [FREE Full text]
42. Rollnick S, Miller WR. What is motivational interviewing? Behav Cognit Psychother 2009 Jun 2009;23(04):325. [doi: 10.1017/S135246580001643X]
43. 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 [FREE Full text] [doi: 10.1001/archinternmed.2010.451] [Medline: 21220660]
44. Swan GE, McClure JB, Jack LM, Zbikowski SM, Javitz HS, Catz SL, et al. Behavioral counseling and varenicline treatment for smoking cessation. Am J Prev Med 2010 May;38(5):482-490 [FREE Full text] [doi: 10.1016/j.amepre.2010.01.024] [Medline: 20409497]
45. Balmford J, Borland R, Li L, Ferretter I. Usage of an Internet smoking cessation resource: the Australian QuitCoach. Drug Alcohol Rev 2009 Jan;28(1):66-72. [doi: 10.1111/j.1465-3362.2008.00009.x] [Medline: 19320678]
46. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

MOST: multiphase optimization strategy
Q2: Questions about Quitting (intervention)
ZIP: zero-inflated Poisson

Edited by G Eysenbach; submitted 21.12.12; peer-reviewed by L Vervoort, N Stanczyk, A Geraghty; comments to author 22.01.13; revised version received 28.01.13; accepted 29.01.13; published 25.03.13.

Please cite as:
McClure JB, Shortreed SM, Bogart A, Derry H, Riggs K, St John J, Nair V, An L
The Effect of Program Design on Engagement With an Internet-Based Smoking Intervention: Randomized Factorial Trial
J Med Internet Res 2013;15(3):e69
URL: http://www.jmir.org/2013/3/e69/
doi:10.2196/jmir.2508
PMID:23529377

Original Paper

Using Text Messaging to Assess Adolescents' Health Information Needs: An Ecological Momentary Assessment

Rebecca Schnall[1], RN, PhD; Anastasia Okoniewski[1], MSN, PMHNP; Victoria Tiase[2], RN, MS; Alexander Low[2]; Martha Rodriguez[1], BS; Steven Kaplan[2], MD

[1]Columbia University, School of Nursing, New York, NY, United States

[2]NewYork-Presbyterian Hospital, New York, NY, United States

Corresponding Author:
Rebecca Schnall, RN, PhD
Columbia University
School of Nursing
617 W 168th Street
New York, NY, 10032
United States
Phone: 1 212 342 6886
Fax: 1 212 305 6937
Email: rb897@columbia.edu

Abstract

Background: Use of mobile technology has made a huge impact on communication, access, and information/resource delivery to adolescents. Mobile technology is frequently used by adolescents.

Objective: The purpose of this study was to understand the health information needs of adolescents in the context of their everyday lives and to assess how they meet their information needs.

Methods: We gave 60 adolescents smartphones with unlimited text messaging and data for 30 days. Each smartphone had applications related to asthma, obesity, human immunodeficiency virus, and diet preinstalled on the phone. We sent text messages 3 times per week and asked the following questions: (1) What questions did you have about your health today? (2) Where did you look for an answer (mobile device, mobile application, online, friend, book, or parent)? (3) Was your question answered and how? (4) Anything else?

Results: Our participants ranged from 13-18 years of age, 37 (62%) participants were male and 22 (37%) were female. Of the 60 participants, 71% (42/60) participants identified themselves as Hispanic and 77% (46/60) were frequent users of mobile devices. We had a 90% (1935/2150) response rate to our text messages. Participants sent a total of 1935 text messages in response to the ecological momentary assessment questions. Adolescents sent a total of 421 text messages related to a health information needs, and 516 text messages related to the source of information to the answers of their questions, which were related to parents, friends, online, mobile apps, teachers, or coaches.

Conclusions: Text messaging technology is a useful tool for assessing adolescents' health behavior in real-time. Adolescents are willing to use text messaging to report their health information. Findings from this study contribute to the evidence base on addressing the health information needs of adolescents. In particular, attention should be paid to issues related to diet and exercise. These findings may be the harbinger for future obesity prevention programs for adolescents.

(J Med Internet Res 2013;15(3):e54) doi:10.2196/jmir.2395

KEYWORDS

text messaging; ecological momentary assessment; mobile health technology

Introduction

Understanding adolescents' health behaviors is important because behavior that is unhealthy and/or risky, rather than infectious or chronic diseases, is the leading cause of morbidity among adolescents [1]. Moreover, adolescents' health behaviors are still not well understood and have been under-studied [2]. We need a better understanding of adolescents' health behaviors in the context of their daily living in order to develop interventions that target their health related activities. This is

particularly important because behaviors and decision-making processes learned and habituated at a young age are likely to be more sustainable over time, and thus may have greater impact than attempting to change the behavior of adults [3].

One approach to addressing behavior change in adolescents is through the use of health information technology (HIT). Effective use of HIT for health communication purposes can facilitate clinical and consumer decision-making and build health skills and knowledge. For example, personal health records are an important and increasingly accepted tool for supporting patient-centered care, self-management, and effective use of health care resources [4]. There is also growing use of mobile health (mHealth) technologies, which are enabling individuals to manage their health. Use of mobile technology affords numerous advantages, including reduced memory bias, the ability to capture time-stamped data, and the potential for personalizing and tailoring information in real-time [5]. The ubiquitous nature of mobile technologies in daily life has created opportunities for applications that were not previously possible and allows for health resource monitoring and management [6]. Mobile technologies have the potential to transform health care by allowing clinicians to help patients address their health care needs in real-time [7]. For example, recent research has demonstrated the usefulness of text messaging on promoting the delivery of immunizations [8-13]. Nonetheless, these studies were limited to the use of text messaging by parents for managing their children and adolescents' health. There has been a dearth of research on adolescents' use of text messaging for managing their own health.

Mobile technology and text messaging are particularly appropriate for use in a study with adolescents since texting has become the preferred channel of communication between adolescents and their friends. In the United States, 2009, 77% of 12-17 year-olds owned cell phones, which increased from 45% in 2004 [14,15]. Cell phones have become indispensable tools in teen communication patterns [16]. In 2009, 88% of adolescent cell phone users used text-messaging, which was a large increase from the 51% in 2006 [14]. Adolescents send or receive an average of 3339 texts a month [17].

Using mobile phone technology and text messaging, we conducted a formative research study using a mixed methods approach to understand adolescents' health information needs in the context of their everyday lives and to assess how adolescents meet their health information needs.

Methods

Sample Recruitment

Prior to the start of the study, approval was obtained from the Columbia University Institutional Review Board (IRB). IRB waiver of parental consent was obtained and participants signed assent forms prior to their participation in the study. Participants were recruited from a local public high school in Bronx, New York. Participants were recruited initially on-site, followed by snowball sampling, a non-probability sampling technique where existing study subjects recruit future subjects from among their friends, to identify additional participants until the desired number is recruited [18]. Participants were allowed to keep the mobile phone as compensation for their participation in the study.

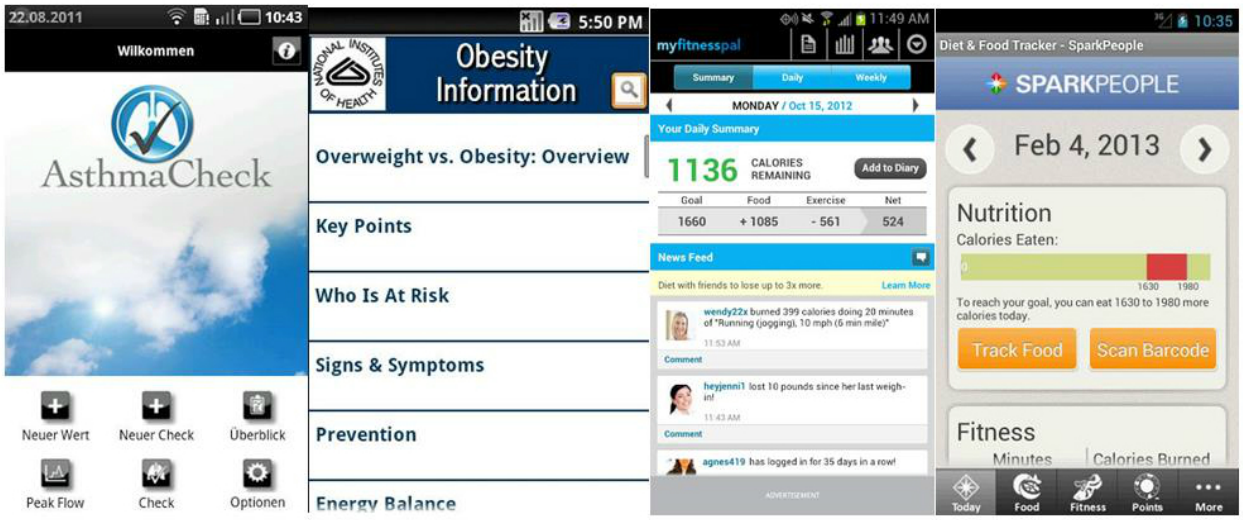
Data Collection

Prior to the start of the study, we collected demographic information and health related quality of life information from the participants using the 36-Item Short Form Health Survey (SF-36), selected from the Medical Outcome Study (MOS) inventory [19]. To evaluate health status in a valid and efficient way, Ware and colleagues developed the SF-36 to reduce questionnaire constraints faced in prior standardized health surveys [19-24]. As opposed to disease-specific measures, the SF-36 is a measure of generic health status or quality of life that includes both physical and mental health concepts. Following completion of the demographics and SF-36 forms, we used an ecological momentary assessment (EMA) to assess our participants' health information needs in the context of their daily lives and determine how they met their needs. EMA is a sampling method developed to assess phenomena at the moment they occur in natural settings and takes place in participants' naturalistic environments, supporting the ecological validity of this approach [25]. We used mobile smartphones and text messaging to conduct an EMA, which involved repeated sampling of subjects' behaviors and experiences in real time, while the participants were in their natural environments [26].

We gave 60 adolescents smartphones with unlimited text messaging and data access, as well as 600 voice minutes for 30 days. Each smartphone had applications related to asthma, HIV, obesity, diet, and exercise preinstalled on the phone. The applications which we installed were freely available on the Android and iPhone market and included: Myfitnesspal, Sparkpeople, NIH obesity, and Asthma Check (Figure 1). We used a hosted messaging gateway service to send text messages 3 times a week to ask adolescents the following questions: (1) What questions did you have about your health today? (2) Where did you look for an answer (mobile device/application, online, friend, book, or parent)? (3) Was your question answered and how? (4) Anything else?

Following 30 days of mobile phone use, we invited participants to attend a follow-up focus group session. The focus groups were held in a conference room at the Columbia University School of Nursing. Food appropriate for time of day was served and participants were compensated $20 for their time. We had a total of 4 sessions with 38 participants across all sessions.

Figure 1. Screenshots of mobile apps which were downloaded on the smartphones prior to the EMA.



Data Analysis

Demographic Data

We used descriptive statistics to calculate the frequencies, means, and standard deviations of our sample population.

EMA Data

Two authors (RS and AO) independently coded each of the text message responses. The text message responses were divided into 3 groups based on the EMA questions: (1) health information needs, (2) information sources, and (3) information need resolved. Coding began after reading each of the text messages at least twice and highlighting the relevant ideas. Codes were created based on a line-by-line analysis. Data was summarized thematically through an iterative process by each author. After coding was completed, the authors met and discussed any discrepancies in their coding until reaching a consensus on an appropriate code.

Focus Group Data

Thematic analysis was used to code focus group data. Data was summarized thematically through an iterative process by the author after each focus group session. Similar procedures were used as in the text messaging coding.

Results

Sample

We had a total of 60 participants in our EMA study. One participant had her phone stolen but we replaced the phone. We turned off 2 cell phones before the end of the 30 days because those students used more than their allotted minutes. In this study, there were 37/60 (62%) males and 22/60 (37%) females with a mean age of 15.9 (SD 1.2) years. Our gender distribution was reflective of the gender makeup of the high school from which we recruited. There were 42/60 (70.0%) participants who described themselves as Hispanic. Participants reported the following racial categories: 16/60 (26.7%) black/African American, 1/60 (1.6%) white, 7/60 (11.7%) multi-racial, and 33/60 (55.0%) other. As would be expected for a population sample primarily composed of healthy adolescents, the response distributions for each of the 8 domains of the MOS SF-36 tended to be skewed in the direction of positive health. The mean physical functioning domain was 46.92 (SD 20.4), slightly below the national mean of 50 for adults.

Adolescents in our study had a lot of previous experience with mobile devices—35/60 (58.0%) had started using a mobile device more than 2 years ago, and 46/60 (76.7%) used their mobile device more than once per day. Usage of their mobile devices during the study period is illustrated in Table 1. There was no difference in mobile phone service use by gender ($\chi^2_{4,57}$=2.68, P=0.612).

Table 1. Mobile phone service usage over 30 days.

	Mean (SD)	Minimum	Maximum
Text	2514 (2751)	50	12,474
Data (megabytes)	4,848,708 (5,938,819)	33,461	28,235,106
Minutes	743 (1045)	0	4436

We had a 90% response rate to our text messages. Participants sent 1935 text messages over the study period (Figure 2). There were 624 text messages that could not be coded because they were irrelevant or provided too little detail to be well-understood (eg, “No nothing else”, “Tomorrow?”). Adolescents sent a total of 421 text messages related to a health information need. The

health information needs are organized by themes in Table 2. Adolescents most often cited health information needs related to diet and exercise.

Adolescents sent 516 text messages related to the source of information to answer their question. Sources of information are categorized in Table 3. The most common source of information was parents (n=202). Following this information source, the Internet (n=200) was the most frequent. Of those who went online, 48/200 (24.0%) used Google as the main source for answering their health question.

In the final question of our EMA study, we sought to determine if adolescents found answers to their health questions. Participants reported that they had 421 health information needs and responded that their question was answered for 332 of their needs and not answered for 42 of their needs. Participants did not disclose whether 47 of their health information needs were met.

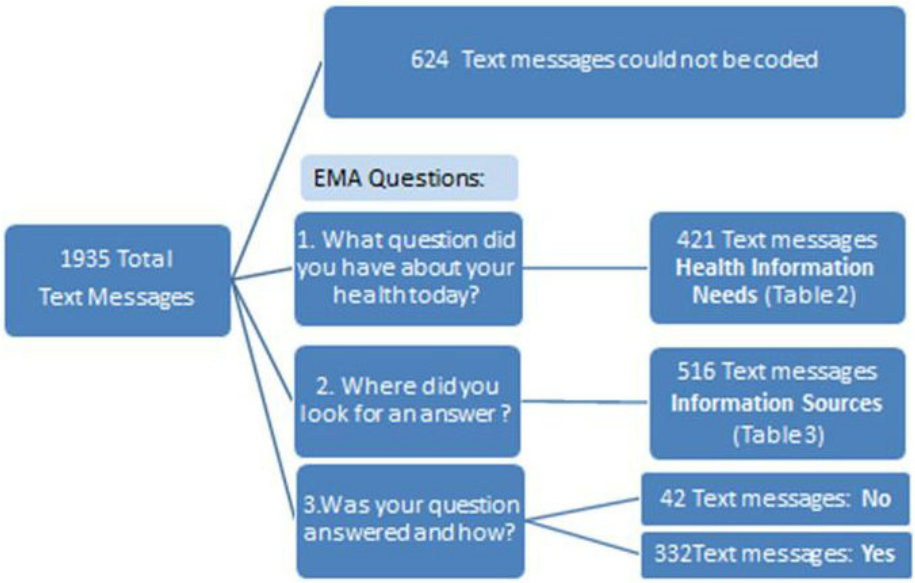
Table 2. Themes of health information needs and sample quotes.

Theme	n	Definition	Example
Diet/ Exercise	120	Nutrition, exercise, intake, calories, weight loss	*How much sugar should you have intake daily?* [EMA 19] *How come do dance everyday and still don't loose [sic] weight?* [EMA 25]
Diagnosis/ clarification	117	Exploration of conditions, definition of medical terminology	*Can allergies go away?* [EMA 13] *Why am I the only one with asthma in my family?* [EMA 8]
Basic medical care	65	Triage of routine conditions, health status, treatment, or relief of condition	*How do I get rid of a headache?* [EMA 60] *What is the easiest way to deal with pimples?* [EMA 12]
Anatomy and physiology	33	Normal body functioning	*How much blood is in the human body?* [EMA 44] *How can I lower my heart rate?* [EMA 4]
Health promotion/ prevention	27	Way to prevent a condition or disease from occurring or reoccurring	*Can you really get diabetes by drinking soda?* [EMA 64] *Does smoking affect your sports life?* [EMA 16]
Reproductive health	23	Symptom identification of SDI's, exposure, prevention, pregnancy	*My question is could you get HIV by kissing some body?* [EMA 20] *Does mirena diminish your possibilities of having kids later on in the future?* [EMA 8]
Psychological health	16	Psychological, psychiatric, or emotional issues	*My question today was what some common signs of depression are?* [EMA 48] *Can my emotional condition effect [sic] my physical condition?* [EMA 20]
Growth and development	11	Developmental norms, weight, height	*Is it healthy to weight [sic] 230 at 17?* [EMA 64] *If I am obese?* [EMA 68]
Emergency/ urgent medical care	9	First aid, bone/joint dislocations, asthma attack	*How you stop asthma attack?* [EMA 21] *What to do when you twist your ankle?* [EMA 1]

Table 3. Types of information sources and examples.

Type of resource	Total n	Example
Parent	202	*I asked my mom.* [EMA 62] *It was answer by my step father showing me different way to do it.* [EMA 19]
Online	200	*It was answers by online using a search site like Google.* [EMA 68] *Yes and by a medical Q and A website.* [EMA 63]
Friend	39	*I asked around to my friends.* [EMA 21]
Mobile device/ application	20	*I found answers on the information app on obesity on the phone.* [EMA 12]
Teacher/ coach	19	*I got the idea from my gym teacher.* [EMA 9]
Doctor/ nurse	14	*Doctor told me this.* [EMA 37]
Family	12	*I asked my sister for a answer.* [EMA 49]
Other	10	*Book at the library.* [EMA 19]

Figure 2. Flowchart of text message responses.



Follow-up Focus Group Sessions

Overview

Of the 60 participants in our EMA study, 38 participated in the follow-up focus group sessions. During the focus group sessions, adolescents reported that they used their mobile phones for music, calendar, alarm, text messaging, pacer, games, navigation, camera, Tumbler, Google, Facebook, and YouTube. One adolescent said, "I just went on Facebook a lot and Google." Another participant said, "If I have any health-related problems I usually just go on Google and it takes to me some doctor website, where real doctors answer the questions." The websites that the adolescents reported were the same as those reported during their EMA, which triangulates our findings. Adolescents reported that they encountered multiple barriers when trying to use the mobile health applications that were installed on the phone. We identified the following barriers to mobile health application use: ease of use, readability, end user needs, and privacy.

Ease of Use

One participant stated, "it looked too complicated". Another said, "I just didn't understand it." Participants also explained that they "didn't want to read the whole thing, terms and agreements". If the terms and agreements of an application were too long, then some participants "would just close the app rather than read it".

Readability

The language contained within the application was also too difficult as one participant described, "They kept using big words." To facilitate easy use of the applications, one participant recommended that we "dumb everything down".

End User Needs

Participants reported that they did not like the apps because they were not tailored for adolescents' use.

> *Honestly, I didn't like the apps, because it's like for older people like, I'm still a teenager, so I don't want to be…boring.*

One participant reminded the moderator that "teenagers think differently from other people", and so the technology should be relevant to them, for example, "it can't be like oh, has your Alzheimer's been better?" Adolescents also emphasized the importance of the user interface and how they choose to use "the ones that looks cooler". One participant explained that the health applications on the phone were "just the way [they looked], like it was boring".

Privacy

Privacy is a consideration for adolescents when deciding to use mobile technology. Nearly all of our participants in our focus groups reported that they put a code on their phone to maintain privacy. Participants cited reasons such as "I want my stuff to be protected, like I don't know who is going to go into my phone and do what with it." One female expressed, "I don't want my mom to see my messages." Another participant said, "I don't want anybody touching my stuff", and the code would discourage others from using his phone.

In addition to putting codes on their phones, participants were concerned that the researchers could see what they were doing with their phones. One participant said, "I felt like there was a person sitting on a computer watching me." Another participant said, "People don't like being watched." Finally, one adolescent said, "I thought that they were looking at our messages and stuff."

Discussion

Findings from our study demonstrated the usefulness of text messaging with health information as part of an EMA study with adolescents. EMA aims to minimize recall bias, maximize ecological validity, and allow study of behavior in real-world contexts. Since the mobile phone has become the favored communication tool for the majority of American teens [27], this tool offers promise for future investigation and data collection. While past studies have demonstrated the use of text messaging for delivering reminders to promote healthy behavior choices, these studies have been limited to parents' use of this tool for managing their children's health. In contrast this study provided preliminary evidence for the acceptability of the use of text messaging for adolescents managing their own health care. Moreover, our study demonstrated that adolescents are willing to use this technology and report on their own health questions in context of their daily lives. EMA with text messaging holds unique promise for future studies with adolescents who may be more resistant and unreliable for follow-up visits but may be willing to use text messaging in its place. In particular, our sample was unique in that they included many underserved and at-risk adolescents. This is noteworthy since developing these interventions for hard-to reach populations are critical and our study findings provide promise for the use of mHealth technology with these adolescents.

Since adolescents text frequently and have multiple health information needs, findings from our study suggested that development of health interventions using mobile technology has great promise. Nonetheless, while adolescents are interested in using applications and features on their phone, the language and interface needs to be tailored for adolescents' use. Our study findings indicated that mHealth technology targeted for adolescent interventions must be designed specifically for adolescents, otherwise it is unlikely that they will be willing to use it. Particular attention should be paid to the user interface as well as developing tools, which must meet the literacy needs of underserved and high-risk adolescents [28-30]. Past research has demonstrated the importance of both of these issues in technology adoption, nonetheless, it is worthwhile to consider the needs of adolescents and tailor mobile apps to meet their literacy needs [31] and designing a user interface which is appealing to them [29,30,32].

Also notable were adolescents' privacy concerns, which might be of interest to policymakers. With the passage of the Health Information Technology for Economic and Clinical Health (HITECH) act, personal health data will soon flow freely in electronic form with the provision that it is available securely and privately [33]. Nonetheless, it is important to consider adolescents' views on the electronic transfer and viewing of their personal health information and their concerns about it. Concerns over the security of their personal health information may be a barrier to seeking health care services.

While the results of our study indicated that adolescents' health information needs are oftentimes met, it is important to educate adolescents who are heavy users of mobile technology and the Internet on how to verify their sources of information. As was evidenced by one of our participants who thought that "real doctors" answered her questions online, adolescents need to be educated so that they can be informed consumers of health information. Further research on how adolescents verify their sources of health information is warranted.

Of note were our participants' most frequently reported questions about diet and exercise. While concerns over childhood obesity in our country continue to grow, our study indicated that our sample of ethnically diverse adolescents has heightened awareness and health information needs regarding issues related to obesity. Despite health information needs related to diet and exercise and past studies on this topic [34], a recent systematic review indicated that the outcomes of all of the existing obesity prevention programs across all age groups were modest and there were few replication studies of any program suggesting that the need for targeted interventions exist [35]. Three of the females in our follow-up focus groups reported that they used the preinstalled diet-tracking application, which was corroborated by data from the EMA text messages. They all reported that they used the application for only a few days because they forgot to continue using it or stopped their diet and did not want to feel guilty. Two participants recommended that we send push notifications to remind them to encourage continual use of the application in the future. These findings may be useful for informing the design of obesity prevention programs for adolescents, but further study is necessary.

Limitations to our study included a convenience sample that may not be representative of all ethnically diverse adolescents, particularly as they were recruited from one area of New York City. Even so, our sample had similar mobile technology use as other adolescents in the United States; past research has shown that 1 in 3 adolescents sends more than 100 text messages a day [16], while our participants sent an average of 84 (SD 91.72) texts per day. Another limitation of our study was the use of EMA. As with all self-report measures, there was no independent check on the veracity of the data, because all data were collected in the absence of the experimenter. In addition, since the browsing history of our participants was not assessed, responses to the text messages could be inaccurate owing to a number of factors including stigma. Finally, our technology was limited and did not have a decision tree associated with it to provide responses and feedback to our participants. This may have limited participants' willingness to continue sharing their health information needs.

Findings from our study indicated the usefulness of text messaging technology as a tool for assessing adolescents' health behavior in the context of their daily lives. Our study demonstrated that adolescents are willing to use text messaging to report their health information. Moreover, adolescents in our study provided useful information for the development of future health behavior tools using mobile technology.

Acknowledgments

This study was funded by HEAL NY Phase 6 - Primary Care Infrastructure "A Medical Home Where Kids Live: Their School" Contract Number C024094 (Subcontract PI: R Schnall) and the Center for Evidence Based Practice in the Underserved NINR P30NR010677 (S. Bakken, Principal Investigator). Findings from this study have been previously presented for poster presentation at AMIA 2012, Chicago, IL and oral presentation at the 2012 mHealth Summit, National Harbor, MD.

Conflicts of Interest

None declared.

References

1. Park MJ, Paul Mulye T, Adams SH, Brindis CD, Irwin CE. The health status of young adults in the United States. J Adolesc Health 2006 Sep;39(3):305-317. [doi: 10.1016/j.jadohealth.2006.04.017] [Medline: 16919791]
2. Lawrence RM, Gootman JA, Sim LJ. Models of care for treatment. In: Adolescent health services: missing opportunities. Washington, DC: National Academies Press; 2009.
3. Mann L, Harmoni R, Power C. Adolescent decision-making: the development of competence. J Adolesc 1989 Sep;12(3):265-278. [Medline: 2687339]
4. Commission for Systemic Interoperability. 2005. Ending the document game: connecting and transforming your healthcare through information technology URL: http://endingthedocumentgame.gov/PDFs/entireReport.pdf [accessed 2013-03-01] [WebCite Cache ID 6Eneqfloh]
5. Istepanian R, Laxminarayan S, Pattichis CS. M-health: emerging mobile health systems. New York, NY: Springer; 2006.
6. mHealth Alliance. 2011. Frequently Asked Questions URL: http://www.mhealthalliance.org/about/faq [accessed 2013-03-01] [WebCite Cache ID 6EnOepnvn]
7. Stone A. The science of real-time data capture: self-reports in health research. Oxford: Oxford University Press; 2007.
8. Kharbanda EO, Stockwell MS, Fox H, Andres R, Lara M, Rickert VI. Text message reminders to promote human papillomavirus vaccination. Vaccine 2011 Mar 21;29(14):2537-2541. [doi: 10.1016/j.vaccine.2011.01.065] [Medline: 21300094]
9. Stockwell MS, Kharbanda EO, Martinez RA, Lara M, Vawdrey D, Natarajan K, et al. Text4Health: impact of text message reminder-recalls for pediatric and adolescent immunizations. Am J Public Health 2012 Feb;102(2):e15-e21. [doi: 10.2105/AJPH.2011.300331] [Medline: 22390457]
10. Stockwell MS, Kharbanda EO, Martinez RA, Vargas CY, Vawdrey DK, Camargo S. Effect of a text messaging intervention on influenza vaccination in an urban, low-income pediatric and adolescent population: a randomized controlled trial. JAMA 2012 Apr 25;307(16):1702-1708. [doi: 10.1001/jama.2012.502] [Medline: 22535855]
11. Kharbanda EO, Vargas CY, Castaño PM, Lara M, Andres R, Stockwell MS. Exploring pregnant women's views on influenza vaccination and educational text messages. Prev Med 2011 Jan;52(1):75-77. [doi: 10.1016/j.ypmed.2010.10.009] [Medline: 21047526]
12. Ahlers-Schmidt CR, Chesser AK, Paschal AM, Hart TA, Williams KS, Yaghmai B, et al. Parent opinions about use of text messaging for immunization reminders. J Med Internet Res 2012;14(3):e83 [FREE Full text] [doi: 10.2196/jmir.1976] [Medline: 22683920]
13. Ahlers-Schmidt CR, Chesser AK, Nguyen T, Brannon J, Hart TA, Williams KS, et al. Feasibility of a randomized controlled trial to evaluate Text Reminders for Immunization Compliance in Kids (TRICKs). Vaccine 2012 Aug 3;30(36):5305-5309. [doi: 10.1016/j.vaccine.2012.06.058] [Medline: 22750044]
14. Lenhart A. Pew Internet & American Life Project. Washington, DC: Pew Internet and American Life Project; 2010. Teens, Cell Phones and Texting URL: http://www.pewinternet.org/Reports/2009/14--Teens-and-Mobile-Phones-Data-Memo/1-Data-Memo/5-How-teens-use-text-messaging.aspx?r=1 [accessed 2013-02-28] [WebCite Cache ID 6Em2acuDJ]
15. Lenhart A. Pew Internet & American Life Project. Washington, DC: Pew Internet & American Life Project Teens, Smartphones & Texting; 2012 URL: http://pewinternet.org/Reports/2012/Teens-and-smartphones.aspx [accessed 2013-02-28] [WebCite Cache ID 6Em2bLXmP]
16. Lenhart A. Pew Internet & American Life Project. Teens and Mobile Phones Over the Past Five Years: Pew Internet Looks Back; 2009. URL: http://www.pewinternet.org/Reports/2009/14--Teens-and-Mobile-Phones-Data-Memo/1-Data-Memo/5-How-teens-use-text-messaging.aspx?r=1 [accessed 2013-02-28] [WebCite Cache ID 6Em2c6its]
17. Nielsenwire. Pew Internet & American Life Project. U.S. Teen Mobile Report: Calling Yesterday, Texting Today, Using Apps Tomorrow. The Nielsen Company; 2010. URL: http://blog.nielsen.com/nielsenwire/online_mobile/u-s-teen-mobile-report-calling-yesterday-texting-today-using-apps-tomorrow/ [accessed 2013-02-28] [WebCite Cache ID 6Em2ceRyy]
18. Goodman LA. Snowball sampling. Ann Math Statist 1961 Mar;32(1):148-170 [FREE Full text] [doi: 10.1214/aoms/1177705148]
19. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993 Mar;31(3):247-263. [Medline: 8450681]

20. Windisch W, Freidel K, Schucher B, Baumann H, Wiebel M, Matthys H, et al. Evaluation of health-related quality of life using the MOS 36-Item Short-Form Health Status Survey in patients receiving noninvasive positive pressure ventilation. Intensive Care Med 2003 Apr;29(4):615-621. [doi: 10.1007/s00134-003-1675-5] [Medline: 12618917]
21. Ware J, 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]
22. Russo J, Trujillo CA, Wingerson D, Decker K, Ries R, Wetzler H, et al. The MOS 36-Item Short Form Health Survey: reliability, validity, and preliminary findings in schizophrenic outpatients. Med Care 1998 May;36(5):752-756. [Medline: 9596066]
23. Perneger TV, Leplège A, Etter JF, Rougemont A. Validation of a French-language version of the MOS 36-Item Short Form Health Survey (SF-36) in young healthy adults. J Clin Epidemiol 1995 Aug;48(8):1051-1060. [Medline: 7775992]
24. McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994 Jan;32(1):40-66. [Medline: 8277801]
25. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavorial medicine. Annals of Behavioral Medicine 1994;16(3):199-202.
26. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [Medline: 18509902]
27. Lenhart A. Pew Internet & American Life Project. Teens and Mobile Phones Over the Past Five Years: Pew Internet looks back; 2009. URL: http://www.pewinternet.org/htm [accessed 2013-02-28] [WebCite Cache ID 6Em2WbDFU]
28. Paek HJ, Hove T. Social cognitive factors and perceived social influences that improve adolescent eHealth literacy. Health Commun 2012 Nov;27(8):727-737. [doi: 10.1080/10410236.2011.616627] [Medline: 22452551]
29. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]
30. Stinson JN, Petroz GC, Tait G, Feldman BM, Streiner D, McGrath PJ, et al. e-Ouch: usability testing of an electronic chronic pain diary for adolescents with arthritis. Clin J Pain 2006;22(3):295-305. [doi: 10.1097/01.ajp.0000173371.54579.31] [Medline: 16514331]
31. Barron AE, Kemker K, Harmes C, Kalaydjian K. Large-scale research study on technology in K-12 school: technology integration as it relates to the national technology standards. Journal of Research on Technology in Education 2003;35(4):489.
32. Ozen C, Basoglu N. Exploring the contribution of information systems user interface design characteristics to adoption process. 2007 Presented at: Management of Engineering and Technology; August 5-9, 2007; Portland p. 2208-2219. [doi: 10.1109/PICMET.2007.4349415]
33. Blumenthal D. Launching HITECH. N Engl J Med 2010 Feb 4;362(5):382-385. [doi: 10.1056/NEJMp0912825] [Medline: 20042745]
34. 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]
35. Haynos AF, O'Donohue WT. Universal childhood and adolescent obesity prevention programs: review and critical analysis. Clin Psychol Rev 2012 Jul;32(5):383-399. [doi: 10.1016/j.cpr.2011.09.006] [Medline: 22681912]

Abbreviations

EMA: ecological momentary assessment
HIT: health information technology
HITECH: Health Information Technology for Economic and Clinical Health
HIV: human immunodeficiency virus
IRB: Columbia University Institutional Review Board
mHealth: mobile health
MOS: Medical Outcome Study
SF-36: 36-Item Short Form Health Survey

Edited by G Eysenbach; submitted 12.10.12; peer-reviewed by A Barman-Adhikari, A Chesser, P Lam; comments to author 21.01.13; revised version received 07.02.13; accepted 21.02.13; published 06.03.13.

Please cite as:
Schnall R, Okoniewski A, Tiase V, Low A, Rodriguez M, Kaplan S
Using Text Messaging to Assess Adolescents' Health Information Needs: An Ecological Momentary Assessment
J Med Internet Res 2013;15(3):e54
URL: http://www.jmir.org/2013/3/e54/
doi:10.2196/jmir.2395
PMID:23467200

Original Paper

Physicians Interrupted by Mobile Devices in Hospitals: Understanding the Interaction Between Devices, Roles, and Duties

Terje Solvoll[1,2]; Jeremiah Scholl[3], PhD; Gunnar Hartvigsen[1,2], PhD

[1]University Hospital of North Norway, Norwegian Centre for Integrated Care and Telemedicine, Tromsø, Norway

[2]University of Tromsø, Department of Computer Science, Tromsø, Norway

[3]Karolinska Institutet, Health Informatics Centre, Jämtlands Län, Sweden

Corresponding Author:
Terje Solvoll
University Hospital of North Norway
Norwegian Centre for Integrated Care and Telemedicine
Po. Box 35
Tromsø, 9038
Norway
Phone: 47 90932183
Fax: 47 77754099
Email: terje.solvoll@telemed.no

Abstract

Background: A common denominator of modern hospitals is a variety of communication problems. In particular, interruptions from mobile communication devices are a cause of great concern for many physicians.

Objective: To characterize how interruptions from mobile devices disturb physicians in their daily work. The gathered knowledge will be subsequently used as input for the design and development of a context-sensitive communication system for mobile communications suitable for hospitals.

Methods: This study adheres to an ethnographic and interpretive field research approach. The data gathering consisted of participant observations, non-structured and mostly ad hoc interviews, and open-ended discussions with a selected group of physicians. Eleven physicians were observed for a total of 135 hours during May and June 2009.

Results: The study demonstrates to what degree physicians are interrupted by mobile devices in their daily work and in which situations they are interrupted, such as surgery, examinations, and during patients/relatives high-importance level conversations. The participants in the study expected, and also indicated, that wireless phones probably led to more interruptions immediately after their introduction in a clinic, when compared to a pager, but this changed after a short while. The unpleasant feeling experienced by the caller when interrupting someone by calling them differs compared to sending a page message, which leaves it up to the receiver when to return the call.

Conclusions: Mobile devices, which frequently interrupt physicians in hospitals, are a problem for both physicians and patients. The results from this study contribute to knowledge being used as input for designing and developing a prototype for a context-sensitive communication system for mobile communication suitable for hospitals. We combined these findings with results from earlier studies and also involved actual users to develop the prototype, CallMeSmart. This system intends to reduce such interruptions and at the same time minimize the number of communication devices needed per user.

(J Med Internet Res 2013;15(3):e56) doi:10.2196/jmir.2473

KEYWORDS

Ethnography; Mobile communication; interruptions; Context-aware computing; Pervasive computing; User-centered design methods; Health care; Pagers; Wireless phones; CSCW; HCI

Introduction

The work setting in hospitals is communication intensive and can lead to significant difficulties related to interruptions from other co-workers [1-4]. One suggested solution for this problem is to implement wireless phone systems [5-7]. Such systems offer a number of advantages over traditional paging systems, such as not requiring the staff to find a phone after being paged.

Moving from wired phones to wireless communication platforms has a significant impact in the future of health care delivery systems [8]. Mobile phones' interference with electronic equipment is, according to [9], not an issue anymore, or it can be solved by the solutions in [10].

The adoption of wireless phones, however, does not come without risk. Psychological theory [11] and empirical evidence [12] both suggest that wireless phones have the potential to create additional problems related to interruptions, such as conversations and questions that would not normally occur, when compared to using traditional paging systems. This has caused some hospital staff to resist the adoption of wireless phone systems when given the opportunity [13]. Further studies regarding the nature and effects of interruptions from wireless phone systems in hospitals have thus been suggested [14].

There have been many suggestions on how to reduce interruptions from mobile devices over the years. Many of these systems are based on using contextual information to reduce interruptions. Some of these systems change the configuration of the phone automatically [15-18] and includes quiet calls where the receiver could negotiate with the caller through text or pre-recorded audio messages. However, this negotiation does not reduce personal interruptions in that the user must still act when receiving a call. Other systems focus on giving the caller information about the receivers' context and thereby helping the caller make decisions on when to call [19-21]. Avrahami et al discovered that if they provided the caller with contextual information about the receiver's situation, it reduced the mismatch between the caller's decision and the receiver's desires [22]. To our knowledge, none of these systems has been tested in hospital settings.

Hospitals are dependent on a wide and reliable communication infrastructure for exchanging different kinds of data, including text, voice, and alarm services. Several studies aiming to improve communication and reduce interruptions have been carried out, also within hospital settings, without major success [3,5-7,23,24]. Personal Digital Assistants (PDAs) have also been tested in a contextual message exchange system without major success [25]. Other systems, like the AwareMedia and AwarePhone systems of Bardram et al [26,27], support context-sensitive communication and form a complete communication system for clinicians in a surgical ward. The AwarePhone system is an application running on a mobile phone (GSM/3G) that is not integrated in the hospital's internal communication infrastructure, therefore, it requires the user to carry another device. The feedback from its users focused on privacy issues as one of the major drawbacks.

In this paper, we present an interpretive case study regarding interruptions from mobile devices at a large hospital in Norway, which started implementing a pervasive wireless phone system in 2006. This setting offers the opportunity to gain perspective from health staff with several years of experience using wireless phone systems. The goal of the study was to characterize a physician's workday, focusing on wireless communication. The aim was to understand the health care workers' communication pattern and how unnecessary interruptions from mobile devices can be reduced: in which situations, what context, and which location. Such design-oriented studies using methods from computer-supported cooperative work have been suggested for medical informatics systems in general [28], and medical collaboration systems in particular [29], in order to improve their overall effectiveness and success rate upon implementation. This knowledge contributes to medical informatics by improving understanding of the effects of the use of wireless phone systems and is used as input to the overall sociotechnical development of a wireless communication system for hospitals.

Methods

Research Setting

The study was conducted at St. Olavs Hospital, Trondheim University Hospital, Norway, which is a health enterprise in Mid-Norway health region that consists of 695,000 inhabitants in total. St. Olavs Hospital is under renovation, and the hospital project consists of the construction of new buildings and a new hospital organization. The first new clinical centers were completed in 2006, and the entire project will be completed in 2015. The new St. Olavs hospital is a technologically advanced hospital, and the communication system is one of the world's most modern hospital communication systems. The existing system is based on wired and wireless (Internet Protocol) IP-phones from Cisco. The wireless IP-phones are the technology of interest for this research. The model in use at the studied hospital is the Cisco 7921G wireless phones. They also use an old paging system, both for on-call duty pagers and personal pagers, and a GSM-based on-call duty phone in parallel as back-up, since the older parts of the hospital do not have coverage for wireless IP-phones. The pagers were from different vendors but were all of the simple type that can receive only the number from the caller. The GSM-phone was an ordinary Nokia mobile phone. Most of the physicians also carried a private mobile phone.

Two separate clinics were included in the study: the Ear, Nose, and Throat (ENT) Clinic and the Child and Youth (CY) Clinic. Two clinics from the same hospital were included in order to provide a broader understanding on the usage and experiences of wireless phones, compared to studying only one clinic. This also provided us the opportunity to compare and contrast the data collected between the separate clinics.

Research Method

The study adheres to an ethnographic and interpretive field research approach [30-33]. Ethnography and participant observations represent a uniquely humanistic interpretive approach [33]. Interpretive research has the potential to explain the human thought and action in a social and organizational context [31]. Principles for trusted pervasive health have also been under consideration [34,35].

Data gathering conducted by the first author consists of participatory observations, nonstructured and mostly ad-hoc interviews, and open-ended discussions of a selected group of physicians at various levels of hierarchy and roles, within two clinics at St. Olavs Hospital. The participatory observer (hereafter referred to simply as "observer") was from a different institution at a different location than the hospital where the

study was performed. The observer's background is computer science, communication, and medical informatics, and he was not medically trained to perform this study. This was fully understood by the participants, who were also aware that the observer was a researcher. The observer signed a nondisclosure agreement regarding sensitive information prior to the observations.

Participant Observations

The fact that peoples' descriptions of their activities often differ from what really occurs in practice is one of the main rationales for doing observations. There can be several reasons for this phenomenon, such as the limitation of the human memory, people not always being aware of their actual behavior, peoples' concern with their image (and therefore giving a better story than the real one), and also the complexity of social life causing different people to report differently.

The observer followed the independent work of 11 physicians, including 2 cancer teams and 2 surgical teams at two clinics, for a total of 135 hours during May and June 2009. The purpose of the work was to observe the daily work of physicians and to interview them, focusing on interruptions from mobile devices. The observer stayed at the ENT clinic for a total of 65 hours: 26 hours observing assistant physicians and 39 hours observing chief physicians. At the CY clinic, he stayed for a total of 70 hours. 15 hours were used to observe assistant physicians, 47 hours observing chief physicians, and 8 hours to observe one chief and one assistant physician working on a cancer team. The observer took the role of a first-year medical student, dressing and acting like a physician to blend in as much as possible for a more realistic picture of the communication situation at the clinic. This technique is often referred to as "shadowing" [36]. He followed each physician in their everyday work at outpatient clinic, surgery, cancer meetings, etc, for at least 2 workdays/nights/duties. The head of the clinic chose physicians and roles to form an average representation of the physicians and roles at the clinic. The observer had to contact each physician to make an appointment for each observation, which was done during the morning meeting at each clinic. There were some changes of the selected physicians and roles due to shifts, which resulted in increased observing hours and did not influence the representativeness of the clinic. The observer registered every call/page/message, type of device, reaction, and context for each physician. Depending on the situation, he also asked questions related to the context and the communication device used.

Data were recorded using pen and paper on a self-constructed four splitter form, registering: situation, device, time (start - stop), and reaction/response (meaning, answer or not, by whom, what context, etc). The reaction/response was recorded as free text notes. The questions made by the observer were related to the context and the communication device (if any). They were asked directly after the event, and the answers written down in free text notes on the form. We do not believe this had any effect on the subjects' behavior and the generalization of the results, due to the nature of the questions: "Was this a call related to your professional role on duty or role-independent personal call?", "Was this considered as an important call that you had to answer, or was it a more general question that another physician could have answered?", etc. In most of the cases, the subject informed the observer unsolicited. The decision about using pen and paper instead of a digital device such as a PDA, as used in [37-40], was made since these physicians did not use these kinds of devices in their daily work. This helped the observer blend in as much as possible in the health care workers settings and avoided unintended attention paid by the workers to the observations.

Interviews/Discussions

Interviewing is also an important method within ethnography. Interviews increase understanding of what has been observed and the subjects' perspective of a specific situation. One way to interview is during an on-going activity like we did, during the observations. In this way, we connected the observations with the interview and therefore got answers to questions about the observed activity, which helped us understand the situation from the subjects' perspective.

During the observation period at each clinic, the observer had office accommodation in an open-plan office among the assistant physicians. This created the opportunity for several discussions and also the collection of input on how to improve the communication system—mostly on improving the user experience, but also on how to reduce the interruptions. While observing the physicians, the observer was allowed to interview and talk to the physicians, but he remained quiet during patient consultations. All questions and discussions, both during observation and office time, were conducted in Norwegian and notes were taken from them, but not recorded. Interview guides were not used, but there were questions related to context and what was observed. Some of the questions were asked of all participants, while others were specific to the situation. The initial focus of the interviews and discussions was the use of wireless mobile communication devices regarding improvement of interruption management.

Data Analysis

Data analysis was conducted concurrently with the data collection process using Grounded Theory. Initial data analysis began with the first author reviewing all the notes and reflecting on some general issues that seemed inherent to the data. It was clear at this point that, despite there being some similarities, there also were differences with respect to the two clinics. A decision was then made to analyze the data from the two clinics separately. For each section the data were analyzed around five basic themes: (1) frequency of interruptions, (2) the clinical situation, (3) context [41], which included location, (4) comparisons of pagers vs wireless/mobile phones, and (5) answered or not vs returned call after a page, and the importance of the call. Exploring these data was an adequate approach to gaining an understanding of the communication in hospitals. The data indicate the communications' frequency, length, and importance, if they were interruptive or not, and the reaction from the user. Analyzing log data is a satisfactory approach to mapping out the actual communication patterns. It is also a useful approach to visualize the numbers and the length of communication sessions, and how often the users were interrupted unnecessarily. Analyzing observed situations and

then comparing them with comments and answers during the interviews and discussions worked well for mapping the communication situation and capturing and understanding potential changes over time.

Results

Although there were many similarities between the two clinics, there were also a number of differences regarding communication patterns and the handling of different situations. The results from the clinics are therefore presented separately.

A general similarity between the clinics was that each physician at both clinics carried at least one wireless IP-phone, one on-call duty pager, and one private mobile phone. Some of them, mostly assistant physicians, also carried a personal pager and, during evening and night duty, a backup GSM phone was also required. To use the wireless IP-phone, the physician had to personally log on to the phone, and if they also were responsible for a role, they had to log on to that role too. When referring to a physician, we will call them Phys-A to K, and nurses as Nurse-A to F.

Ear, Nose, and Throat Clinic

The Ear, Nose, and Throat (ENT) clinic has a regional function for the Mid-Norway health region. The clinic treats problems that the general practitioners or private specialists in ear, nose, and throat, have not been able to treat. The clinic also incorporates jaw- and eye- units, but we concentrated our observations to the ENT-unit, which consisted of the following sections: outpatient clinic, surgery, and inpatient ward. The clinic takes care of work-up, diagnostic, medical, and surgical treatment within cancer, nose, sinus treatment, etc. We observed assistant physicians' role as on-call duty at the inpatient ward and outpatient clinic, in surgery, and also observed the chief physician's role as on-call duty at the inpatient ward, outpatient clinic, in surgery, and in cancer meetings. A cancer meeting was a team of physicians meeting/examining and treating cancer patients.

Frequency of Interruptions

One of the assistant physicians, Phys-A, said on the first day of observations: "It varies a lot how much we are interrupted from various mobile devices. The busiest time is definitely between 8 in the morning and 7-8 in the evening". However, only a few hours that we observed at this clinic were considered as "normal" regarding interruptions. Table 1 shows all interruptions from mobile devices received by the participating physicians at this clinic.

During the observations the physicians and/or their nearest colleagues commented on the frequency of interruptions during the day. A chief physician, Phys-B, stated one day during outpatient clinic: "Unusually quiet on the phone today or actually it has been this quiet all week. It seems like many have started their vacation", and a theatre nurse, Nurse-A, said during a surgery, "This has to be the first surgery I've performed together with NN where he has not been constantly interrupted by the phone". During a cancer meeting, the chief physician, Phys-B, said, "This is a quiet day regarding calls, but it is publicly known at the clinic which days I have these meetings", which was continued by the head nurse, Nurse-B: "The phone should not ring when we are in these kind of meetings. The rule is that a secretary answers the phone and takes messages". Even though most of the observation hours at this clinic were considered as "quiet", Table 1 show that they were interrupted several times in situations when they should not have been unnecessarily interrupted. Most of the interruptions were mentioned afterwards as "not important" and could have been postponed until the physician becomes available, or answered by others.

Table 1. Overview of total amount of interruptions from mobile devices at ENT during observation time (only interruptions from the followed physicians' devices, except for *g*, which was a nurses' phone brought to the physician).[a]

			Assistant physician	Chief physician
Preparatory/complementary work				
		Answered	3a	8a
		Ignored		
Outpatient ward				
	No patient			
		Answered	1b	3a
		Ignored		
	With patient			
		Answered	2a,8b,1c, 3e,1g	
		Ignored		
Inpatient ward				
	No patient			
		Answered		1a
		Ignored		
	With patient			
		Answered		
		Ignored		1a
Surgical theatre				
		Answered	1a	5a
		Ignored		1a
Meeting				
		Answered	3a	2a
		Ignored		
Cancer meetings				
		Answered		3a, 1b
		Ignored		1a, 2b
Other situations				
		Answered	1a, 1b	11a
		Ignored		

[a]a=wireless IP-phone; b=on-call-duty pager; c=backup on-call-duty GSM-phone; e=Wired IP-phone; g=other

While observing in the outpatient clinic, we discovered that the health care workers at the clinic often sought the physician in outpatient duty "in-person" instead of calling him/her. In this way, they grasped the context immediately and knew if they could interrupt or not. We also observed that the less experienced physicians often used the nurses to contact the experienced physicians, instead of calling them themselves. In this way, they could continue their own work while waiting for an answer. We observed such situations at the outpatient clinic, where normally a nurse came into the room and brought her own wireless IP-phone to the physician, and asked if he/she could answer a question from the caller.

Another situation that was observed was that calls/pages during meetings/lectures at the clinic normally were answered. This was also pointed out by the head of the clinic:

> *There are only two physicians that have on-call duty during each meeting/lecture, but regardless, everyone immediately answers incoming calls/pages… This is not necessary, and the most annoying part is when the person picks up the phone and starts to talk on their way out.*

Interruptions During Surgery

Table 1 show that physicians at this clinic also responded to phone calls during surgery. In this case, rather than answering the phone themselves, somebody else in the room answered the

call for them, normally a nurse or another physician. The person who answered the phone either took a message or held the phone up to the physician's ear. After such an incident during a surgery where the theatre nurse, Nurse-B, answered the chief physician's phone and held the phone up the physician's ear, she said:

> *Normally we only take messages or convey the question to the physician, but, like now, if it is ok for the physician, we could hold the phone for the physician, which happens rarely. Most of the calls are questions that could be conveyed by the nurse and then answered through the nurse. Normally after a surgery, there are more or less 10 messages waiting for the physician that could be difficult to manage. Most of these calls have to be returned anyhow.*

We also observed that the pager was left in the physician's coat outside the surgical theatre, and they brought with them only their private mobile phone and the wireless IP-phone.

Wireless Phones vs Pagers

The physicians at the clinic appeared to be pleased with carrying a wireless IP-phone, and during discussions about carrying wireless phones in the new hospital compared with pagers, a chief physician, Phys-B, said:

> *People are generally satisfied with carrying phones compared to the old days when we only had pagers, but it took a while to adjust to the new system. After some time you learn to screen the calls, the important ones from those that are not so important. With a pager you never know, and you therefore have to return the call as soon as possible.*

In response to the question about the phone being more interruptive compared to the pager-only situation, he said: "The phone is less interruptive compared to the pager. I do not know if it is me who has been better at telling people when I'm available and managing the communication, or if it is the others who have been better." This also seemed to be the general view at this clinic. Another important point regarding why there may be fewer interruptions from the phone compared to when they had only pagers, was made by a chief physician, Phys-E: "If you call someone and they are busy, and your interruption is not important enough, you risk an unpleasant situation, while paging someone is not that risky…".

Child and Youth Clinic

The Child and Youth (CY) clinic treats patients aged between 0-16 years, mainly from the mid-Norway health region, but also from other parts of Norway. The clinic consists of 10 units, where we concentrated on the outpatient child and youth unit, cancer and blood diseases unit, and child and newborn intensive care units. We observed one general chief physician, one newborn intensive care chief physician, one chief physician in cancer and blood diseases, two assistant physicians, in two roles: primary watch duty and intermediate watch duty, which are the primary and secondary on-call-duties at this clinic.

Frequency of Interruptions

The frequency of interruptions from mobile devices during observation time at this clinic were considered as "normal" most of the observation days, except for one day observing a chief physician, Phys-F, at the outpatient clinic who (for a period) worked there only once a week: "It will probably be quiet on the phone today, I'm seldom here, so it depends if anybody knows I'm here". The other days were more normal, like another chief physician, Phys-I, illustrated: "This is an average day regarding the number of calls. You know we pediatricians are so 'kind' and therefore easy to call/contact…". Table 2 shows all interruptions from mobile devices received by the participating physicians at this clinic. However, some of the days had periods that were rather busy, which was illustrated by an assistant physician, Phys-H, when he announced in frustration after a phone call and an incoming page: "If somebody tries to call me on the phone, and it is busy, they page me, why? I'm busy on the phone and cannot return the call immediately." Another assistant physician, Phys-G, also expressed her frustration after several interruptions: "I really want a system that could separate the important/critical calls from those that could wait."

Table 2. Overview of total amount of interruptions from mobile devices at the ENT clinic during observation time (only interruptions from the followed physicians' devices, except for *g*, which was a nurse's phone, brought to the physician).[a]

			Assistant physician	Chief physician
Preparatory/complementary work				
		Answered	2a, 2b	7a
		Ignored		
Outpatient ward				
	No patient			
		Answered		5a
		Ignored		
	With patient			
		Answered		1a, 1d
		Ignored		1a
Maternity/intensive/labor ward				
	No patient			
		Answered	2a,2b,1f	3a, 1d
		Ignored		
	With patient			
		Answered		6a, 1b, 1g
		Ignored	1b	1a, 1g
Surgical theatre				
		Answered		1a
		Ignored		
Meeting				
		Answered		6a, 1d
		Ignored		
Conversation room				
		Answered	1a	
		Ignored		1a, 1d
Other situations				
		Answered	6a,2b,1f	17a, 3d
		Ignored	1b, 1d	

[a]a=wireless IP-phone; b=on-call-duty pager; d=private GSM; f=personal pager; g=other

As shown in Table 2, the physicians at CY also ignored incoming pages/calls, but the ignored pages/calls were not actually completely ignored. They were either answered and the caller was told that they would call back, or they hurried with the examination and the call was ignored until they finished or had a natural break in the examination.

Clinic Settings

The setting in the CY was slightly different than at ENT. The pagers were used less, and role or function was seldom used when they contacted each other. An assistant physician, Phys-G, pointed out: "Nobody uses role or function. When they contact each other, they call or page you personally", and:

> *It is only an old habit if somebody still uses the pager to contact you while you are on on-call duty. A wireless phone should be enough to carry; actually I hate the sound from the pager, I have turned it off and I'm only using the vibration, and we don't carry the GSM-based on-call duty phone unless we're leaving the hospital.*

The observer found that the on-call duty pager was in use also at this clinic, but not as much as wireless phones. It was observed only once that a chief physician was paged during observation time, referring to a critical situation that the chief physician had to attend immediately.

Wireless Phones vs Pagers

The physicians at this clinic also stated their satisfaction in carrying a wireless phone, and the majority of physicians also argued that the phone is not more interruptive than the pager. A chief physician, Phys-K, said: "The fact that we are carrying a wireless phone these days is brilliant compared to previously when we only had pagers. We are here to work and are supposed to be available within the working hours, and it is also easier to reach the other health care workers". When questioned by the observer on interruptions from today's phones compared with yesterday's pagers, he answered:

> *With a phone in our pocket, you do not need to search for a phone whenever you need to call or answer a call. Another point is that I think there are fewer interruptions compared with when we only had pagers. Actually I think that pagers were slightly more interruptive since we had to search for a phone every time we wanted to return a call.*

However, not all the physicians at this hospital were too happy about being equipped with a wireless phone, and some of them refused to carry one and would use only pagers. The observer met an older surgeon several times who did not carry a phone. This was commented on by one of the chief physicians, Phys-F:

> *Not everybody carries a phone even though they should. They think that there is a higher threshold to page than just to call? But they also need to locate a phone and that takes a lot of their time. There is not a phone on every corner anymore. These people have not given the new phone system time; they do not want changes and only want to keep the old pager system.*

The observer discovered that this surgeon was paged several times when we met, and then borrowed the wireless phone from one of the other physicians/nurses nearby and in doing so also disturbed another health care worker. Another chief physician, Phys-K, announced after such an incident: "I am not too happy with those who "hide" away…".

Discussion

The purpose of this study was to learn about a physician's workday, focusing on wireless communication, and to identify the potential to reduce unnecessary interruptions from mobile devices: in which situations, what context, and which location they should not be interrupted. Our aim was also to generally understand the health care workers' communication pattern. Based on this study, but also previous studies done by members of our project group [13,42], we argue that there is strong evidence that interruptions from mobile devices represent a problem in hospitals and that a solution to reduce such interruptions is needed. Another important point, which some of the participants in this study address, is that they may have experienced fewer interruptions from wireless phones compared to pagers, once people has been used to carrying a phone. This issue has not been mentioned in previous studies and contrasts a finding in Sweden where widespread use of phones seemed to cause people to contact each other more often [43]. It suggests that, although interruptions might be a problem when using wireless phones, cultural shifts that develop over time in order to handle these interruptions might be able to effectively reduce the problem at some hospitals. An additional point made by one of the chief physicians, Phys-K, at CY, was that he thought a pager was slightly more interruptive since they had to locate a phone, even interrupt somebody else to borrow a phone, every time they wanted to return a page. This is related to the concern about an increased level of interruptions after introduction of wireless phones to all health care workers at a clinic. Some of the physicians in [13] expected an increased interruptions rate if wireless phones were introduced to all health care workers in the department.

The results from our study are used as input in an ongoing project [42,44,45] on designing an interruption management system to reduce some of the unnecessary interruptions a physician experiences throughout the day, and especially in situations where they should not be disturbed. Examples of such situations are: in surgery, during patient examinations, or having high-importance level conversations with patients or relatives.

Strengths and Weakness of the Study

At the ENT clinic, we experienced quiet days with few interruptions from mobile devices. On one hand, this could be explained by the Hawthorn effect [46], but on the other hand, most likely not, since this effect reduces over time; our study lasted for almost 2 months and this did not change during our observations. That is, most of the clinics' workers were aware of the project and that the observer was there to record interruptions from mobile devices and therefore may have been more careful in interrupting by calling or paging the physicians. We also observed that the nurses and other health care workers often sought each other out in-person and could therefore be aware of the context before they interrupted. During discussions, we became aware that this was not unusual and that it probably did not happen more frequently during our observations. The fact that the time we spent at the clinic was considered "quiet" by the health care workers could be explained by findings from earlier research regarding health care workers' inaccuracy when reporting incidents [47,48], which also could include interruptions. But since it also was considered "quiet" regarding patient consultations, it strengthens our belief that this was representative of the normal situation. Therefore, we do not consider the result from the study to be significantly influenced by our presence at the clinic or by health care workers' behaving differently. However, the fact that the observer dressed like a physician and was presented as a researcher made him accepted as "one of them", which seemed vital to their communication pattern and hopefully strengthened the study.

Findings That Conflict With Earlier Studies

The physicians in [13,42] were concerned about increased interruptions if they carried a wireless phone instead of a pager and that a phone call interrupts more than a page. When we asked the physicians we met at St. Olavs Hospital if they thought the phone was more interruptive than the pager after the introduction of wireless IP-phones, some of them told us that in the beginning, after providing wireless phones to every worker at the clinic, this might be the case. However, they said that after a while it would go back to the same level of

interruptions, or even fewer interruptions with the phone. Their explanation was that it was more unpleasant to call someone and interrupt their work, than just page them. Even though the participating physicians normally answered the phone and said that they would call back, they thought that a page was more interruptive than a phone call. When they were paged, they had to locate a phone, which could result in interrupting others to borrow a phone to return the call, since they never knew how important the page was and therefore felt that they had to return it right away. With a wireless phone of their own, they could just pause what they were doing, answer the phone, sense if they had to answer right away, or just say that they would return the call, and finish up what they were doing before the interruption. This interpretation from the participants could be related to the fact that wired phones were not located "on every corner" at this hospital, since everybody was supposed to carry a wireless phone, and by borrowing a phone, they felt that they were disturbing others. A solution could be a combination of a pager for incoming calls and a wireless phone for outgoing calls like some of the physicians in [13] used. However, this solution would not solve the problem of knowing who is calling or the importance of the call. Another important consideration to have in mind about interruptions by phones vs pagers, is that since the subjects of the study had carried their own wireless phone for up to 3 years, it could be related to inaccurate reporting and retention and that the users got more used to the phone compared to the pager.

The display on top of the pager seemed to be an important feature for the physicians at the University Hospital of North Norway [13]. This display made it easier for them to see who was paging them. The physicians in this study, at St. Olavs Hospital, did not think this was important. The feedback was that such a display could be useful but not critical for the overall usability, and it is difficult to know anything about the importance of the call just from a number. We also discovered that only a few of the physicians who carried a phone and a pager, did carry the pager in a way where it could be easy to read the display.

Summary and Future Work

We conclude from this study, but also from earlier studies [13,42], that physicians in hospitals are interrupted unnecessarily by mobile devices in situations where such interruptions should be avoided. The introduction of IP-based phones at St Olavs Hospital has shown that this transition in itself is not sufficient to reduce the number of interruptions for the physicians. The study illustrates the need for an integrated context-sensitive phone system that reduces unnecessary interruptions and eliminates the use of multiple communication devices.

We believe, by knowing and understanding the physicians' working conditions and the nature of such interruptions and also by involving the physicians in the design process, it is possible to make a system suited for their communication pattern and working conditions. The lack of user involvement is an important issue to consider when designing and developing eHealth applications [49]. This study contributes to such knowledge and is used as input in an ongoing project on designing and developing a context-sensitive communication system: CallMeSmart. Context sensing when developing mobile medical applications has also been a success within other studies of medical Internet research [50]. Our system is designed to reduce unnecessary interruptions from mobile devices in situations where interruptions should be avoided, such as, when they are in surgery dressed in sterile clothing, during patient examination in outpatient clinic, having high-importance level conversations with patients/relatives. It is also designed to eliminate the need of multiple communication devices for each user. The system is supposed to sense the context of each user automatically, change the physicians' availability and the phones' profile according to the context information, and also give the caller feedback about the physicians' availability. As such, we are developing a prototype for a context-sensitive mobile communication system suitable for hospitals, called CallMeSmart. CallMeSmart is being developed using input from this study in combination with outcomes from [42,44,45], and we are involving real users in the design and testing process.

Acknowledgments

This research is supported by the Research Council of Norway, Grant No. 176852/S10, and by the Center for Research-Based Innovation, Tromsø Telemedicine Laboratory, through Norwegian Research Council Grant No. 174934. We would like to thank all the physicians at St. Olavs Hospital for all their help and collaboration so far.

Conflicts of Interest

None declared.

References

1. Blum NJ, Lieu TA. Interrupted care. The effects of paging on pediatric resident activities. Am J Dis Child 1992 Jul;146(7):806-808. [Medline: 1496947]
2. Katz MH, Schroeder SA. The sounds of the hospital. Paging patterns in three teaching hospitals. N Engl J Med 1988 Dec 15;319(24):1585-1589. [doi: 10.1056/NEJM198812153192406] [Medline: 3200267]
3. 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]

4. Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, et al. An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study. J Med Internet Res 2011;13(3):e59 [FREE Full text] [doi: 10.2196/jmir.1655] [Medline: 21875849]
5. Eisenstadt SA, Wagner MM, Hogan WR, Pankaskie MC, Tsui FC, Wilbright W. Mobile workers in healthcare and their information needs: are 2-way pagers the answer? Proc AMIA Symp 1998:135-139 [FREE Full text] [Medline: 9929197]
6. Minnick A, Pischke-Winn K, Sterk MB. Introducing a two-way wireless communication system. Nurs Manage 1994 Jul;25(7):42-47. [Medline: 8044475]
7. Spurck PA, Mohr ML, Seroka AM, Stoner M. The impact of a wireless telecommunication system on time efficiency. J Nurs Adm 1995 Jun;25(6):21-26. [Medline: 7776003]
8. Laxminarayan S, Istepanian RS. UNWIRED E-MED: the next generation of wireless and internet telemedicine systems. IEEE Trans Inf Technol Biomed 2000 Sep;4(3):189-193. [Medline: 11026588]
9. van Lieshout EJ, van der Veer SN, Hensbroek R, Korevaar JC, Vroom MB, Schultz MJ. Interference by new-generation mobile phones on critical care medical equipment. Crit Care 2007;11(5):R98 [FREE Full text] [doi: 10.1186/cc6115] [Medline: 17822524]
10. Hanada E, Watanabe Y, Nose Y. A screening gate to prevent entry of mobile telephone handsets in the speaking/stand-by mode into prohibited and restricted areas. IEEE Trans Inf Technol Biomed 2000 Dec;4(4):320-323. [Medline: 11206817]
11. Parker J, Coiera E. Improving clinical communication: a view from psychology. J Am Med Inform Assoc 2000;7(5):453-461 [FREE Full text] [Medline: 10984464]
12. Coiera E. When conversation is better than computation. J Am Med Inform Assoc 2000;7(3):277-286 [FREE Full text] [Medline: 10833164]
13. Scholl J, Hasvold P, Henriksen E, Ellingsen G. Managing Communication Availability and Interruptions: A Study of Mobile Communication in an Oncology Department. In: LaMarca A, editor. Pervasive Computing. Berlin: Springer; 2007:234-250.
14. 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 [FREE Full text] [doi: 10.1016/j.ijmedinf.2007.04.006] [Medline: 17569576]
15. Khalil A, Connelly K. Improving Cell Phone Awareness by Using Calendar Information. In: Costabile M, Paterno F, editors. Human-Computer Interaction. Berlin: Springer; 2005:588-600.
16. Nelson L, Bly S, Sokoler T. Quiet calls: talking silently on mobile phones. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.: ACM Press; 2001 Presented at: SIGCHI Conference on Human factors in Computing Systems; March 31-Apr. 5, 2001; Seattle, Washington.
17. Schmidt A, Takaluoma A, Mäntyjärvi J. Context-Aware telephony over WAP. Personal and Ubiquitous Computing 2000;4(4):225-229.
18. Siewiorek D, Smailagic A, Furukawa J, Krause A, Moraveji N, Reiger K, et al. SenSay: a context-aware mobile phone. In: Proceedings of the IEEE International Symposium on Wearable Computers.: IEEE; 2003 Presented at: Seventh IEEE International Symposium on Wearable Computers; Oct. 21-23, 2003; Washington, DC.
19. Milewski A, Smith T. Providing presence cues to telephone users. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work.: ACM; 2000 Presented at: ACM conference on Computer Supported Cooperative Work; 2000; Philadelphia, PA.
20. Pedersen ER. Calls.calm: enabling caller and callee to collaborate. In: CHI '01 Extended Abstracts on Human factors in Computing Systems. New York: ACM; 2001:235-236.
21. Tang JC, Yankelovich N, Begole J, Kleek MV, Li F, Bhalodia J. ConNexus to awarenex: extending awareness to mobile users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Seattle, Washington, United States; 2001 Presented at: SIGCHI Conference on Human Factors in Computing Systems; Mar. 31-Apr. 5, 2001; Seattle, Washington.
22. Avrahami D, Gergle D, Hudson SE, Kiesler S. Improving the match between callers and receivers: A study on the effect of contextual information on cell phone interruptions. Behaviour & Information Technology 2007;26(3):247-259.
23. Acuff RD, Fagan LM, Rindfleisch TC, Levitt BJ, Ford PM. Lightweight, Mobile E-Mail for Intra-Clinic Communication. in 1997 AMIA Annual Fall Symposium. Nashville, TN 1997.
24. Sammon MJ, Karmin LSB, Peebles E, Seligmann DD. MACCS: Enabling communications for mobile workers within healthcare environments. In: MobileHCI'06. New York: ACM; 2006.
25. Munoz MA, Rodriguez M, Favela J, Martinez-Garcia AI, Gonzalez VM. Context-aware mobile communication in hospitals. Computer 2003;36(9):38-46.
26. Bardram JE, Hansen TR. The AWARE architecture: supporting context-mediated social awareness in mobile cooperation. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work.: ACM; 2004 Presented at: ACM Conference on Computer Supported Cooperative Work; 2004; Chicago, IL.
27. Bardram JE, Hansen TR, Soegaard M. AwareMedia: a shared interactive display supporting social, temporal, and spatial awareness in surgery. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work. Banff, Alberta, Canada; 2006 Presented at: 20th Anniversary Conference on Computer Supported Cooperative Work; Nov. 4-8, 2006; Banff, AB.

28. Pratt W, Reddy MC, McDonald DW, Tarczy-Hornoch P, Gennari JH. Incorporating ideas from computer-supported cooperative work. J Biomed Inform 2004 Apr;37(2):128-137. [doi: 10.1016/j.jbi.2004.04.001] [Medline: 15120659]
29. Weerakkody G, Ray P. CSCW-based system development methodology for health-care information systems. Telemed J E Health 2003;9(3):273-282. [doi: 10.1089/153056203322502669] [Medline: 14611695]
30. Blomberg J, Burrell M, Guest G. An ethnographic approach to design. In: Jacko J, Sears A, editors. The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications. Mahwah, New Jersey: Lawrence Erlbaum Associates Inc; 2003:964-986.
31. Klein HK, Myers MD. A set of principles for conducting and evaluating interpretive field studies in information systems. MIS Q 1999;23(1):67-93.
32. Walsham G. Interpretive case studies in IS research: nature and method. Eur J Inf Syst 1995;4(2):74-81.
33. Atkinson P, Hammersley M. Ethnography and participant observation. In: Denzin N, Lincoln Y, editors. Handbook of qualitative research. Thousand Oaks, CA: Sage Publications; 1994.
34. Ruotsalainen PS, Blobel BG, Seppälä AV, Sorvari HO, Nykänen PA. A conceptual framework and principles for trusted pervasive health. J Med Internet Res 2012;14(2):e52 [FREE Full text] [doi: 10.2196/jmir.1972] [Medline: 22481297]
35. Bardram JE. Pervasive healthcare as a scientific discipline. Methods Inf Med 2008;47(3):178-185. [Medline: 18473081]
36. Wasson C. Ethnography in the field of design. Human Organization 2000;59(4):377-388.
37. Westbrook JI, Ampt A. Design, application and testing of the Work Observation Method by Activity Timing (WOMBAT) to measure clinicians' patterns of work and communication. Int J Med Inform 2009 Apr;78 Suppl 1:S25-S33. [doi: 10.1016/j.ijmedinf.2008.09.003] [Medline: 18951838]
38. Westbrook JI, Braithwaite J, Georgiou A, Ampt A, Creswick N, Coiera E, et al. Multimethod evaluation of information and communication technologies in health in the context of wicked problems and sociotechnical theory. J Am Med Inform Assoc 2007;14(6):746-755 [FREE Full text] [doi: 10.1197/jamia.M2462] [Medline: 17712083]
39. Westbrook JI, Coiera E, Dunsmuir WT, Brown BM, Kelk N, Paoloni R, et al. The impact of interruptions on clinical task completion. Qual Saf Health Care 2010 Aug;19(4):284-289. [doi: 10.1136/qshc.2009.039255] [Medline: 20463369]
40. Westbrook JI, Woods A, Rob MI, Dunsmuir WT, Day RO. Association of interruptions with an increased risk and severity of medication administration errors. Arch Intern Med 2010 Apr 26;170(8):683-690. [doi: 10.1001/archinternmed.2010.65] [Medline: 20421552]
41. Solvoll T. Mobile Communication in Hospitals: What is the Problem? In: Ruckeman CR, editor. Integrated Information and Computing Systems for Natural, Spatial, and Social Sciences. Hershey, PA: IGI Global; 2013:287-301.
42. Solvoll T, Scholl J. Strategies to reduce interruptions from mobile communication systems in surgical wards. J Telemed Telecare 2008;14(7):389-392. [doi: 10.1258/jtt.2008.007015] [Medline: 18852324]
43. Scholl J, Groth K. Of organization, device and context: Interruptions from mobile communication in highly specialized care. Interacting with Computers 2012;24(5):358-373.
44. Solvoll T, Fasani S, Ravuri AB, Tiemersma A, Hartvigsen G. Evaluation of an Ascom/trixbox system for context sensitive communication in hospitals. In: Proceedings of the 8th Scandinavian Conference on Health Informatics. 2010 Presented at: 8th Scandinavian Conference on Health Informatics; 2010; Copenhagen.
45. Botsis T, Solvoll T, Scholl J, Hasvold P, Hartvigsen G. Context-aware systems for mobile communication in healthcare - A user oriented approach. In: Proceedings of the 7th Wseas International Conference on Applied Informatics and Communications. 2007 Presented at: 7th Wseas International Conference on Applied Informatics and Communications; Aug. 24-26, 2007; Athens p. 69-74.
46. Roethlisberger FJ, Dickson WJ, Wright HA. Management and the worker: an account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago. Cambridge, Mass: Harvard University Press; 1939.
47. Ampt A, Westbrook J, Creswick N, Mallock N. A comparison of self-reported and observational work sampling techniques for measuring time in nursing tasks. J Health Serv Res Policy 2007 Jan;12(1):18-24. [doi: 10.1258/135581907779497576] [Medline: 17244393]
48. Flynn EA, Barker KN, Pepper GA, Bates DW, Mikeal RL. Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. Am J Health Syst Pharm 2002 Mar 1;59(5):436-446. [Medline: 11887410]
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. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]

Abbreviations

CY: Child and Youth
ENT: Ear, Nose, and Throat
GSM: Global System for Mobile (communications)
IP: Internet Protocol

Edited by G Eysenbach; submitted 07.12.12; peer-reviewed by A Talaei-Khoei, C Granja; comments to author 04.01.13; revised version received 21.01.13; accepted 05.02.13; published 07.03.13.

Please cite as:
Solvoll T, Scholl J, Hartvigsen G
Physicians Interrupted by Mobile Devices in Hospitals: Understanding the Interaction Between Devices, Roles, and Duties
J Med Internet Res 2013;15(3):e56
URL: http://www.jmir.org/2013/3/e56/
doi:10.2196/jmir.2473
PMID:23470528

Original Paper

A Text Messaging Intervention to Improve Heart Failure Self-Management After Hospital Discharge in a Largely African-American Population: Before-After Study

Shantanu Nundy[1], MD; Rabia R Razi[2], MD, MPH; Jonathan J Dick[3], MD; Bryan Smith[4], MD; Ainoa Mayo[5], MA; Anne O'Connor[6], MD; David O Meltzer[5], MD, Ph.D

[1]Section of General Internal Medicine, Department of Medicine, University of Chicago Medical Center, Chicago, IL, United States

[2]Cedars-Sinai Heart Institute, Los Angeles, CA, United States

[3]Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, United States

[4]Department of Medicine, University of Chicago Medical Center, Chicago, IL, United States

[5]Section of Hospital Medicine, Department of Medicine, University of Chicago Medical Center, Chicago, IL, United States

[6]Section of Cardiology, Department of Medicine, University of Chicago Medical Center, Chicago, IL, United States

Corresponding Author:
Shantanu Nundy, MD
Section of General Internal Medicine
Department of Medicine
University of Chicago Medical Center
5841 S Maryland Ave
MC2007
Chicago, IL, 60637
United States
Phone: 1 7737020620
Fax: 1 7738342238
Email: shantanu.nundy@gmail.com

Abstract

Background: There is increasing interest in finding novel approaches to reduce health disparities in readmissions for acute decompensated heart failure (ADHF). Text messaging is a promising platform for improving chronic disease self-management in low-income populations, yet is largely unexplored in ADHF.

Objective: The purpose of this pre-post study was to assess the feasibility and acceptability of a text message–based (SMS: short message service) intervention in a largely African American population with ADHF and explore its effects on self-management.

Methods: Hospitalized patients with ADHF were enrolled in an automated text message–based heart failure program for 30 days following discharge. Messages provided self-care reminders and patient education on diet, symptom recognition, and health care navigation. Demographic and cell phone usage data were collected on enrollment, and an exit survey was administered on completion. The Self-Care of Heart Failure Index (SCHFI) was administered preintervention and postintervention and compared using sample *t* tests (composite) and Wilcoxon rank sum tests (individual). Clinical data were collected through chart abstraction.

Results: Of 51 patients approached for recruitment, 27 agreed to participate and 15 were enrolled (14 African-American, 1 White). Barriers to enrollment included not owning a personal cell phone (n=12), failing the Mini-Mental exam (n=3), needing a proxy (n=2), hard of hearing (n=1), and refusal (n=3). Another 3 participants left the study for health reasons and 3 others had technology issues. A total of 6 patients (5 African-American, 1 White) completed the postintervention surveys. The mean age was 50 years (range 23-69) and over half had Medicaid or were uninsured (60%, 9/15). The mean ejection fraction for those with systolic dysfunction was 22%, and at least two-thirds had a prior hospitalization in the past year. Participants strongly agreed that the program was easy to use (83%), reduced pills missed (66%), and decreased salt intake (66%). Maintenance (mean composite score 49 to 78, *P*=.003) and management (57 to 86, *P*=.002) improved at 4 weeks, whereas confidence did not change (57 to 75, *P*=.11). Of the 6 SCHFI items that showed a statistically significant improvement, 5 were specifically targeted by the texting intervention.

Conclusions: Over half of ADHF patients in an urban, largely African American community were eligible and interested in participating in a text messaging program following discharge. Access to mobile phones was a significant barrier that should be addressed in future interventions. Among the participants who completed the study, we observed a high rate of satisfaction and preliminary evidence of improvements in heart failure self-management.

(J Med Internet Res 2013;15(3):e53) doi:10.2196/jmir.2317

KEYWORDS

heart failure; self-care; patient education; cellular phone; text messaging; African Americans

Introduction

Despite major scientific advances, heart failure continues to be a common and costly condition, and each year over 1 million people are admitted to an inpatient setting for acute decompensated heart failure (ADHF) [1,2]. National attention has turned toward reducing 30-day readmissions for ADHF, partially because financial penalties from the Centers for Medicare and Medicaid Services (CMS) for higher than expected rates of readmissions began in October 2012 [3,4]. This problem is particularly salient to hospitals serving larger proportions of African Americans because these patients have higher rates of readmissions than white patients [5-7]. Thus, there is an urgent need for low-cost solutions to reduce heart failure readmissions in African Americans.

Approximately 40% of ADHF hospitalizations are preventable because of varied factors, such as dietary indiscretion, medication nonadherence, and lack of timely medical consultation [8]. Patient nonadherence to heart failure drugs ranges from 30% to 60% and nonadherence to lifestyle recommendations ranges from 50% to 80%, with higher rates occurring in socioeconomically disadvantaged groups [9].

Mobile technology, in particular text messaging (also known as SMS: short message service), is emerging as a promising platform for chronic disease management in low-income populations [10,11], in part because it has high rates of utilization across socioeconomic groups [12,13]. Recent studies of mobile phone-based telemonitoring interventions in heart failure have demonstrated mixed success in reducing heart failure readmissions [14,15]. However, these interventions were not designed or evaluated in African American patients and may not be as effective in these populations. For one, these interventions typically required Internet- or Bluetooth-enabled phones, which may not always be available in these communities. Second, these interventions largely focused on telemonitoring rather than self-management support. The results of a recent large multicenter clinical trial found that self-management education was effective only in low-income patients (<$30,000 family income) [16], suggesting that a one-size-fits-all approach to improving heart failure outcomes does not work, and that self-management education is particularly effective in vulnerable health populations.

In this study, we pilot-tested a text message–based self-management intervention in an urban, largely African American population for 30 days following hospitalization for ADHF. Our study aims were to assess the feasibility and acceptability of the intervention and to test the hypothesis that the intervention was associated with improvements in self-management.

Methods

Patient Recruitment

After study approval was obtained from the Institutional Review Board, patients were recruited from the University of Chicago Medical Center (UCMC) inpatient cardiology service. Patients were recruited until target enrollment was achieved from November 2011 to January 2012 for a 4-week study. Informed consent was obtained prior to recruitment. Eligible patients included adult patients over the age of 18 years who were diagnosed with ADHF either with decreased or preserved systolic function as determined by the admitting physicians. Because the intent of the study was to provide self-management support, individuals who were not their own primary caregiver, were being discharged to a rehab facility, or who had poor mental status (Mini-Mental score<17) were excluded [17]. In addition, patients who did not have access to a personal mobile phone were not eligible for enrollment.

Consent was obtained from the attending physicians for contacting their patients for enrollment. Discharge planners were encouraged to notify the study team of any new admissions for ADHF. Study participants received $30 for study participation and to offset the costs of text messages. They were also provided with a scale to measure their weight.

Study Design

The pilot was designed as a single-arm prospective study. The primary endpoint was change in the Self-Care of Heart Failure Index (SCHFI), a well-described measure of self-management in heart failure [18], which was administered at enrollment and at the end of the 30-day intervention. In addition, a mobile phone usage survey was administered on enrollment [19], and demographic and clinical data were obtained through chart review. At the completion of the intervention, a telephone-based patient experience survey, including Likert-scale and open-ended questions, was administered [19].

Study Intervention

A text message communication platform developed for health researchers, SMS-Care (mHealth Solutions LLC, New York, NY, USA), was used for this study. Participants were enrolled in SMS-Care prior to discharge from the hospital and began receiving text messages on their personal cell phones the day after discharge. Text messages were composed to reflect literature published for patient education by the American Heart Association [20]. In addition, language similar to that used by

the UCMC inpatient heart failure education team was incorporated into the texts. For a 30-day period, each participant received automated messages in the following domains:

1. Medication adherence: a daily reminder message (eg, "Time to take your heart failure medications") and a biweekly adherence question (eg, "Did you take all your heart failure medications today?")

2. Dietary compliance: educational messages (eg, "Remember to avoid salt. Items high in salt include canned soups, deli meats, and fried foods.")

3. Appointment adherence: a reminder 48 hours before and the day of their cardiology or primary care follow-up appointments (eg, "Please remember to go to your appointment with Dr. Smith today. Take all your medicines with you.")

4. Heart failure signs and symptom recognition: warning signs of heart failure (eg, "Know the signs of fluid buildup: your weight going up, swelling of your legs, and having trouble breathing.")

5. Management if experiencing symptoms ("Have you noticed that your legs are swollen or are you having trouble fitting into your shoes? If yes, call your physician.")

6. Health care navigation: knowing how to get in touch with cardiologist, obtaining medications after discharge ("If you have not done so already, make sure you have all the medicines that you were discharged on."), and dealing with complications of paying for medications ("If you're having trouble paying for your medicines, please make sure your doctor knows about this.")

Each participant's text message programming was personalized to reflect his/her medication regimen and follow-up appointments. Participants were provided a tutorial on receiving, reading, and sending text messages on enrollment. At the time of enrollment, each patient was sent a test message and replied to it, ensuring basic competency with the use of text messages. Participants were regularly reminded that the system was automated and was not an emergency response system.

Data Analysis

Per the most recent scoring procedure [18], raw scores from the SCHFI were tabulated into standardized 100-point scales: maintenance, management, and confidence. Preintervention and postintervention scores for each scale were compared using paired *t* tests. Individual items were compared using Wilcoxon rank sum tests. Stata version 11 was used for the analysis (StataCorp LP, College Station, TX, USA).

Results

Study Recruitment and Sample Characteristics

Of 61 patients initially identified for ADHF, 51 were successfully approached for enrollment prior to discharge and 27 agreed to participate. Twelve of the patients approached did not own a personal cell phone. An additional 6 patients did not meet inclusion criteria because they failed their Mini-Mental exam (n=3), needed a health care proxy (n=2), or were hard of hearing (n=1). Only 3 patients approached for the study who met all inclusion criteria refused to participate. Of the 27 patients who met inclusion criteria, 15 were successfully enrolled. The remainder were unable to be enrolled due to logistical barriers (eg, off the floor, discharged early). Eight of 15 enrollees completed the text messaging portion of the study. Of the remaining 7 participants, 2 died, 1 was admitted to a subacute facility, and 4 had technology issues, including their cell phone being disconnected. A total of 6 participants completed the entire study including preintervention and postintervention surveys.

All but 1 participant in the study was African American (Table 1). The average age of participants was 50 years (range 23-69) and 40% (6/15) were female. The majority had Medicaid as primary or secondary insurance with Medicare or were uninsured. Approximately half of participants (47%, 7/15) had systolic heart failure. The mean ejection fraction for those with systolic dysfunction was 22%, and two-thirds (67%, 10/15) of all participants had at least 1 prior hospitalization in the past year. Most participants were on evidence-based heart failure therapies on admission including angiotensin-converting enzyme (ACE) inhibitors (53%, 8/15) and beta-blockers (86%, 13/15).

Cellular Phone Use

Most participants (93%, 14/15) carried their cell phone with them always or almost always (Table 2). All participants reported being somewhat or very comfortable with text messaging, although actual usage varied widely from 0 to 60 text messages per day. All but 1 participant had an unlimited text messaging plan, and only one-third (33%, 5/15) of participants in our sample had a smartphone capable of accessing the Internet and running applications (apps).

Table 1. Participant characteristics at enrollment (N=15).

Baseline characteristic		Statistic
Age, mean (range)		50 (23-69)
Race, n (%)		
	African American	14 (93)
	White	1 (7)
Gender, n (%)		
	Women	6 (40)
	Men	9 (60)
Mini-Mental status, mean (range)		21 (18-22)
Insurance status, n (%)		
	Medicare only	3 (20)
	Medicaid only	3 (20)
	Dual eligible	4 (27)
	Private insurance	3 (20)
	Uninsured	2 (13)
Medical history		
	Preserved ejection fraction (EF) heart failure, n (%)	8 (53)
	Average EF for those with systolic heart failure (%), mean (range)	22.4 (9-47)
	Hypertension, n (%)	11 (73)
	Diabetes, n (%)	8 (53)
	Smoker, n (%)	5 (33)
	ACE inhibitor on admission, n (%)	8 (53)
	Beta-blocker on admission, n (%)	13 (86)
Admissions		
	1 admission in the prior year, n (%)	6 (40)
	2 or more admissions in the prior year, n (%)	4 (27)
	Number of admissions in the prior year, median (range)	1 (0-7)

Table 2. Prior participant experience with cellular phone calling and text messaging (N=15).

Baseline characteristic	n (%)
Owns a smartphone	5 (33)
Unlimited text messaging	14 (93)
Landline in addition to cell phone	5 (33)
Frequency with which carry cell phone	
Always	12 (80)
Almost always	2 (13)
Sometimes	1 (6)
Comfort level making or receiving calls	
Very comfortable	14 (93)
Somewhat comfortable	1 (6)
Total calls made/received per day	
0	0 (0)
1-5	3 (20)
6-10	3 (20)
11-20	4 (27)
>20	5 (33)
Have used text messaging feature before	15 (100)
Comfort level using text messages	
Very comfortable	10 (66)
Somewhat comfortable	5 (33)
Total text messages sent/received per day	
0	1 (6)
1-5	4 (27)
6-10	5 (33)
11-20	1 (6)
>20	4 (27)

Participant Engagement

Although not required, participants were encouraged to text back comments or responses to questions sent via text message. Although responses were not read by research staff during the course of the study, response rate was considered to be a marker of patient engagement and could inform future program design. Over the 30-day intervention, participants sent an average of 5.7 text messages (range 0-27) or approximately 1 message every 5 days. Five participants did not send any text messages; 2 participants sent over 20 messages. Interestingly, both of these participants were near the median in terms of prior usage of text messaging, and 1 reported being only somewhat comfortable with text messaging prior to the study.

Participant Experience

All (100%, 6/6) participants reported the highest level of satisfaction with the mobile phone–based heart failure self-management program (Table 3). Although most participants (66%, 4/6) strongly agreed that the text messaging system was easy to use and was helpful in improving self-management, a minority (33%, 2/6) strongly disagreed with these statements. Despite this, all participants agreed that they would recommend the program to a friend or family member.

Table 3. Participant evaluation postintervention (n=6).

Survey question	Likert scale response, n (%)			
	Strongly agree	Moderately or slightly agree	Slightly or moderately disagree	Strongly disagree
Overall, I was satisfied with this study	6 (100)			
It was easy to receive and read the text messages from the research team	5 (83)			1 (17)
It was easy to send text messages to the research team	4 (66)			2 (33)
I found the text message reminders to be helpful at decreasing the number of pills I missed	4 (66)	1 (17)		1 (17)
I found the text message reminders to be helpful at decreasing the amount of salt in my diet	4 (66)	1 (17)		1 (17)
I found the text message reminders to be helpful at decreasing the number of doctor visits that I missed	3 (50)	1 (17)		2 (33)
I would recommend this cell phone reminder system to my friends/family that have heart failure	5 (83)	1 (17)		

During the open-ended survey, participants reported that the intervention improved self-management directly by providing reminders, but also indirectly by increasing disease awareness and reinforcing the importance of self-management. They liked that the system served as a reminder and provided feedback. One participant stated, “I knew I wasn’t being forgotten,” and another that, “It’s nice to know that someone cares.” Only 2 participants identified elements they did not like about the system: 1 complained that it was hard for him to text and the other wanted more text messages. Participants suggested improvements such as providing more instruction on how to text with the system and making the cost of texting with the system free.

Change in Heart Failure Self-Management

Participants’ responses to the SCHFI suggested that the intervention was associated with improvements in heart failure self-management (Table 4). On a 100-point standardized scale, self-care maintenance improved from 49 to 78, representing an increase of 28 points (95% CI 15-42, *P*=.003). Self-care management increased from 57 to 86, or 30 points (95% CI 17-42, *P*=.002). There was no statistically significant change in self-care confidence (57 to 75, 95% CI –6 to 43, *P*=.11).

Of the 22 individual items comprising the SCHFI, improvements were seen in 6 items: weighing self, eating a low salt diet, forgetting to take medicines, avoiding getting sick, contacting physician in case of worsening symptoms, and confidence in evaluating symptoms (Table 5). The text messaging intervention specifically targeted content areas covered by 7 of the 22 SCHFI items. For example, the intervention did not include any messages about exercising or using pill counters. Of the 6 individual SCHFI measures that improved, 5 were specifically targeted by the texting intervention. Of the 7 items targeted by the intervention, 5 improved and 2 had no statistically significant change.

Table 4. Self-Care of Heart Failure Index (SCHFI) scales preintervention and postintervention (n=6).

SCHFI scale	Preintervention score, mean (95% CI)	Postintervention score, mean (95% CI)	Difference, mean (95% CI)	*P* value
Maintenance	49 (38-61)	78 (68-88)	28 (15-42)	.003
Management	57 (42-71)	86 (72-100)	30 (17-42)	.002
Confidence	57 (32-81)	75 (70-80)	19 (–6 to 43)	.11

Table 5. Self-care Heart Failure Index (SCHFI) individual items preintervention and postintervention (n=6).

SCHFI Item	Preintervention		Postintervention		*P* value[a]
	Mean	SE	Mean	SE	
Maintenance					
Weighing self[b]	1.6	0.27	3.3	0.21	*.03*
Eating low salt diet[b]	2.4	0.27	3.7	0.21	*.03*
Forgetting to take medicines (lower is better)[b]	1.9	0.22	1.3	0.21	*.02*
Keeping to low salt diet when eating out	1.5	0.22	3	0.52	.05
Checking ankles for swelling[b]	3.3	0.25	3.8	0.17	.16
Keeping appointments[b]	3.3	0.23	3.7	0.33	.45
Avoiding getting sick	2.9	0.27	3.8	0.17	*.046*
Physical activity	2.4	0.24	3.2	0.4	.16
Exercising for thirty minutes	1.3	0.15	2.5	0.56	.11
Using a pill system	3.1	0.34	2.7	0.61	.51
Management					
Likelihood of contacting physician[b]	2	0.3	3	0.37	*.02*
Realization of symptoms	2.4	0.31	4	0	.09
Likelihood of reduction in salt intake	3	0.3	4	0	.16
Likelihood of reduction in fluid intake	2.8	0.28	4	0	.08
Likelihood of taking extra diuretics	2.8	0.28	3	0.63	.28
Thought that remedy helped symptoms	1.9	0.31	3.3	0.33	.11
Confidence					
Confidence in evaluating symptoms[b]	2.8	0.22	3.3	0.21	*.046*
Confidence in remaining free of heart failure	2.4	0.13	3	0.37	.16
Confidence in following treatment advice	2.7	0.23	3.5	0.22	.10
Confidence in recognizing changes	3.1	0.27	3.7	0.21	.23
Confidence in action to relieve symptoms	2.6	0.24	3.2	0.4	.52
Confidence in evaluation of remedy	2.6	0.24	2.8	0.31	.73

[a] Statistically significant values in italics.

[b] Relates to specific text messages.

There were no statistically significant differences in preintervention SCHFI measures between the 6 patients who completed the entire intervention and the 9 patients who did not complete the intervention or who were lost to follow-up.

Discussion

We report the results of a text message–based self-management intervention in patients discharged from the hospital with acute decompensated heart failure (ADHF). To the best of our knowledge, this is the first study of a text messaging intervention for ADHF piloted in an urban, largely African American population.

Principal Results

The text messaging intervention was associated with improvements in self-care maintenance and management. Guidance from the developers of the SCHFI suggests that an improvement in either of these scales of 0.5 standard deviations or 8 points is clinically relevant and that a cutoff score of 70 can be used to judge self-care adequacy [18]. We observed increases in maintenance and management of 28 and 30 points, respectively, suggesting that our findings are clinically significant. Moreover, none of participants scored below 70 preintervention for maintenance, whereas all but 1 improved to above 70 at follow-up; for management, only 1 participant scored above 70 preintervention, but following the intervention all participants scored above the 70 threshold.

Without a control, these results must be interpreted with caution. In addition to a Hawthorne effect, participants were discharged from the hospital and likely seen in clinic at least once or twice between the preintervention and postintervention periods. Self-management teaching at any of these time points, or simply

improvement in health status, may account for the observed improvements. However, it is notable that the specific SCHFI items that improved were generally those targeted by the texting intervention, implicating a causal link. Future studies should validate these preliminary findings in a controlled trial.

A major aim of this study was to access the feasibility of text messaging interventions in our patient population. Among patients who met the inclusion criteria, 90% agreed to participate in the study, which is high for our institution. Most participants in our study were comfortable with text messaging and had unlimited text messaging plans. This suggests that in our study population, among those with access to a cell phone, text messaging is a familiar and acceptable means of health care communication. In contrast, only one-third of patients had smartphones, suggesting that mobile phone–based interventions requiring Web access or apps would have low feasibility in our study population.

However, 24% of patients approached for recruitment did not own a personal cell phone. This differs from national surveys in which high rates of mobile phone access in low-income populations are observed [12]. However, these surveys were conducted in the general population, not hospitalized patients with complex medical needs. Future studies in low-income populations should consider providing participants cell phones to improve accessibility.

There were considerable challenges to texting in our study. Although all participants reported comfort in texting prior to enrollment, we observed low participant response rates and requests for additional training in texting. Although most participants reported high levels of satisfaction with the system, a few found the system difficult to use and not helpful in improving self-management. More research is needed in how to design technologies that are usable across a wide range of patients and how to best target mobile phone–based interventions to patients most likely to benefit.

Limitations

A major limitation of our study was the rate of completion and loss to follow-up. Of the 15 participants enrolled, only 6 received the entire intervention and completed preintervention and postintervention surveys. Although this largely reflects the challenges with research in our study population, it may bias the results toward those who responded favorably to the system. The study also had no control group, so the effects on self-management should be interpreted with caution. Finally, this was a single institution study with a small sample size and the results may not generalize to other patient populations.

Comparison With Prior Work

Prior mobile phone-based heart failure interventions have had mixed results when studied in clinical trials [14,15]. In contrast to our intervention, which focused on providing self-management support, these studies used mobile phones largely as telemonitoring devices. The Telemonitoring to Improve Heart Failure Outcomes (Tele-HF) study, a large negative multicenter clinical trial, demonstrated that telemonitoring is not effective in reducing death or hospitalization in patients recently admitted for ADHF [21]. In our study, rather than facilitating in-between visit care by providers, we aimed at building self-care skills through reminders, encouragement, and patient education. Our study builds on prior qualitative work suggesting that the effects of mobile phone-based interventions in heart failure may go beyond telemonitoring and increase patient empowerment [22], and a recent clinical trial suggesting that self-management education is particularly effective in low-income patients with heart failure [16]. It also extends findings from other chronic diseases, including diabetes, asthma, and human immunodeficiency virus/acquired immunodeficiency syndrome, on the utility of text message-based interventions in urban African Americans [19,23-25].

Conclusions

Text messaging may be a useful tool for improving heart failure self-management in urban African Americans after hospital discharge. More research is needed to target enrollment to those patients most likely to benefit and to evaluate outcomes in a controlled trial.

Acknowledgments

Dr Nundy is supported by the Agency for Healthcare Research and Quality Health Services Research Training Program (T32 HS00084). This research was supported by a University of Chicago Clinical and Translational Science Award (UL1 RR024999).

Conflicts of Interest

Dr Nundy previously cofounded and was part owner of mHealth Solutions, LLC, a mobile health software company which developed the text message communication platform used in this study, but currently has no financial relationship or affiliation with the company. Dr Dick cofounded and is part owner of mHealth Solutions, LLC. The other authors have no conflicts of interest to disclose.

References

1. Writing Group Members, Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation 2010 Feb 23;121(7):e46-e215 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192667] [Medline: 20019324]

2. Centers for Disease Control and Prevention (CDC). Centers for Disease Control and Prevention (CDC). 2010 Jan 01. Heart failure fact sheet URL: http://www.cdc.gov/DHDSP/data_statistics/fact_sheets/docs/fs_heart_failure.pdf [accessed 2013-02-20] [WebCite Cache ID 6EZze2sa2]
3. Centers for Medicare & Medicaid Services, Department of Health and Human Services. Federal Register, Vol 77, No 92. 2012 May 11. Proposed rules URL: http://www.gpo.gov/fdsys/pkg/FR-2012-05-11/html/2012-11421.htm [accessed 2013-02-20] [WebCite Cache ID 6EZzgljEt]
4. Centers for Medicare & Medicaid Services. 2012 Apr 24. Readmission reduction program URL: http://cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program.html/ [accessed 2013-02-20] [WebCite Cache ID 6EZzmWMwM]
5. Joynt KE, Orav EJ, Jha AK. Thirty-day readmission rates for Medicare beneficiaries by race and site of care. JAMA 2011 Feb 16;305(7):675-681 [FREE Full text] [doi: 10.1001/jama.2011.123] [Medline: 21325183]
6. Bhalla R, Kalkut G. Could Medicare readmission policy exacerbate health care system inequity? Ann Intern Med 2010 Jan 19;152(2):114-117. [doi: 10.1059/0003-4819-152-2-201001190-00185] [Medline: 19949133]
7. Hernandez AF, Curtis LH. Minding the gap between efforts to reduce readmissions and disparities. JAMA 2011 Feb 16;305(7):715-716. [doi: 10.1001/jama.2011.167] [Medline: 21325191]
8. Vinson JM, Rich MW, Sperry JC, Shah AS, McNamara T. Early readmission of elderly patients with congestive heart failure. J Am Geriatr Soc 1990 Dec;38(12):1290-1295. [Medline: 2254567]
9. van der Wal MH, Jaarsma T. Adherence in heart failure in the elderly: problem and possible solutions. Int J Cardiol 2008 Apr 10;125(2):203-208. [doi: 10.1016/j.ijcard.2007.10.011] [Medline: 18031843]
10. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
11. Boland P. The emerging role of cell phone technology in ambulatory care. J Ambul Care Manage 2007;30(2):126-133. [doi: 10.1097/01.JAC.0000264602.19629.84] [Medline: 17495681]
12. Smith A. Mobile access 2010. Washington, DC: Pew Internet & American Life Project; 2010 Jul 07. URL: http://pewinternet.org/Reports/2010/Mobile-Access-2010.aspx [accessed 2013-02-20] [WebCite Cache ID 6Ea09sLl6]
13. Blumberg SJ, Luke JV. National Center for Health Statistics. 2012 Jun. Wireless substitution: early release of estimates from the National Health Interview Survey URL: http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201206.pdf [accessed 2013-02-20] [WebCite Cache ID 6Ea0QLAfK]
14. Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, MOBITEL Investigators. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009;11(3):e34 [FREE Full text] [doi: 10.2196/jmir.1252] [Medline: 19687005]
15. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
16. Powell LH, Calvin JE, Richardson D, Janssen I, Mendes de Leon CF, Flynn KJ, HART Investigators. Self-management counseling in patients with heart failure: the heart failure adherence and retention randomized behavioral trial. JAMA 2010 Sep 22;304(12):1331-1338. [doi: 10.1001/jama.2010.1362] [Medline: 20858878]
17. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Validation of a telephone version of the mini-mental state examination. J Am Geriatr Soc 1992 Jul;40(7):697-702. [Medline: 1607586]
18. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. J Cardiovasc Nurs 2009 Dec;24(6):485-497 [FREE Full text] [doi: 10.1097/JCN.0b013e3181b4baa0] [Medline: 19786884]
19. Dick JJ, Nundy S, Solomon MC, Bishop KN, Chin MH, Peek ME. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. J Diabetes Sci Technol 2011 Sep;5(5):1246-1254 [FREE Full text] [Medline: 22027326]
20. American Heart Association. HeartHub for Patients. Heart failure URL: http://www.hearthub.org/hc-heart-failure.htm [accessed 2013-02-20] [WebCite Cache ID 6Ea0UqOzU]
21. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
22. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Perceptions and experiences of heart failure patients and clinicians on the use of mobile phone-based telemonitoring. J Med Internet Res 2012;14(1):e25 [FREE Full text] [doi: 10.2196/jmir.1912] [Medline: 22328237]
23. Nundy S, Dick JJ, Solomon MC, Peek ME. Developing a behavioral model for mobile phone-based diabetes interventions. Patient Educ Couns 2013 Jan;90(1):125-132. [doi: 10.1016/j.pec.2012.09.008] [Medline: 23063349]
24. Macdonell K, Gibson-Scipio W, Lam P, Naar-King S, Chen X. Text messaging to measure asthma medication use and symptoms in urban African American emerging adults: a reasibility study. J Asthma 2012 Oct 30. [doi: 10.3109/02770903.2012.733993] [Medline: 23106138]

25. Cornelius JB, Dmochowski J, Boyer C, St Lawrence J, Lightfoot M, Moore M. Text-messaging-enhanced HIV intervention for African American adolescents: a reasibility study. J Assoc Nurses AIDS Care 2012 Nov 2. [doi: 10.1016/j.jana.2012.06.005] [Medline: 23122907]
26. Nundy S, Dick JJ, Solomon MC, Peek ME. Developing a behavioral model for mobile phone-based diabetes interventions. Patient Educ Couns 2013 Jan;90(1):125-132. [doi: 10.1016/j.pec.2012.09.008] [Medline: 23063349]

Abbreviations

ADHF: acute decompensated heart failure
EF: ejection fraction
SCHFI: Self-Care of Heart Failure Index
UCMC: University of Chicago Medical Center

Edited by G Eysenbach; submitted 17.08.12; peer-reviewed by S Inglis, CJ(Wu; comments to author 22.12.12; revised version received 07.01.13; accepted 08.01.13; published 11.03.13.

Please cite as:
Nundy S, Razi RR, Dick JJ, Smith B, Mayo A, O'Connor A, Meltzer DO
A Text Messaging Intervention to Improve Heart Failure Self-Management After Hospital Discharge in a Largely African-American Population: Before-After Study
J Med Internet Res 2013;15(3):e53
URL: http://www.jmir.org/2013/3/e53/
doi:10.2196/jmir.2317
PMID:23478028

Original Paper

Internal Versus External Motivation in Referral of Primary Care Patients with Depression to an Internet Support Group: Randomized Controlled Trial

Benjamin W Van Voorhees[1*], MD, M.Public Health; Robert C Hsiung[2*], MD; Monika Marko-Holguin[1*], MSS; Thomas K Houston[3,4], MD, M.Public Health; Joshua Fogel[5], Ph.D; Royce Lee[6], MD; Daniel E Ford[7], MD, M.Public Health

[1]Section of General Pediatrics Department of Pediatrics, University of Illinois at Chicago, Chicago, IL, United States

[2]Dr. Bob LLC, Chicago, IL, United States

[3]Division of Health Informatics and Implementation Science, University of Massachusetts Medical School, Worcester, MA, United States

[4]eHealth QUERI, Bedford VA Medical Center, Bedford, MA, United States

[5]Department of Finance and Business Management, Brooklyn College of the City University of New York, Brooklyn, NY, United States

[6]Department of Psychiatry, The University of Chicago, Chicago, IL, United States

[7]Division of General Internal Medicine, Department of Medicine, Johns Hopkins Medical Institutions, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Benjamin W Van Voorhees, MD, M.Public Health
Section of General Pediatrics Department of Pediatrics
University of Illinois at Chicago
840 S Wood St
Suite 1437, MC 856
Chicago, IL, 60612
United States
Phone: 1 3129960023
Fax: 1 3129960243
Email: bvanvoor@uic.edu

Abstract

Background: Depressive disorders and symptoms affect more than one-third of primary care patients, many of whom do not receive or do not complete treatment. Internet-based social support from peers could sustain depression treatment engagement and adherence. We do not know whether primary care patients will accept referral to such websites nor do we know which methods of referral would be most effective.

Objective: We conducted a randomized clinical trial to determine whether (1) a simple generic referral card (control), (2) a patient-oriented brochure that provided examples of online postings and experience (internal motivation), or (3) a physician letter of recommendation (external motivation) would generate the greatest participation in a primary care Internet depression treatment support portal focused around an Internet support group (ISG).

Methods: We used 3 offline methods to identify potential participants who had not used an ISG in the past 6 months. Eligibility was determined in part by a brief structured psychiatric interview based on the Patient Health Questionnaire-9 (PHQ-9). After consent and enrollment, participants were randomly assigned to 1 of 3 groups (control, internal motivation, or external motivation). We constructed a portal to connect primary care patients to both fact-based information and an established ISG (Psycho-Babble). The ISG allowed participants to view messages and then decide if they actually wished to register there. Participation in the portal and the ISG was assessed via automated activity tracking.

Results: Fifty participants were assigned to the 3 groups: a motivation-neutral control group (n=18), an internal motivation group (n=19), and an external motivation group (n=13). Of these participants, 31 (62%) visited the portal; 27 (54%) visited the ISG itself. The internal motivation group showed significantly greater participation than the control group on several measures. The external motivation group spent significantly less time logged onto the portal than the control group. The internal motivation group showed significantly greater participation than the external motivation group on several measures.

Conclusions: Referral of primary care patients with depressive disorders and symptoms to an ISG is feasible even if they have never previously used one. This may best be accomplished by enhancing their internal motivation.

Trial Registration: Clinicaltrials.gov: NCT00886730; http://clinicaltrials.gov/show/NCT00886730 (Archived by WebCite at http://www.webcitation.org/6F4981fDN)

(J Med Internet Res 2013;15(3):e42) doi:10.2196/jmir.2197

KEYWORDS

depressive disorder; Internet; primary care; support groups

Introduction

Mood disorders have a lifetime prevalence of 20.8% [1-3]. Many Americans with a depressive disorder do not receive any treatment, and many of those who do receive treatment either do not receive high quality treatment or do not complete a full course of treatment [4]. Chronic care and collaborative care models have demonstrated benefit in improving process of care, symptoms, and functional outcomes [5-7]. A key component of these models appears to be support provided by case management, which may act by improving adherence or increasing patient activation [6,8]. However, these models are often expensive and cumbersome to implement [7]. As the number and quality of Internet models that provide self-directed psychotherapy [9], education [10], or social support [11] increases, so does the prospect for improving depression outcomes without costly person-to-person interventions or complex organizational changes [12]. Technology-based depression interventions have shown considerable promise in treating depression in select populations. However, the breadth of appeal of Internet-based treatment outside of highly structured surroundings, such as in school settings for children and adolescents, is not known [13-19].

Internet sites offering peer-to-peer interaction may be particularly attractive to laypersons. In Australia, peer-to-peer and gaming models are widely used, including Reach Out! which provides psycho-education regarding common adolescent mental health problems [20]. Similarly, in the United States, social media is commonly used to address health concerns. Approximately 28% of Internet users are estimated to have contacted an Internet support group (ISG) for a medical condition or personal problem [21]. ISGs have several advantages, including transcending geographic barriers, facilitating disclosure (people will often tell a computer more than to another person), and increasing access to diverse sources of information. There is some evidence that use of ISGs for depression may be associated with both reductions in depressed mood and an increase in learning about depression [11,22]. Although we know little about how to engage primary care patients with Internet-based models, it is theorized that social engagement will be an essential component [17,23]. Unfortunately, many Internet-based interventions have high dropout rates, ranging up to 75% of all participants [24]. Participants' longer and more consequential involvement with the Internet site and its direct personal relevance to the user predict greater behavior change in Internet-based programs [25,26]. Peer support and frequent updates may support participation in ISGs [27,28]. A prerequisite for further evaluations of the potential benefits of ISGs for primary care patients with depressive illness is developing an effective method of referral and ascertaining key predictors of participation. Self-determination theory provides a framework for evaluating possible referral methods. In this model, internal source motivations based upon preferences for autonomy, competence, and connection are superior to external source motivations based upon financial incentives or recommendation of an authority figure [2].

The aim of this study was to determine the most feasible and effective methods to refer primary care patients with depressive disorders and symptoms to an ISG. We examined whether baseline attitudes based on the theories of motivational interviewing, planned behavior, and self-determination predicted Internet site participation [2,29], and we assessed whether a structured approach to goal setting would increase ISG participation [30]. We conducted a randomized controlled trial to compare 3 different methods of referring primary care patients with depressive disorders and symptoms to an ISG: emphasizing internal motivation (patient-oriented brochure), emphasizing external motivation (physician recommendation), and employing neither motivational strategy (referral card with Internet address). Interventions that require minimal physician involvement are particularly useful in actual practice settings. Because of this, the simple referral card is an important consideration. Our hypotheses were that (1) primary care referral (all methods) would be effective in engaging participants with the ISG (ie, >30% of participants would visit the ISG), (2) the physician letter recommendation (external motivation) group would show the greatest participation, (3) attitudes toward ISG participation would be important predictors of subsequent use, and (4) email reminders would increase participation.

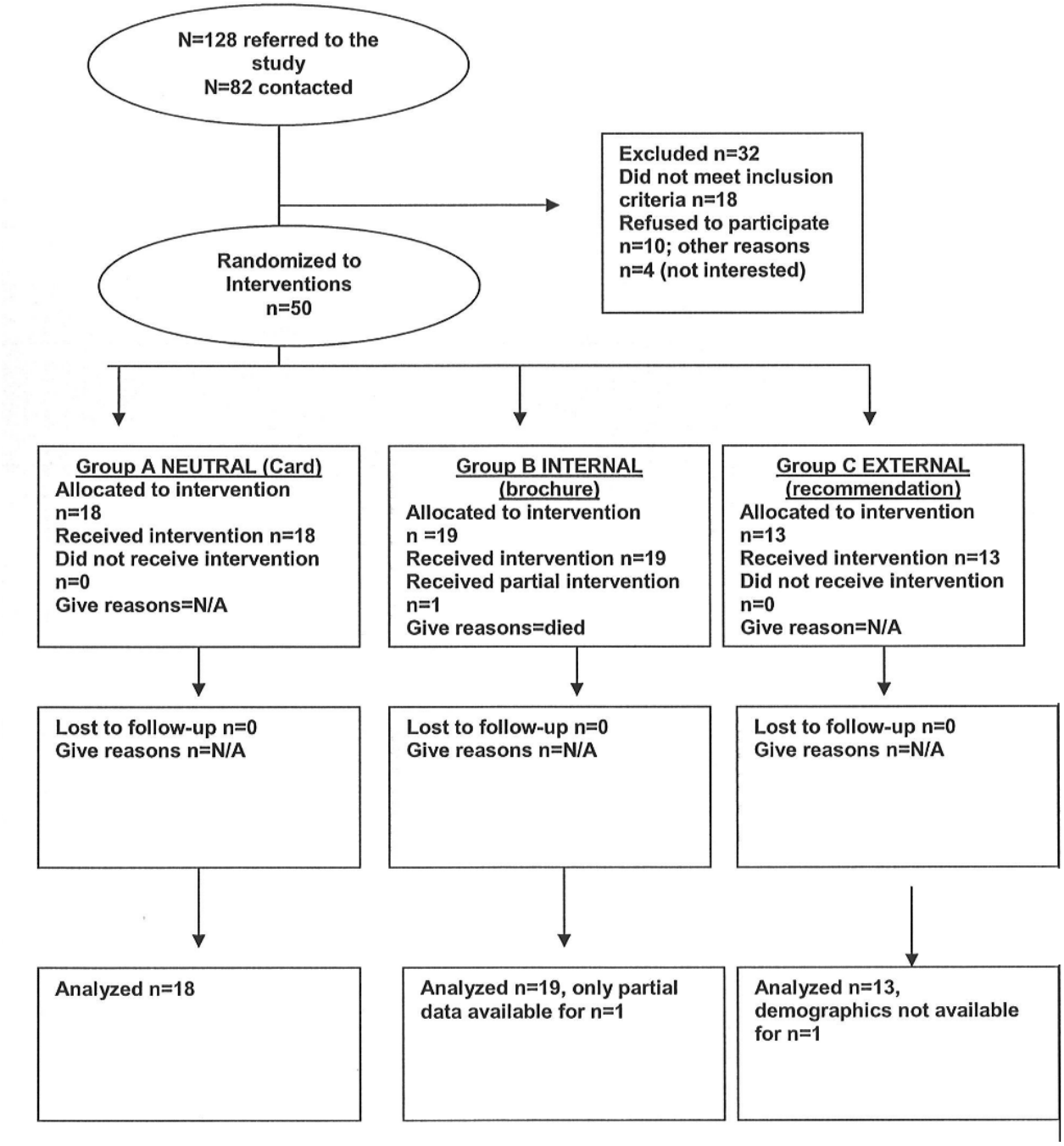
Methods

Study Design

This study included 9 key steps from initial recruitment to study completion. We identified potential participants using 3 methods: (1) reading a poster in the waiting room and self-referred, (2) completion of a statement of interest form in their physician's office after discussion with their physician, or (3) completing a statement of interest at a public information table in the primary care office. In step 2, those recruited through reading a poster and self-referred or completing a statement of interest form after discussion with their physician provided consent to be called by study staff to learn more about the study (those who signed up at the public information table in the primary care office already knew about the study). Step 3 was

a phone eligibility assessment. After providing consent for phone eligibility assessment, the study coordinator conducted a brief structured psychiatric interview based on the Patient Health Questionnaire-9 (PHQ-9) [31]. Following eligibility confirmation by both the principal investigator (internist-pediatrician) and co-investigator (psychiatrist), participants were offered enrollment in the study (step 4). There was approximately a 7-day delay between initial eligibility assessment and actual enrollment. At step 5, the study coordinator met the participants at or near the primary care office and conducted informed consent and enrollment. At this time, the participants also completed a baseline written questionnaire. In step 6 after consent and enrollment, participants were randomly assigned (using a sealed envelope with equal likelihood of assignment to all arms) to one of 3 groups: (1) referral card (neutral motivation group), (2) patient-oriented brochure providing examples of online postings and experience (internal motivation group), or (3) physician letter of recommendation (external motivation group) (see Figure 1). Participation on the Internet was assessed via automated activity tracking (step 7). Study staff contacted each participant 6 to 8 weeks after enrollment to evaluate depressed mood and any concerns about their website experiences (step 8). In the final step (step 9), each participant was asked to complete a written poststudy questionnaire. The study was approved by the University of Chicago Institutional Review Board. The trial was prospectively registered (ClinicalTrials.gov NCT00886730).

Figure 1. CONSORT diagram.



Inclusion and Exclusion Criteria

All eligible participants met the following criteria: (1) a PHQ-9 score of 8 or above with either depressed mood or anhedonia and/or were considering treatment for depressed mood, (2) were accepting of at least one form of treatment for depression, (3) had not viewed or posted messages more than once in the past month on any ISG, (4) did not self-report being diagnosed with bipolar disorder by a health professional, (5) was 18 years of age or older, (6) attends a primary care clinic and had visited it in the past 6 months, and (7) had Internet access for the subsequent 4 weeks, had been on the Internet at least 3 times, and was able to use email. We excluded those considered to be at high risk for suicide attempts, which included those with past psychiatric hospitalizations, past suicide attempts, diagnosed bipolar disorder, or a score of greater than 1 on the PHQ-9 suicide assessment item (bothered by thoughts that he/she would be better off dead or of hurting him/herself in some way more than half the days in past 2 weeks), or who reported intent for self-harm as per assessment.

Setting

We recruited physicians at 6 primary care sites in urban and suburban areas of Chicago. These included a federally qualified health center (n=1), a private practice (n=1), university-affiliated off-site practices (n=2), a student health clinic (n=1), and an on-site academic resident practice (n=1). Physicians were recruited using lunchtime education programming.

Adaptations Made After Study Initiation

Several barriers to implementation were addressed after the study was fielded. During the course of the study, 2 practice sites closed and several physicians left the practices. After initial attempts using nurses to approach patients were unsuccessful, physicians were asked to approach appropriate patients directly to determine if they might be interested in the study. The original intention was to have the recommendation letter be from the participant's personal physician. Instead, the letter was signed by the principal investigator (a primary care physician) because of time feasibility limitations.

Randomization, Blinding, and Concealment

We randomly assigned participants (using sealed envelopes) to 1 of the 3 groups described subsequently. A data manager who was not involved in the study execution prepared the sealed envelopes. Participants were aware of group assignments. The main outcome measures were collected in an automated fashion online. The database manager and the safety caller (described subsequently) were blinded to group assignments.

Randomized Intervention Groups

Neutral Motivation Orientation (Referral Card)

Participants received a simple 3"×5" card with the name of the ISG and the following description: "Online support/coping group for people feeling blue, stressed, depressed, down, low, or sad" (see Multimedia Appendix 1 for full text).

Internal Motivation Orientation (Patient-Centered Brochure)

Participants received an 8"×11" handout that provided a more complete description of the ISG. The handout was based on a patient perspective with samples of Internet postings from current users. This card emphasized peer-to-peer support and did not mention health care organizations or health care provider endorsements. The information addressed potential barriers to participation: participants would not be identified, posting would not take much time, information from peers could be checked for accuracy with other peers and providers, and patients could learn how to tell their providers about their activities on the Internet (see Multimedia Appendix 1 for full text).

External Motivation Orientation (Physician Recommendation)

Participants received a medical recommendation in the form of a letter signed by the principal investigator, a primary care physician (see Multimedia Appendix 1 for full text).

Email Reminder

An email reminder was sent to all participants who did not visit the portal within 7 days of enrollment. Email reminders have been associated with higher levels of participation [27].

Internet Portal Intervention

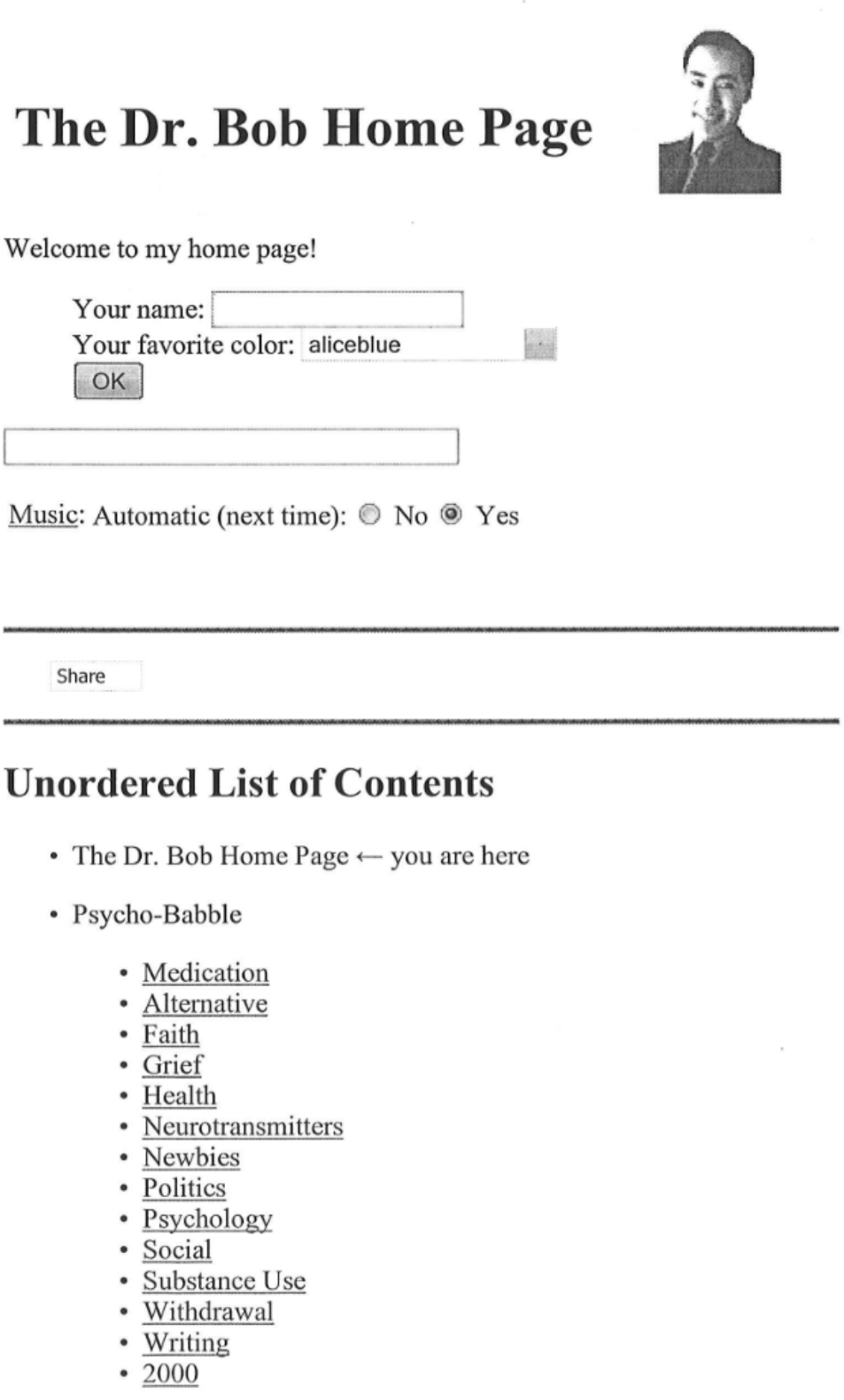
We constructed a portal to connect primary care patients to both fact-based information and an established ISG. Two focus groups identified several key ideas with regard to the development of a primary care portal: (1) offer opportunities to learn from others, (2) offer opportunities to help others, (3) offer fact-based resources to evaluate accuracy of information, and (4) protect users from disturbing content, such as self-harm or salacious topics. To address these needs [32-34], we chose to build a portal (Psycho-Babble) that would: (1) offer access to fact-based information sites (eg, NIMH [35] and MOODGym [36]), (2) present aspects of the ISG that focus group participants suggested would be of greatest interest, and (3) offer access to the ISG. Participants could use the portal to visit the ISG and also the other websites. The ISG allowed participants to view messages and then decide if they actually wished to register there.

Internet Support Group

Psycho-Babble [37,38] is a mental health peer support group started in 1998 (see Figure 2). Online mental health groups can be classified as autonomous self-help groups or support groups led by mental health professionals. Psycho-Babble is a hybrid that combines the empowerment of self-help with the supportiveness of a milieu maintained by a mental health professional. The asynchronous online (message board) format makes the group more accessible and in some ways safer than groups that meet face-to-face [38]. The original forum continues to focus on biological treatments, and additional forums have been added for psychological treatments, complementary and alternative treatments, religious faith, social support, and discussion of the administration of the site. The portal through which participants reached Psycho-Babble highlighted 6 forums. The introduction in the main forum lists 3 other forums, and a

link takes participants to a list of all 10. We expected participants to have been aware of the other Psycho-Babble forums. Higher user control has been associated with favorable perceptions by participants [39]. We believed Psycho-Babble would be a valuable ISG to use because of its well-developed participant base with history, loyalty, and repeat visits [40].

Figure 2. Screenshot of Psycho-Babble home page.



Physician and Nurse Training

As part of this intervention, investigators met with each primary care provider and nurse at the participating primary care sites for 30 minutes during a lunchtime education program. This program included (1) explanation of the study design and the ISG and (2) summary of problems in depression care, self-management approaches currently used by individuals with depression, and the state of research on ISGs. A group of resident physicians received additional training on how to conduct randomized trials as an added inducement to participate in the study.

Safety Monitoring

A structured approach to managing participant safety was used. After eligibility assessment, all participants with elevated depressed mood received a recommendation for a medical evaluation. All those with self-harm ideation received a structured assessment to determine need for service. While enrolled in the study, the existing Psycho-Babble adverse event reporting mechanisms were available to participants. A safety call was made by a licensed clinical social worker 6 to 8 weeks after enrollment [37,38].

Outcome Measures

We evaluated use of the portal and the ISG by using data automatically logged by the portal and ISG servers (Table 1). We originally planned to assess outcomes across several categories: visits, time logged in, posts viewed, and posts posted. Few posts were posted; therefore, we added a broader measure of interaction, communications, which we use to refer to all submissions of data by participants. Participants submitted data not only for postings, but also to preview posts, to take the quiz required to register at the ISG, to enter a synchronous (real-time) chat room, to actually chat, and to contact the ISG administration.

Portal Measures

The portal server tracked visits and minutes logged in by each username. Participants were automatically logged out after 6 minutes. Time actually engaged with the portal may have been longer (eg, if they browsed the NIMH site for 30 minutes) or shorter (eg, if they stepped away from their computer after 3 minutes). If participants used the portal to visit the ISG, the time they spent at the ISG was included in the time they were logged in to the portal. However, if they visited the ISG directly, it was not. It was possible to access the ISG directly because it was open to the public and not restricted to study participants.

Internet Support Group Measures

At the ISG, viewing posts was unrestricted, but submitting data, including postings, was restricted to registered group members. The ISG server automatically tracked viewing of posts by participants by using cookies (information stored by a server on a user's computer) which it set when participants visited the ISG from the portal. Because participants could visit the ISG directly or delete the cookies, viewing of posts (ie, ISG visits) by study participants may have been undercounted. The ISG server automatically tracked communications, including postings, by username.

Independent Variables

Sociodemographic Factors

We obtained self-reported age (years), gender, race/ethnicity, and education level (eg, some high school or high school graduate).

Mood

For the participant's dimensional measures of depressed mood, we used the 10-item Center for Epidemiologic Studies Depression Scale (CES-D) (Cronbach alpha =.87) [41] and the PHQ-9 [31] score (Cronbach alpha=.77).

Self-efficacy

Perceived control was assessed with the Mastery Scale [42]. The 7 items on the scale measure the extent to which participants see themselves as being in control of the forces that significantly affect their lives. Higher scores indicate greater self-efficacy (Cronbach alpha=.81).

Loneliness

The loneliness item of the CES-D was used to report loneliness. Respondents were asked the frequency of feeling lonely in the past week on a 4-point scale (0=less than a day, 1=1-2 days, 2=3-4 days, and 3=5-7 days).

Self-determination Theory

Self-determination theory posits that humans seek experiences in which they will develop autonomy, connection, and competence. Furthermore, the stronger the internal motivation (as opposed to external factors), the higher the quality and enduring the motivation [2]. Based on a focus group, we developed items to address attitudes in each of these domains and 1 additional one, concern for adverse experiences (see Multimedia Appendix 1). We measured 9 items based on a Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree) (Cronbach alpha=.69).

Theory of Planned Behavior

We revised the items from a previous questionnaire, which was modified based on the preventive health model and the theory of planned behavior [43,44] for the purpose of participation in our ISG. According to the theory of planned behavior, intention is the most proximal cognitive measure to actual behavior. Intention is directly influenced by attitudes and beliefs toward a behavior (eg, attitudes toward intervention), subjective norms (eg, concerns with regard to family, peer, or employer opinions), and perceived behavioral control (ie, controllability and self-efficacy) [29]. We measured items based on a Likert-type scale (1=strongly disagree to 5=strongly agree) for each domain (see Multimedia Appendix 1). The reliability (Cronbach alpha=.89) and predictive validity of the original instrument has previously been demonstrated in adolescents [45].

Transtheoretical Model of Change

Items were adapted from Miller and Rollnick's 3-item assessment of motivation [46]. In terms of validity, this scale and many individual items predicted adherence to an Internet-based depression prevention intervention (modified form) [45]. On a scale from 1 to 10, participants rated the

importance, their ability, and their readiness to overcome depressed mood over the next 2 months. Higher numerical values indicate higher levels for agreement with the item.

Shared Decision Making

We previously demonstrated improved outcomes with shared decision making for depression [47]. We asked participants to rate agreement with "I can talk with my providers in a way so my preferences for treatment are included" on a Likert-type scale (1=strongly disagree to 5=strongly agree).

Analysis

Descriptive statistics were calculated for the demographic variables. For between-group categorical comparisons, we used Pearson's chi-square test or the Fisher exact test when there were <5 observations per cell. For continuous outcomes, we used analysis of variance (ANOVA) for between-group comparisons at the same time points. For continuous between-group data with skewed distribution, we used the Mann-Whitney test for 2-level comparisons. We used a similar analytical approach for comparisons between the 3 groups and also used the Kruskal-Wallis test for 3-level comparisons with a skewed distribution. We developed regression models for 3 main outcomes: (1) total time on-site, (2) number of posts viewed, and (3) number of posts attempted. Separate regression analyses were performed for the outcome variables. For total time on-site, the outcome had skewed data and there were values of zero, precluding a direct logarithmic transformation. Therefore, we added a value of 1 to all the variables so that they could be logarithmically transformed. Linear regression was then performed. For the count data (posts viewed and posts attempted), negative binomial regression was used. Because of the limited number of observations, we chose to develop only 2 models for each analysis. Model 1 was a univariate analysis. Model 2 adjusted for all items that were significant at $P<.05$ on Model 1 for each outcome. STATA 11 (StataCorp LP, College Station, TX, USA) was used to conduct the analyses. We also calculated effect sizes for between-group comparisons with statistical significance (Cohen's *d*) [48].

Sample Size Calculations and Interim Analyses

A sample size of 225 was estimated to have a power of 0.8 to detect the difference between both the neutral group and the internal or external group (50%, 35%, and 10% participation). This sample size was not achieved because of an extensive delay in starting the study due to an extensive Institutional Review Board review (1 year). An interim analysis was conducted for purposes of safety review and end of grant funding. On the basis of the results of the interim analysis, which showed significant differences between randomization groups for the main study endpoints and significant predictors for measures of participation, the data monitoring and safety board (DSMB) and the investigators jointly agreed to end the trial.

Table 1. Outcome variables.

Items	Description	Source[a]
Sessions and time		
Number of sessions	Number of individual visits	Portal
Percent those in sessions with at least 1 session	Percent with at least 1 visit	Portal
Total time	Number of minutes spent logged in (participants were timed out after 6 minutes of inactivity)	Portal
Minutes elapsed per session	Number of minutes/session	Portal
Viewing		
Posts viewed	Number posts viewed by participant	ISG
Percent who viewed at least once	Percent who viewed at least 1 post	ISG
Posts viewed by each participant who registered	Number of posts viewed by those who registered on the ISG	ISG
Boards viewed	Number of boards viewed per participant	ISG
Boards viewed by each participant who registered	Number of boards viewed per participant who registered with the ISG	ISG
Posting		
Communications	Posts and data combined	ISG
Data submitted	Information submitted to the portal	ISG
Posts submitted	Number of posts submitted	ISG
Posts attempted	Number of posts attempted	ISG
Percent of those who attempted at least once	Percent who attempted to post at least 1 time	ISG
Posts attempted by each participant who registered	Number of posts submitted who registered on the ISG	ISG
Posts posted	Number of posts attempted who registered on the ISG	ISG
Percent of those who posted at least once	Percent who attempted to post at least 1 time who registered on the ISG (created post, but failed to confirm)	ISG
Posts posted by each participant who registered	Mean number of posts submitted who registered on the ISG	ISG
Visitation		
Visited portal or ISG	Percent who visited either the portal or directly went to the ISG	ISG and portal
Participants who visited portal	Percent who visited the portal	ISG and portal
Participants who registered ISG	Percent who registered on the ISG	ISG and portal

[a] ISG: Internet support group.

Results

Sample Characteristics

A total of 50 participants were assigned to the 3 groups. At baseline, there were no significant differences between the 3 groups on any demographic characteristics (Table 2). With regard to the race/ethnicity of the sample, 47% were white (23/49), 41% were African American (20/49), 2% were Hispanic (1/49), 6% were Asian (3/49), and 4% were other (2/49). One participant did not report race/ethnicity. The mean age of participants was 37.49 years (SD 17.15). This was a moderately depressed cohort with PHQ-9 scores ranging between 7.22 to 10.24 for the 3 randomization groups with a mean for the entire sample of 9.18 (SD 4.64). With regard to education and socioeconomic status, 41% (20/48) were college graduates and the mean reported income was US $45,407 (SD 59,496). Fifty-one percent (25/49) reported discussing depression with their physician, 43% (21/49) reported being treated before for depression, and 31% (15/49) reported a family member had been depressed in the past. One participant did not provide any demographic data. Over the course of the study (November 2008 to June 2009), 128 individuals were referred by their physicians or answered advertisements posted in the clinics, of which 82 were successfully contacted and had eligibility assessed with 32 being excluded (see Multimedia Appendix 1). Reasons for exclusion included 10 who were eligible but did not enroll, 4 who were not interested, and 18 who did not meet eligibility criteria (8 with past psychiatric hospitalization, 8 with a self-harm attempt, and 2 with bipolar disorder). Of those who

could be reached (82), only 4 reported no interest in the study. Primary care physicians recruiting for the study did not report many refusals for initial contact by study staff. Of those referred to the study staff for contact, 64% were contacted. Partial data were available for 2 participants, 1 who died of natural causes after enrollment (internal motivation group) and 1 participant enrolled in the external motivation group.

Participation for Entire Cohort

A total of 31 of 50 participants (62% of the entire study cohort) visited the portal or the ISG directly, whereas 27 of 50 (54%) registered on the ISG itself. Participants who visited the portal logged in for a mean of approximately 2 hours (118.5 minutes) each. Participants who visited the ISG viewed an average of 20.9 posts, communicated an average of 5.8 times, and posted an average of 0.4 posts. Table 3 shows the frequency distributions of the number of visits to the portal, time on the portal, and the number of posts viewed, communications, and posts posted. There was a skewed distribution spent on time on portal with 20 (40%) having no time, but 1 (2%) with 1301 minutes. Similarly, 4 (8%) participants viewed more than 46 posts, whereas 25 (50%) viewed no posts. In terms of communications, most (40/50, 80%) made no communications whereas a few (3/50, 6%) communicated more than 15 times. In terms of actual postings, 47 (94%) never posted, whereas 1 posted 1, 4, and 6 times. Most posts and replies were not replied to (new posts: 3 replied to, 6 not replied to).

Participation by Randomization Group

The internal motivation group had significantly greater participation for several participation measures than the neutral motivation group with more individual data submissions (mean 7.53 vs mean 0.39, *P*=.01, effect size [ES] 0.67, CI –0.02 to 1.32), were more likely to attempt to submit at least once (43% vs 6%, *P*=.02) and had significantly more posts attempted overall and by those who registered (mean 5.21 vs mean 0.17, *P*=.02, ES 0.62, CI –0.07 to 1.27 and mean 0.19 vs mean 0.01, *P*=.02, ES 0.55, CI –0.19 to 1.26, respectively). The neutral motivation group spent significantly more time logged onto the ISG as compared to the external motivation group (mean 9.27 vs mean 0.23, *P*=.04, ES 0.44, CI –0.29 to 1.15). There was only one 3-group comparison (see Multimedia Appendix 1) that demonstrated significance, which was the percent posting at least once (*P*=.02). The internal motivation group had significantly greater participation for several participation measures than the external motivation group (Tables 4 and 5). Specifically, members of the internal motivation group had more individual data submissions (mean 7.53 vs mean 0.54, *P*=.03, ES 0.60, CI –0.14 to 1.31), were more likely to submit at least once (42% vs 8%, *P*=.05), were more likely to have visited with the portal or ISG directly (79% vs 46%, *P*=.05) and were more likely to have attempted to post overall and by those who registered (mean 5.21 vs mean 0.23, *P*=.05, ES 0.56, CI –0.18 to 1.27 and mean 0.19 vs mean 0.01, *P*=.05, ES 0.55, CI –0.19 to 1.26, respectively). No other significant differences were observed.

Email Reminders

During the course of the study, 22 participants received an email reminder because they had not visited the portal within 7 days of enrollment. Of these 22 participants, 5 subsequently visited the portal; of those 5 participants, 3 did so within 2 days of receiving the email reminder.

Predictors of Participation

Time On-Site

In the univariate analysis, demographic characteristics, severity of depressed mood, and self-efficacy did not predict minutes logged in (Table 6). One attitude item, “might read troubling comments,” was a statistically significant negative predictor (*P*=.03). Three items, “participating is important,” “want to do what physicians want me to do,” and “going to a depression Internet support website is easy,” were statistically significant (*P*=.049, *P*=.02, and *P*=.047, respectively), and another 3, “benefits outweigh difficulty,” “want to do what family thinks,” and “physicians think I should go” were not statistically significant (*P*=.06, *P*=.08, and *P*=.07, respectively). The motivation items and participation in decision making were not statistically significant. In the exploratory multivariate analysis of the significant univariate predictors, only the item of “might read troubling comments” remained statistically significant (*P*=.03).

Posts Viewed

Demographic characteristics and mood did not predict posts viewed (Table 7). Self-efficacy was not statistically significant (*P*=.09). No attitude items were statistically significant. Two intention items were statistically significant (“friends think I should go” *P*=.03, and “physicians think I should go” *P*=.01) and 1 was not statically significant (“want to do what friends think” *P*=.08). The motivation items and participation in decision making were not statistically significant. In the exploratory multivariate analysis of the significant univariate predictors, none remained statistically significant.

Posts Attempted

Demographic characteristics did not predict posts attempted (Table 8). Baseline CES-D, was not statically significant (*P*=.06). Self-efficacy was statistically significant (*P*=.04). Two intention items, “my family supports me going” and “physicians think I should go,” were statistically significant (*P*=.04 and *P*=.046, respectively), and another 2, “friends think I should go” and “want to do what physicians think,” were not statically significant (*P*=.06 and *P*=.06). Participation in decision making was not statistically significant. In the exploratory multivariate analysis of the significant univariate predictors, only self-efficacy remained statistically significant (*P*=.05).

Safety and Unintended Events

Fourteen of 49 participants received follow-up calls from the principal investigator related to concerns identified by the assessment caller (ie, depressed mood or self-harm ideation). One participant died of natural causes shortly after enrollment. No study participant reported adverse experiences while on the ISG and there were no self-harm events.

Table 2. Demographics of study participants.

Items	Participants (N=50)			*P* value
	Neutral motivation (card) n=18	Internal motivation (brochure) n=19	External motivation (recommendation) n=13	
Age (years), mean (SD)	34.83 (12.72)	39.39 (16.83)	34.42 (19.49)	.62
PHQ-9 score, mean (SD)	9.13 (5.84)	10.24 (3.72)	7.22 (3.83)	.29
Gender, n (%)				
Male	4 (22)	6 (33)	5 (42)	.54
Female	14 (78)	12 (67)	7 (58)	
Ethnicity, n (%)				
White	10 (56)	8 (44)	5 (42)	.71
Non-white	8 (44)	10 (56)	7 (58)	
Education, n (%)				
Some high school	3 (17)	1 (6)	2 (17)	.72
High school graduate	1 (6)	2 (11)	2 (17)	
Some college	5 (28)	7 (39)	5 (42)	
College graduate	9 (50)	8 (44)	3 (25)	
Marital status, n (%) [a]				
Married	1 (6)	1 (24)	1 (8)	.62
Divorced/separated/widowed	2 (11)	2 (12)	1 (8)	
Never married	3 (15)	11 (65)	10 (83)	
Talk to health care provider, n (%)				
Yes	11 (61)	10 (59)	3 (27)	.19
Treatment, n (%)				
Yes	8 (57)	9 (69)	4 (50)	.75
Family treatment, n (%)				
Yes	5 (28)	8 (47)	2 (17)	.22
Counseling, n (%)				
Yes	10 (56)	12 (71)	6 (50)	.56
Income (US$), mean (SD)	35,025.00 (30,344.51)	57,553.67 (84,277.86)	43,056.67 (52,325.73)	.90

[a] Marital status does not equal the total sample size in each group because data were not obtained from some participants.

Table 3. Time on portal, posts viewed, and number of communications for entire cohort.

Items	Participants, n (%)[a]
Time on portal (minutes) [b]	
0	20 (40)
1	6 (12)
2	2 (4)
3	2 (4)
4	3 (6)
5	2 (4)
6	1 (2)
9	1 (2)
13	1 (2)
15	1 (2)
16	1 (2)
19	1 (2)
30	1 (2)
31	1 (2)
45	1 (2)
51	1 (2)
53	1 (2)
89	1 (2)
543	1 (2)
774	1 (2)
1301	1 (2)
Number of posts viewed	
0	25 (50)
1	1 (2)
2	4 (8)
3	2 (4)
4	2 (4)
5	1 (2)
6	1 (2)
9	2 (4)
12	1 (2)
15	2 (4)
16	1 (2)
17	2 (4)
20	1 (2)
46	1 (2)
54	1 (2)
55	1 (2)
92	1 (2)
152	1 (2)

Items	Participants, n (%)[a]
Number of communications	
0	40 (80)
3	1 (2)
4	1 (2)
7	3 (6)
8	1 (2)
12	1 (2)
16	1 (2)
34	1 (2)
59	1 (2)

[a] For simplicity, all those enrolled in the study but who did not visit the site are recorded as 0.

[b] The time in minutes is the actual time each participant spent logged in to the portal.

Table 4. Participation measures reporting only with $P<.10$.[a]

Items	Neutral motivation (card) n=18	Internal motivation (brochure) n=19	External motivation (recommendation) n=13
Time, mean (SD)			
Total time		74.68 (228.13)	4.5 (8.63)
Minutes logged (portal)		56.21 (175.66)	4.08 (8.3)
Minutes logged (ISG)	9.27 (26.59)	18.47 (53.24)	0.23 (0.6)
Posting			
Data submitted, mean (SD)	0.39 (1.65)	7.53 (15.09)	0.54 (1.94)
Post attempted, mean (SD)	0.17 (0.71)	5.21 (11.54)	0.23 (0.83)
Percent of those who attempted to post at least once, n (%)	1 (6)	2 (42)	1 (8)
Post attempted by each participant who registered, mean (SD)	0.01 (0.03)	0.19 (0.43)	0.01 (0.03)
Portal and ISG, mean (%)			
Visited portal or ISG		15 (78.95)	6 (46.15)

[a] For full data, see Multimedia Appendix 1.

Table 5. Group comparisons of participation measures reporting only with $P<.10$.[a]

Items	2-group *P* value		
	Neutral (card) vs internal (brochure)	Neutral (card) vs external (recommendation)	Internal (brochure) vs external (recommendation)
Time			
Total time			.08
Minutes logged (portal)			.09
Minutes logged (ISG)		.04	.08
Posting			
Data submitted	.01		.03
Post attempted	.02		.05
Percent of those who attempted at least once	.02		.05
Post attempted by each participant who registered	.02		.05
Portal and ISG			
Visited portal or registered ISG			.05

[a] For full data, see Multimedia Appendix 1.

Table 6. Predictors of time on-site reporting only with $P<.10$.[a]

Predictors	n	Mean (SD)	Model 1			Model 2		
			Beta	SE	*P*	Beta	SE	*P*
Attitudes toward Internet including self-determination theory								
Concern for adverse experiences								
I might read troubling comments about depression on the Internet	48	3.67 (1.12)	-0.58	0.25	.03	-0.62	0.28	.03
Preventive health model questions								
Beliefs about intervention								
Participating in a depression Internet social support website is an important thing to do	46	3.59 (1.02)	0.58	0.29	.049	0.27	0.32	.40
Attitudes toward intervention								
The benefits of a depression Internet social support site outweigh any difficulty	46	3.48 (0.86)	0.66	0.34	.06			
Social norms								
I want to do what the members of my immediate family think I should do	43	2.98 (1.03)	0.54	0.30	.08			
Physicians and other health care professionals want me to go to an Internet intervention site	45	3.02 (1.2)	0.47	0.25	.07			
I want to do what the physicians and other health care professionals want me to do	45	3.4 (0.91)	0.8	0.32	.02	0.45	0.36	.23
Self-efficacy								
Visiting this depression Internet social support site is an easy thing to do	46	3.89 (0.97)	0.62	0.30	.047	0.43	0.29	.16

[a] For full data, see Multimedia Appendix 1.

Table 7. Predictors of posts viewed reporting only with $P<.10$.[a]

Predictors	n	Mean (SD)	Model 1			Model 2		
			Beta	SE	*P*	Beta	SE	*P*
Social factors								
Self-efficacy scale	45	0.44 (0.87)	0.63	0.37	.09			
Preventive Health Model								
Social norms								
My close friends support me going to depression social support Internet site	45	2.78 (1.13)	0.85	0.4	.03	0.13	0.55	.81
I want to do what my close friends think I should do	45	2.82 (0.98)	0.93	0.53	.08			
Physicians and other health care professionals want me to go to an Internet intervention site	45	3.02 (1.20)	0.88	0.33	.01	0.79	0.49	.11

[a] For full data, see Multimedia Appendix 1.

Table 8. Predictors of posts attempted reporting only those with $P<.10$.[a]

Predictors	n	Mean (SD)	Model 1			Model 2		
			Beta	SE	*P*	Beta	SE	*P*
Mood								
Baseline CES-D	43	9.19 (4.64)	0.28	0.15	.06			
Social factors								
Self-efficacy scale	45	0.44 (0.87)	1.56	0.76	.04	1.5	0.77	.05
Attitudes toward Internet including self-determination theory								
Connection seeking								
Able to help others by sharing my experiences on the Internet	47	3.53 (1.08)	2.05	1.07	.06			
Preventive health model								
Social norms								
My close friends support me going to depression social support Internet site	45	2.78 (1.13)	2.04	1.07	.06			
My immediate family supports me going to a depression Internet social support site	45	2.71 (1.25)	1.58	0.75	.04	1.07	0.57	.06
Physicians and other health care professionals want me to go to an Internet intervention site	45	3.02 (1.2)	0.81	0.41	.046	0.23	0.43	.60
I want to do what the physicians and other health care professionals want me to do	45	3.4 (0.91)	1.28	0.67	.06			
Transtheoretical model of change/MI								
Readiness	48	7.81 (2.73)	0.65	0.36	.07			

[a] For full data, see Multimedia Appendix 1.

Discussion

This study suggests that low-cost and relatively low-intensity primary care referral methods are useful for engaging ISG-naive patients with depressed mood with a depression ISG. Our results support hypothesis 1 that referral of primary care patients to an Internet-based social support group portal was effective (>30% visited the site in all groups). Hypothesis 2 was not supported, whereby a patient-oriented brochure primarily focused on eliciting internal motivations demonstrated greater participation on multiple measures with moderate effect sizes than a generic medical letter of recommendation. Support for hypothesis 3

was attained with the findings that items from several models, but particularly the theory of planned behavior (in particular social norms), predicted participation. Furthermore, there was some evidence that email reminders to eligible patients were associated with the first visit to the portal (hypothesis 4).

This is the first trial that we are aware of comparing low-intensity methods for referring primary care patients with depression to online social support or other online depression resources. The percentage of patients visiting the portal (79%) in the internal motivation (brochure) group was similar to those in other studies using minimal contact approaches in medical settings of brief advice (85%) and email reminder (77%) [49,50] versus participation in a mental health information website without any referral method in the general population (24%) [10]. We previously demonstrated that more intensive interventions (motivational interviewing versus brief advice) increases participation for Internet-based depression prevention interventions for adolescents, with similar moderate effect sizes [50]. Others have demonstrated that email-based reminders may increase participation [49]. Resnicow [51] demonstrated that health messaging using brochures that relate closely to the patients' current concerns and preferred method of motivation may significantly increase prohealth behaviors such as diet change. The preparation of our brochure based on themes identified in 2 focus groups may have been particularly helpful. Similarly, relying on the brochure may have enabled our participants to develop higher quality intrinsic motivations as opposed to external ones. It is noteworthy that the greatest impact of the referral methods was on measures of posting, rather than simply viewing, which perhaps supports the premise that higher quality motivation is critical to enacting higher levels of engagement. Goal setting and having an established motivational structure to bring to the Internet experience has been noted to influence participation and could explain the added benefit of the intrinsic group (brochure) as well [30]. Similarly, perhaps the letter (external motivation) tended to crowd out intrinsic motivations in favor of external ones and serve as a negative incentive [52]. The finding that email reminders may increase participation is consistent with the findings of a prior systematic review [20]. However, it should be noted that the level of active involvement, such as posting, was very low (only 4 individuals posted).

Attitudes and clinical factors, rather than demographic factors, provided a broader understanding of the process by which participants may choose to participate in our mental health ISG. Typically, factors that predict participation in a self-directed psychotherapy website include demographic factors (younger age, higher education, and greater illness severity), clinical factors (greater dysfunctional thinking and illness severity), health services factors (referral by a mental health professional), and attitudes (importance, self-efficacy, and perceived benefits) [50,53,54]. Although the attitudinal predictors of participation in an ISG are similar to those for Internet-based self-directed psychotherapy, social norms appear to be more important factors in predicting participation in an ISG [44,50,55-57]. Perhaps participants considered using an Internet-discussion group to be more of a social choice than a treatment decision or event, and when considering something unfamiliar they defer to the judgment of others. It is uncertain how much participation on an ISG could be as sensitive to comparisons of activity with other individuals such as that occurs with gaming on the Internet [28].

These sites reflected the challenges of contemporary primary care, as 2 of the practice sites closed during the study and were relocated as a result of the financial crisis of 2008. In terms of internal validity, the participants were effectively randomized with no significant between-group differences and received 3 distinct interventions. Although small, the sample size demonstrated between-group differences and helped illuminate several preliminary associations. However, several key limitations with internal validity should be recognized: (1) the sample was quite small and the smallest group had only 13 participants (we only analyzed 12), (2) a letter from the patient's own physician personal letter of recommendation might have performed better than a generic one, (3) the closure of 2 clinics during the study disrupted recruiting and in some cases delivery of the intervention elements, (4) delivery of emails to all groups and direct physician recruitment of nearly all participants may have resulted in reduced between-group differences in participation, and (5) poor engagement may simply have reflected lack of easy access to the Internet or experience with interfaces like those of the portal and ISG. The observation that the external motivation (recommendation) group assignment had a mean PHQ-9 of 7.22, which was below the eligibility criteria, suggests this group had disproportionate improvement in depressed mood before actual enrollment. Two forms of bias may explain this: (1) practice effect bias suggests that if the same test is administered within 2 weeks of time then practice effect bias is present (ie, a person remembers the questions and might try to change the answers), or (2) experimenter bias (participants may have guessed we were expecting to find decreased depression and answered to satisfy us). This could suggest that this group varied from the others in some unmeasured characteristic, such as motivation, but this seems unlikely to alter the result.

In terms of external validity, we should note several concerns as well. Although it is possible that some participants were not ISG-naive, the exclusion of only those who had not visited a depression ISG within the past 6 months is unlikely to have resulted in recruitment of participants who were significantly more disposed to depression ISGs than the general population. Because we did not collect data on prior use of other depression ISGs, we cannot determine the extent to which prior use may have influenced participation in the ISG of this study. Similarly, the recruiting methods may have self-selected for more motivated individuals, a fact that may suggest caution in estimating future effectiveness in primary care. However, recruiting physicians did not report many patients refusing referrals to the study and 64% of those referred were contacted. Given that this was perhaps a somewhat motivated sample, just half of the participants actually registered on the ISG. Lack of Internet access in this relatively lower income population and a somewhat outdated interface may have reduced participation versus a population with greater Internet access and skill [58]. Similarly, the degree to which the website may not have met

current expectations for frequently updated materials, such as the style of the site, may have affected participation [58,59].

In conclusion, primary care patients may be amenable to referral to Internet-based social support and information portals, even if they have not utilized ISGs in the past. Relatively low-cost materials and minimal physician interaction may be required to accomplish this referral process. Motivational approaches focused on intrinsic factors may be superior to extrinsic ones. Referral to Internet-based social support using relatively low-cost methods, such as brochures, may offer a key method to moving toward patient-centered models, particularly for an illness in which treatment adherence and perhaps even outcomes are substantially influenced by prevailing social norms and attitudes. Clinicians may wish to consider referral to such reliable and reputable Internet-based social media as an early step to enable patients to explore with others treatment options or barriers to adherence. Health administrators and health services researchers should consider streamlined approaches to engaging patients with reliable and reputable Internet-based supportive social media. The next logical step would be a randomized clinical trial in a primary care setting to determine if referral to Internet-based social support for those with depressive illness improves quality of care, symptoms, and function.

Acknowledgments

Supported by NIMH R-34 R34 MH 073742-01 and NIMH K-08 MH 072918-01A2 (Dr Van Voorhees salary support during study).

Conflicts of Interest

Dr Van Voorhees operates a limited liability company, RISE Consulting, that provides know-how support for health technology companies. Dr Van Voorhees has served as a consultant for Prevail Health Solutions, Inc, Mevident Inc, Verimed Inc, and Social Kinetics, Inc. Dr Van Voorhees has also been a paid speaker/academic visitor at the University of Hong Kong, Chinese International School (Hong Kong), and Wuhan University. The University of Chicago has granted a no-cost license to Mevident, Inc to adapt the CATCH-IT intervention. Dr Van Voorhees has agreed to support the company working 5.5 days as consultant at $1000/day. Dr Hsiung is the single member of Dr Bob LLC, which owns and operates the Psycho-Babble Internet support group. The net revenue of Dr Bob LLC is less than $1000/year.

Multimedia Appendix 1

Group assignment samples and extended analysis of the tables embedded in the body of the manuscript.

[PDF File (Adobe PDF File), 119KB - jmir_v15i3e42_app1.pdf]

Multimedia Appendix 2

CONSORT eHealth Checklist V1.6.2 [36].

[PDF File (Adobe PDF File), 1008KB - jmir_v15i3e42_app2.pdf]

References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
2. Vansteenkiste M, Sheldon KM. There's nothing more practical than a good theory: integrating motivational interviewing and self-determination theory. Br J Clin Psychol 2006 Mar;45(Pt 1):63-82. [doi: 10.1348/014466505X34192] [Medline: 16480567]
3. Kessler RC, Nelson CB, McGonagle KA, Liu J, Swartz M, Blazer DG. Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. Br J Psychiatry Suppl 1996 Jun(30):17-30. [Medline: 8864145]
4. Kessler RC, Zhao S, Katz SJ, Kouzis AC, Frank RG, Edlund M, et al. Past-year use of outpatient services for psychiatric problems in the National Comorbidity Survey. Am J Psychiatry 1999 Jan;156(1):115-123. [Medline: 9892306]
5. Wells K, Sherbourne C, Schoenbaum M, Ettner S, Duan N, Miranda J, et al. Five-year impact of quality improvement for depression: results of a group-level randomized controlled trial. Arch Gen Psychiatry 2004 Apr;61(4):378-386. [doi: 10.1001/archpsyc.61.4.378] [Medline: 15066896]
6. Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med 2006 Nov 27;166(21):2314-2321. [doi: 10.1001/archinte.166.21.2314] [Medline: 17130383]

7. Gilbody S, Bower P, Whitty P. Costs and consequences of enhanced primary care for depression: systematic review of randomised economic evaluations. Br J Psychiatry 2006 Oct;189:297-308 [FREE Full text] [doi: 10.1192/bjp.bp.105.016006] [Medline: 17012652]
8. Gensichen J, Beyer M, Muth C, Gerlach FM, Von Korff M, Ormel J. Case management to improve major depression in primary health care: a systematic review. Psychol Med 2006 Jan;36(1):7-14. [doi: 10.1017/S0033291705005568] [Medline: 16356292]
9. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
10. Santor DA, Poulin C, LeBlanc JC, Kusumakar V. Online health promotion, early identification of difficulties, and help seeking in young people. J Am Acad Child Adolesc Psychiatry 2007 Jan;46(1):50-59. [doi: 10.1097/01.chi.0000242247.45915.ee] [Medline: 17195729]
11. 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]
12. Landback J, Prochaska M, Ellis J, Dmochowska K, Kuwabara SA, Gladstone T, et al. From prototype to product: development of a primary care/internet based depression prevention intervention for adolescents (CATCH-IT). Community Ment Health J 2009 Oct;45(5):349-354. [doi: 10.1007/s10597-009-9226-3] [Medline: 19641992]
13. Whittaker R. MEMO--a mobile phone depression prevention intervention for adolescents: development process and postprogram findings on acceptability from a randomized controlled trial. J Med Internet Res 2012;14(1).
14. Ritterband LM, Ardalan K, Thorndike FP, Magee JC, Saylor DK, Cox DJ, et al. Real world use of an Internet intervention for pediatric encopresis. J Med Internet Res 2008;10(2):e16 [FREE Full text] [doi: 10.2196/jmir.1081] [Medline: 18653440]
15. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
16. 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]
17. Van Voorhees B, Watson N, Bridges JF, Fogel J, Galas J, Kramer C, et al. Development and pilot study of a marketing strategy for primary care/internet-based depression prevention intervention for adolescents (the CATCH-IT intervention). Prim Care Companion J Clin Psychiatry 2010;12(3) [FREE Full text] [doi: 10.4088/PCC.09m00791blu] [Medline: 20944776]
18. Van Voorhees BW, Mahoney N, Mazo R, Barrera AZ, Siemer CP, Gladstone TR, et al. Internet-based depression prevention over the life course: a call for behavioral vaccines. Psychiatr Clin North Am 2011 Mar;34(1):167-183 [FREE Full text] [doi: 10.1016/j.psc.2010.11.002] [Medline: 21333846]
19. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
20. Burns J, Morey C, Lagelée A, Mackenzie A, Nicholas J. Reach Out! Innovation in service delivery. Med J Aust 2007 Oct 1;187(7 Suppl):S31-S34. [Medline: 17908022]
21. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform 2006 Jan;75(1):8-28. [doi: 10.1016/j.ijmedinf.2005.07.032] [Medline: 16125453]
22. Griffiths KM, Calear AL, Banfield M, Tam A. Systematic review on Internet Support Groups (ISGs) and depression (2): What is known about depression ISGs? J Med Internet Res 2009;11(3):e41 [FREE Full text] [doi: 10.2196/jmir.1303] [Medline: 19793718]
23. 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]
24. Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. Health Technol Assess 2006 Sep;10(33):iii, xi-ixiv, 1 [FREE Full text] [Medline: 16959169]
25. Strecher VJ, Shiffman S, West R. Moderators and mediators of a web-based computer-tailored smoking cessation program among nicotine patch users. Nicotine Tob Res 2006 Dec;8 Suppl 1:S95-101. [Medline: 17491176]
26. 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]
27. 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]
28. Lin RJ, Zhu X. Leveraging social media for preventive care-A gamification system and insights. Stud Health Technol Inform 2012;180:838-842. [Medline: 22874310]

29. Ajzen I. The directive influence of attitudes on behavior. In: Gollwitzer PM, Bargh JA, editors. The Psychology of Action: Linking Cognition and Motivation to Behavior. New York: Guilford Press; 1996:385-403.
30. 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]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
32. Sturm R. Tracking changes in behavioral health services: how have carve-outs changed care? J Behav Health Serv Res 1999 Nov;26(4):360-371. [Medline: 10565097]
33. Sturm R. How expensive is unlimited mental health care coverage under managed care? JAMA 1997 Nov 12;278(18):1533-1537. [Medline: 9363977]
34. Shuman CR. Managed psychiatric care: a suburban medical department activity model. Mil Med 1996 Sep;161(9):557-561. [Medline: 8840798]
35. National Institute of Mental Health. URL: http://www.nimh.nih.gov/index.shtml [accessed 2013-02-11] [WebCite Cache ID 6EMPzizsY]
36. Australian National University. The MoodGYM Training Program URL: https://moodgym.anu.edu.au/welcome [accessed 2013-02-11] [WebCite Cache ID 6EMQ2CtRA]
37. Hsuing R. Psycho-Babble. URL: http://www.dr-bob.org/babble/ [accessed 2013-02-11] [WebCite Cache ID 6EMQ4HcHk]
38. Hsiung R. The best of both worlds: An online self-help group hosted by a mental health professional. CyberPsychology & Behavior 2000 Dec 2000;3(6):935-950. [doi: 10.1089/109493100452200]
39. Crutzen R, Cyr D, de Vries NK. The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. J Med Internet Res 2012;14(2):e45 [FREE Full text] [doi: 10.2196/jmir.1922] [Medline: 22532074]
40. 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 [FREE Full text] [doi: 10.2196/jmir.1837] [Medline: 21946128]
41. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measures 1977;1:385-401.
42. Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT. The stress process. J Health Soc Behav 1981 Dec;22(4):337-356. [Medline: 7320473]
43. Myers RE, Wolf TA, McKee L, McGrory G, Burgh DY, Nelson G, et al. Factors associated with intention to undergo annual prostate cancer screening among African American men in Philadelphia. Cancer 1996 Aug 1;78(3):471-479. [doi: 10.1002/(SICI)1097-0142(19960801)78:3<471::AID-CNCR14>3.0.CO;2-W] [Medline: 8697393]
44. Van Voorhees BW, Fogel J, Houston TK, Cooper LA, Wang NY, Ford DE. Attitudes and illness factors associated with low perceived need for depression treatment among young adults. Soc Psychiatry Psychiatr Epidemiol 2006 Sep;41(9):746-754. [doi: 10.1007/s00127-006-0091-x] [Medline: 16896516]
45. Marko M, Fogel J, Mykerezi E, Van Voorhees BW. Adolescent Internet depression prevention: preferences for intervention and predictors of intentions and adherence. Journal of CyberTherapy and Rehabilitation 2010;3(1):9-30.
46. Miller W, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behaviors. New York: Guilford Press; 1991.
47. Clever SL, Ford DE, Rubenstein LV, Rost KM, Meredith LS, Sherbourne CD, et al. Primary care patients' involvement in decision-making is associated with improvement in depression. Med Care 2006 May;44(5):398-405. [doi: 10.1097/01.mlr.0000208117.15531.da] [Medline: 16641657]
48. Cohen J. Statistical Power Analysis for the Behavioral Sciences 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
49. Ritterband LM, Borowitz S, Cox DJ, Kovatchev B, Walker LS, Lucas V, et al. Using the internet to provide information prescriptions. Pediatrics 2005 Nov;116(5):e643-e647 [FREE Full text] [doi: 10.1542/peds.2005-0404] [Medline: 16263978]
50. Van Voorhees BW, Fogel J, Pomper BE, Marko M, Reid N, Watson N, et al. Adolescent Dose and Ratings of an Internet-Based Depression Prevention Program: A Randomized Trial of Primary Care Physician Brief Advice versus a Motivational Interview. J Cogn Behav Psychother 2009;9(1):1-19 [FREE Full text] [Medline: 20694059]
51. Resnicow K, Davis RE, Zhang G, Konkel J, Strecher VJ, Shaikh AR, et al. Tailoring a fruit and vegetable intervention on novel motivational constructs: results of a randomized study. Ann Behav Med 2008 Apr;35(2):159-169. [doi: 10.1007/s12160-008-9028-9] [Medline: 18401673]
52. Hoek W, Marko M, Fogel J, Schuurmans J, Gladstone T, Bradford N, et al. Randomized controlled trial of primary care physician motivational interviewing versus brief advice to engage adolescents with an Internet-based depression prevention intervention: 6-month outcomes and predictors of improvement. Transl Res 2011 Dec;158(6):315-325. [doi: 10.1016/j.trsl.2011.07.006] [Medline: 22061038]
53. 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]
54. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]

55. Veroff J, Kulka RA, Douvan EAM. Mental Health in America: Patterns of Help Seeking from 1957-1976. New York: Basic Books; 1981.
56. Van Voorhees BW, Fogel J, Houston TK, Cooper LA, Wang NY, Ford DE. Beliefs and attitudes associated with the intention to not accept the diagnosis of depression among young adults. Ann Fam Med 2005;3(1):38-46 [FREE Full text] [doi: 10.1370/afm.273] [Medline: 15671189]
57. Nutting P, Rost K, Dickinson M, Werner JJ, Dickinson P, Smith JL, et al. Barriers to initiating depression treatment in primary care practice. J Gen Intern Med 2002 Feb;17(2):103-111 [FREE Full text] [Medline: 11841525]
58. Schneider F, van Osch L, de Vries H. Identifying factors for optimal development of health-related websites: a delphi study among experts and potential future users. J Med Internet Res 2012;14(1):e18 [FREE Full text] [doi: 10.2196/jmir.1863] [Medline: 22357411]
59. Brouwer W, Oenema A, Crutzen R, de Nooijer J, de Vries NK, Brug J. An exploration of factors related to dissemination of and exposure to internet-delivered behavior change interventions aimed at adults: a Delphi study approach. J Med Internet Res 2008;10(2):e10 [FREE Full text] [doi: 10.2196/jmir.956] [Medline: 18417443]

Abbreviations

ANOVA: analysis of variance
CES-D: Center for Epidemiologic Studies Depression Scale
ES: effect size
ISG: Internet support group
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 01.06.12; peer-reviewed by C Barker, J Barnes; comments to author 25.08.12; revised version received 26.10.12; accepted 19.12.12; published 12.03.13.

Please cite as:
Van Voorhees BW, Hsiung RC, Marko-Holguin M, Houston TK, Fogel J, Lee R, Ford DE
Internal Versus External Motivation in Referral of Primary Care Patients with Depression to an Internet Support Group: Randomized Controlled Trial
J Med Internet Res 2013;15(3):e42
URL: http://www.jmir.org/2013/3/e42/
doi:10.2196/jmir.2197
PMID:23482332

Original Paper

Online Cognitive Behavioral Therapy for Bulimic Type Disorders, Delivered in the Community by a Nonclinician: Qualitative Study

Carrie-Anne McClay[1], BA(Hons); Louise Waters[1], PhD; Ciaran McHale[1], BA(Hons); Ulrike Schmidt[2], MD, PhD; Christopher Williams[1], MB ChB, BSC, MMedSc, MD, FRCPsych

[1]Institute of Health and Wellbeing, Mental Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

[2]Institute of Psychiatry, King's College London, London, United Kingdom

Corresponding Author:
Christopher Williams, MB ChB, BSC, MMedSc, MD, FRCPsych
Institute of Health and Wellbeing
Mental Health and Wellbeing
University of Glasgow
Administration Building, Gartnavel Royal Hospital
1055 Great Western Road
Glasgow, G12 0XH
United Kingdom
Phone: 44 0044 141 211 ext 39
Fax: 44 0044 141 211 38
Email: chris.williams@glasgow.ac.uk

Abstract

Background: Cognitive behavioral therapy is recommended in the National Institute for Clinical Excellence guidelines for the treatment of bulimia nervosa. In order to make this treatment option more accessible to patients, interactive online CBT programs have been developed that can be used in the user's own home, in privacy, and at their convenience. Studies investigating online CBT for bulimic type eating disorders have provided promising results and indicate that, with regular support from a clinician or trained support worker, online CBT can be effective in reducing bulimic symptoms. Two main factors distinguish this study from previous research in this area. First, the current study recruited a wide range of adults with bulimic type symptoms from the community. Second, the participants in the current study had used cCBT with support from a nonclinical support worker rather than a specialist eating disorder clinician.

Objective: To investigate participants' experiences of using an online self-help cognitive behavioral therapy (CBT) package (Overcoming Bulimia Online) for bulimia nervosa (BN) and eating disorders not otherwise specified (EDNOS).

Methods: Eight participants with a mean age of 33.9 years took part in semi-structured interviews. Interviews were transcribed and analyzed using a 6-step thematic analysis process.

Results: Saturation was achieved, and 7 themes were identified in the dataset. These were: (1) conceptualizing eating disorders, (2) help-seeking behavior, (3) aspects of the intervention, (4) motivation to use the online package, (5) privacy and secrecy with regard to their eating problems, (6) recovery and the future, and (7) participant engagement describing individuals' thoughts on taking part in the online research study.

Conclusions: Participants suggested that online CBT self-help represented a generally desirable and acceptable treatment option for those with bulimic type eating problems, despite some difficulties with motivation and implementation of some elements of the package.

Trial Registration: International Standard Randomized Controlled Trial Number of the original RCT that this study is based on: ISRCTN41034162; http://www.controlled-trials.com/ISRCTN41034162 (Archived by WebCite at http://www.webcitation.org/6Ey9sBWTV)

(J Med Internet Res 2013;15(3):e46) doi:10.2196/jmir.2083

KEYWORDS

bulimia nervosa; self-help; cCBT; qualitative research; cognitive behavioral therapy

Introduction

Bulimia nervosa (BN) is an eating disorder characterized by recurrent episodes of binge eating and reversing behaviors and has many negative psychological, physical, and social consequences. Cognitive behavioral therapy (CBT) is recommended in the National Institute for Clinical Excellence [1] guidelines for the treatment of bulimia nervosa. In order to make CBT more accessible to individuals who are unable or unwilling to engage in face-to-face therapy, computerized CBT (cCBT) packages have been developed for the treatment of BN. Computerized CBT has been shown to result in improvements in bulimic symptoms when delivered in CD-ROM format [2,3,4] or online [5,6] and can be used in the individual's own home, in privacy, and at their convenience.

To date, two studies of an online CBT package, "Overcoming Bulimia Online" (OBO) [7] have included a qualitative investigation of users' attitudes towards and experience of using online CBT. In Pretorius, Rowlands, Ringwood, and Schmidt's [8] study of adolescents' experiences of using OBO, the main themes identified included: motivation, support, factors influencing choice of treatment and recovery, and the future. Participants reported choosing online CBT because of its accessibility and flexibility, citing a desire to have autonomy and control over their eating disorder and recovery. Many participants indicated that they considered cCBT as a first step in their treatment and perhaps as a first step towards additional treatment options due to increased confidence in discussing their eating problems, but they reported that low motivation was a barrier to using this intervention.

Sanchez-Ortiz et al [9] more recently carried out a qualitative study of students' experiences of OBO. Interviewees cited flexibility, confidentiality, and ease of use as positive aspects of the online package. Interestingly, many participants discussed other treatments they had accessed and compared these with their experience of OBO. Participants indicated that the online CBT package was more structured than counseling. Also, some participants expressed feelings of apprehension with regard to approaching their GP about their eating problems. This may be one of the reasons that some individuals see cCBT as a more desirable treatment option. However, that study and the randomized controlled trial (RCT) from which participants were drawn focused on students [6]. Students are a highly selected group educationally, in terms of age, their access via universities to high-speed Internet, and their information technology literacy.

The current study aimed to add to this growing body of literature regarding a wider population of user's experiences of using a cCBT package for bulimia nervosa. Two main factors distinguish this study from previous research in this area. First, the current study recruited a wide range of adults with bulimic type symptoms from the community. This meant that a more representative sample of the population could be recruited. The two studies previously mentioned recruited individuals who were highly likely to favor this online treatment approach: students and adolescents. Second, in contrast to the studies of Pretorius et al [8] and Sanchez-Ortiz et al [9], all participants in the current study had used OBO with support from a nonclinical support worker rather than specialist eating disorder clinicians. This is significant since, having similarly favorable results, such as when the support is provided by specialists, would support the view that low intensity CBT can be delivered and supported by trained but nonspecialist support staff. This would save resources in services and aid the increase in access to psychological therapies, a main treatment target in the National Health Service.

Therefore, the purpose of this qualitative study was to investigate whether similar themes emerged with this new population. The study also aimed to investigate reasons for choosing the self-help approach and to compare attitudes towards cCBT with other treatments that participants had accessed in the past. A final aim was to determine participants' attitudes towards taking part in a predominantly online research study where recruitment, questionnaires, and follow-up assessment were carried out using online web pages and emailed links to online evaluations.

Methods

Recruitment

All participants invited for interview were taking part in an RCT of Overcoming Bulimia Online. These participants had been recruited to take part in an initial survey about attitudes towards online self-help for eating disorders [10] and were then screened and invited to take part in the RCT. Various community-based recruitment methods were employed in the study. Advertisements were placed on mental health websites, in mental health newsletters, and an article about the study appeared in an eating disorders magazine (*Anorexia and Bulimia Care*). Posters were also placed in public places such as libraries, gyms, universities, colleges, and student counseling services. Additionally, flyers giving details of the study were put into university students' orientation packages.

Participants

Ten months into the RCT of OBO, 55 participants who had had access to the online package for at least 10 weeks were grouped according to how many sessions of the package they had completed—defined as low users (0-2 sessions), medium users (3-5 sessions), and high users (6-8 sessions) (the OBO package contains 8 sessions in total). A purposive sampling approach was taken with 12 participants from each group who were randomly selected to be invited for interviews using the randomization function in SPSS 12.

Nine participants agreed to take part in the interviews. One interview could not be transcribed due to a poor recording, so 8 interviews were used in the final analysis. As expected, the number of sessions completed had increased by the time of interview as participants still had access to the online package, therefore the sample contained no low users (0-2 sessions), 2 medium users (3-5 sessions), and 6 high users (6-8 sessions). The mean number of sessions completed was 6.4 out of the 8 session online course. All participants were women; the mean age of the sample was 33.9 years (range = 28-50 yrs). All individuals met the criteria for Bulimia Nervosa or Eating Disorder Not Otherwise Specified (EDNOS), as established

prior to entry into the RCT using the Eating Disorders Examination [11]. Duration of their eating disorders ranged from 2 years to 30 years; the mean was 16.6 (SD 8.6).

Procedure

The Online Intervention: "Overcoming Bulimia Online"

Participants interviewed had been given access to the online CBT based self-help package OBO as part of an RCT. The package contained 8 interactive sessions that aimed to change users' thoughts, feelings, and behavior with regard to food. As illustrated in Figure 1, the sessions contained written text and video clips, which were supplemented by audio information and instructions. The 8 topics covered in the package were: (1) Introduction and What is Bulimia?, (2) Understanding why I have Bulimia, (3) How do I change?, (4) The role of thoughts in Bulimia, (5) Assertiveness and increasing activity, (6) Problem solving, (7) Living Life to the Full, and (8) Planning for the future and review of what I have learned.

Participants were encouraged to work through the sessions independently in their own time and chose to receive weekly telephone, email, or text message support. These 8 support sessions were provided along with the online package by a trained self-help support worker. The support worker was a research assistant with experience of working in the area of eating disorders and guided CBT research. She had written guidelines on how to deliver guided self-help and was supervised by the clinical lead of the project (CW). At the beginning of the intervention, 3 of the participants chose to receive their support sessions over the telephone, 4 requested email support, and 2 opted for text message support to their mobile phones. The support time varied depending on the extent to which the participant engaged with the support, ie, replied to the support emails.

At the end of each module, participants were asked to complete various homework assignments such as using the anxiety control training audio daily until the next session, implementing some of the rules of healthy eating, and keeping food, thought, and activity diaries. Additionally, participants were asked to rate their mood and give information about their bingeing and vomiting at the beginning of each session. This information was recorded and presented in a graph that could be viewed by the participant and any authorized support workers, researchers, or clinicians (see Figure 2).

Figure 1. Screenshot from OBO, Session 3: "How do I change?"

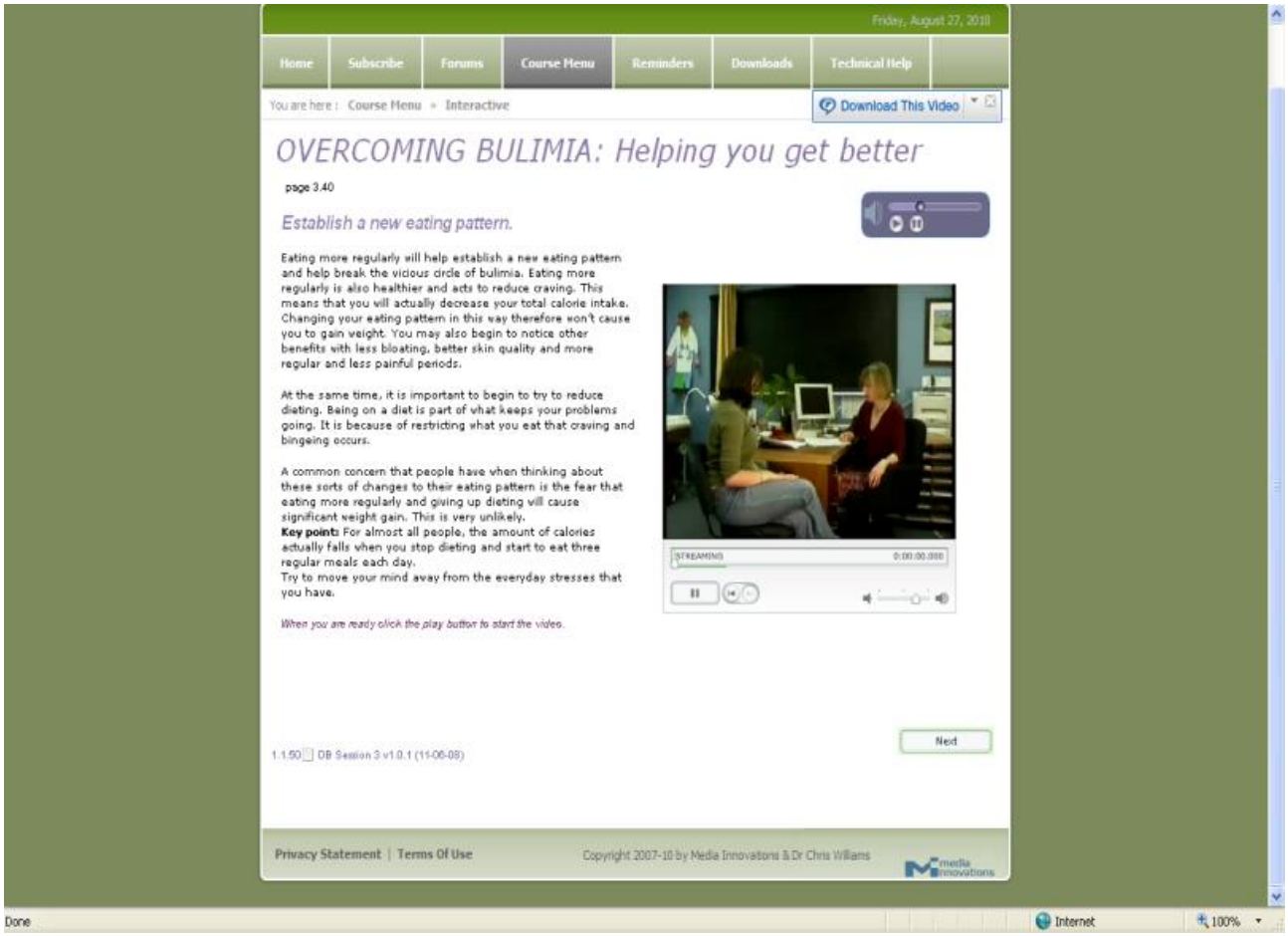
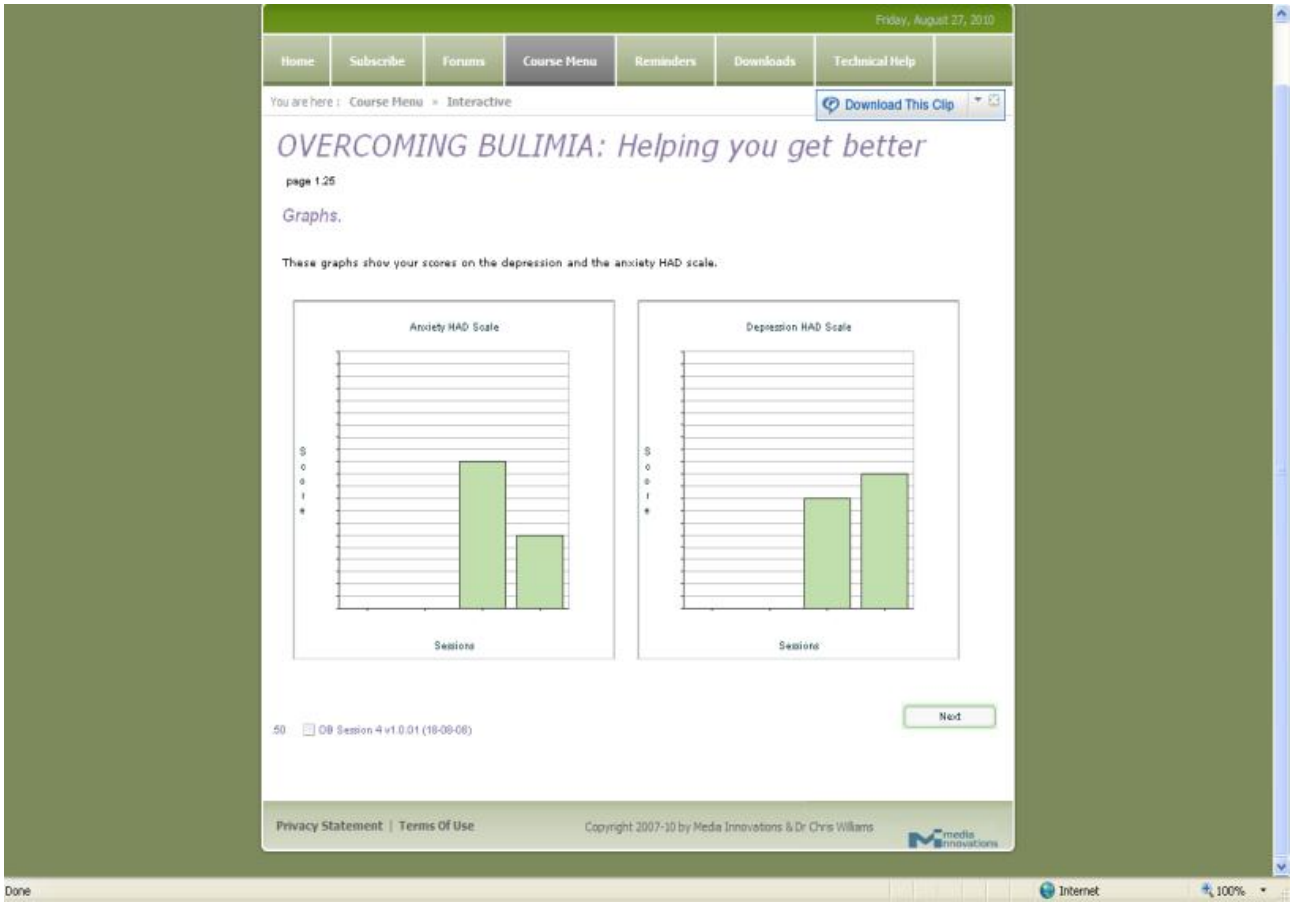


Figure 2. Screenshot of the mood rating graphs in Overcoming Bulimia Online.



The Interviews

Following informed consent, semi-structured interviews were carried out by C-A M over the telephone and were recorded using a digital recording device. A question guide was devised by the research team to probe issues such as participants' reasons for choosing the self-help approach, their experience of using the online package, motivational issues while using the package, how the intervention compared to other treatments they had accessed in the past, and aspects of the online research process. The mean duration of the interview was 19.4 minutes.

Analysis

Interviews were recorded, transcribed verbatim, and analyzed using the 6-step inductive thematic analysis approach outlined by Braun and Clark [12]. This method of analysis was chosen because the identified themes are data driven. As such, themes are identified without trying to fit them into a pre-designed coding framework.

Two researchers (C-A M and CM) carried out the coding and identification of themes limiting the subjective nature of the analysis, thus increasing the validity. Both researchers independently generated initial codes for all interesting extracts. Using these codes, themes were then searched for. Once this stage was complete, the analyzers discussed the themes that they had identified and agreed upon the final themes that would be used. All coded extracts were then put into the appropriate themes and subthemes. In some cases, extracts were themed differently by the researchers; C-A M discussed these with a third member of the research team (LE) who was experienced in the process of qualitative analysis and then made the final decision on where these extracts belonged. All themes were reviewed and the extent to which they encapsulated the data was judged.

Results

Analysis of transcripts identified seven main themes, as illustrated in Table 1.

Table 1. Themes and subthemes identified in the semi-structured interviews.

Themes	Subthemes
Conceptualizing eating disorders	Impact of and feelings about bulimia
	Perceptions of eating disorders/people with eating disorders
	Acknowledgement of/acceptance of the problem
Help-seeking	Past experiences
	Barriers to help
	Reasons for choosing self-help
	Prior knowledge
The intervention	Support worker
	Positive aspects
	Negative aspects/difficulties
Motivation	Aids
	Challenges
Participant engagement	Opportunity to help self and others
	Experience as an online research participant
Progress and recovery	Improvements in bulimic symptoms
Privacy	Secrecy
	Talking about bulimia

Conceptualizing Eating Disorders

Participants stated that their eating problems had resulted in many negative consequences including low self-esteem, poor self-concept, a negative view of the future, and concerns about the impact of the disorder on their family. Participants also spoke of feelings of isolation and shame associated with BN, and there was a sense that their eating problems were a significant burden in their lives: “I thought I was just a horrible, horrible person and all I deserved was to be who I am”. [P9]

Several statements were made relating to perceptions of people with eating disorders. There was a sense that views such as these were damaging and affected confidence in help-seeking and talking about eating problems: “...there’s a standard joke that people who are bulimic are failed anorexics...and that’s an extra thing that you carry around with you”. [P1]

Although participants had very negative feelings about their eating problems, some interviewees expressed relief in having someone acknowledge their eating problems: “I could take a deep breath and think oh my god (laughs) this is real. Because of all those years of people saying to me oh you’re just jumping on the band wagon and all the rest of it”. [P1]

The statements by these participants seemed to indicate a turning point and a sense that now their problems had been acknowledged by another person (the researchers), that they then could accept them and start to tackle these problems.

Help-seeking

All participants discussed their past help-seeking and their experience of various treatment and support options including: their GP, formal therapy, self-help books, and support groups. Negative experiences described included: difficulty in understanding self-help books, unsatisfactory support services, and the feeling that their problems were dismissed or inadequately dealt with: “...that I’ve tried over the years is talking to my GP, various GPs about it and just been fobbed off so... Yeah I’ve been turned, turned away by I think it’s 4 GPs who’ve all said it’s not a problem”. [P1]

The most common barrier identified by the majority of participants was the lack of access to treatment through the NHS. There was a general sense of low expectations of the NHS and a portrayal that these individuals felt that they had to arrange help for themselves: “And also because there’s not much help out there for like, you know if you go to your doctors there’s a long waiting list, that’s if you get help, so it’s just wanting to recover and get whatever help is available...”. [P2]

Practical issues and readiness for change were also identified as barriers to accessing help in the past. Travelling to an eating disorders center or therapy session, taking time off work, and thus having to disclose the eating disorder was identified as a problem. Individuals voiced being reluctant to seek help or had chosen to manage their problems alone, often due to the common feelings of shame and embarrassment known to be associated with the disorder: “I’ve tried, er, self-help books, you know looked at books in the library...Em, I never really wanted to take them out because of the embarrassment of asking for them over the counter”. [P1]

The issue of prior knowledge about eating disorders and their treatments was raised by a few interviewees. The extent to which participants understood eating disorders varied with one participant expressing confusion concerning the causes of her BN and another stating that despite having a sound knowledge

of the problem, she was unable to use this information to help herself: "I used to work with adolescents and the adolescents either had eating disorders or were self-harmers so I had a lot of knowledge that I never applied to myself". [P1]

Participants outlined various reasons for choosing self-help. There was a sense that the self-help approach seemed like a desirable alternative to traditional face-to-face treatment, primarily because of the private, convenient, flexible, and anonymous nature of the treatment. Autonomy and the ability to deal with these problems independently were also highlighted as attractive aspects of self-help materials: "you didn't have to sit there with a patronising (pause) um person, being judged every week on whether you follow what she actually said to you or not". [P7]

Two participants also stated that feelings of desperation led them to seek help on the Internet, an indication again that perhaps other routes of accessing support for their eating problems had been unhelpful or undesirable. The following statement highlights the serious situation that these undiagnosed and untreated individuals are in and the importance of community initiatives in order to identify and support such individuals: "I was really, really desperate...when people are looking through Internet websites for online help...it's the question between living and dying really and that's where I was". [P9]

The Intervention

Interviewees were largely positive about the online package. Many participants stated that OBO was helpful because it increased their understanding of their problems. Participants also said that the intervention was very specific and gave them useful skills and tools that they could use to improve their situation. There was a general indication that the package was educational and easy to understand: "It helps you tackle other stuff as well, and being assertive, it sort of just really—it teaches you the whole lot of it, if you get what I mean. It's just not one thing, it's lots of little things I think now that cause it". [P8]

Participants expressed that a particularly useful element of the package was the "Letter from the future" and also the "challenging unhelpful thoughts" session with the thought flashcards: "I think the most helpful idea was the idea of the letters from the future...actually just thinking about it, in 10 years time do I actually still want to be in the same position and thinking about the things that I might be missing out on". [P6]

Again, participants highlighted the notion of privacy and autonomy as being a positive aspect of such a treatment. The fact that the treatment could be accessed without others knowing about it was cited as a main advantage of OBO. The convenient and quickly accessible nature of the package was also outlined as an advantage: "...it was good because it was convenient so uh huh you could fit it into, you know, the rest of your life". [P5]

Participants also mentioned a number of difficulties or challenges that they faced when using the self-help package including technical difficulties, a lack of support, and motivation. A key difficulty identified by participants was the fact that the package was not tailored enough to their needs. Participants described the fact that some content was not new to them, the package was designed to meet the needs of a group of people with BN rather than the individual, and one individual got the impression that the package was aimed at younger users: "...if you click back by accident you end up losing everything so that's a bit rubbish" [P7] and "there was probably was a need for a personal element but em, a bit more tailored like em, like that fitted somebody who was 20 years down the line". [P4]

However, it was acknowledged that this difficulty was perhaps an inevitable aspect of using an online self-help package. Participants also expressed difficulties in implementing aspects of the package such as a healthy eating regime, using the anxiety control training, and challenging unhelpful thoughts: "I did start to kind of engage with a healthier eating pattern but I immediately started to put on weight so that was, that really put me off". [P6]

Other difficulties include the lack of discretion of the web pages of OBO with participants saying they could not use the package unless they were alone because the topic of the package was very obvious: "...the package itself when you go online it, it was quite clearly that it is to do with bulimia, bulimia so I couldn't use it a lot of the time". [P4]

This links closely to another difficulty experienced by some participants: lack of privacy to use the package. Two participants said that they found it difficult to get time alone to use the online package without people such as family members interrupting them.

Interviewees referred to the workload involved in completing the package and the pace at which they completed the sessions. Many said that they were unable to complete the sessions at the pace suggested, one session per week ,and although the majority of participants referred to the workbooks, they were often not actively used: "I think it's quite a bit to be involved in the package...found er um, almost impossible to do one a week" [P6] and "the thought of having to fill more books in and logs and things when you do that sort of thing all day it just, it's just a complete turn off". [P4]

Participants stated that the support worker was a valued element of the intervention and that it was good to have someone there to help with their progress through the package without judgment: "...she has made a huge huge difference so I appreciate all the chats I've had with her and she's good...yeah I think it's made, it made a massive massive difference". [P7]

Additionally, email support was considered an effective means of receiving the support. Participants also referred to the flexibility of the support. One stated that it was helpful to change their support medium when required. Others said that they wish they had taken advantage of this flexibility, perhaps indicating that this option should have been promoted more: "I think that it might have been better for me for sometimes to actually email and say I want to talk over the phone about this". [P5]

However, some participants outlined the fact that the weekly contact from the support worker made them feel pressured to complete the sessions, and one participant indicated that she did not engage in the support sessions because she felt guilty when she had not made any progress with the package: "I didn't

find it supportive because it asked about 5 or 6, if not more questions of which I couldn't answer any of them because I'd only just literally been given access to the online package". [P4]

Another participant said that she did not feel that the support helped her progress, although she did appreciate the texts from the support worker: "Em, that, that didn't really help me either way...I think (the support worker) just used to text me everyone now and then to see if I was ok. Which was nice". [P8]

Some participants suggested that it would be helpful to have continued support after the 8-10 week intervention period. They said that it would beneficial to have follow-up sessions with their support worker in the months following the intervention to discuss how things are going and perhaps keep them on track with the improvements they had made: "You know like not completely cut off the contact but...Yeah. But stay in touch even for a longer period of time". [P9]

Motivation

A general feeling of low motivation was portrayed by some participants, indicating that this was a major obstacle to their progression with the intervention: "Well, er actually, it was easier to just not really do it". [P7]

Two individuals said they had to work on getting the motivation to start the package: "...it's been quite hard to sort of em, I suppose, motivate myself... because I always thought that once I got treatment, it would, you know, getting the treatment would be the difficult thing and then once I started, you know, it would be easy". [P5]

A lack of support was also cited as something that affected participants' motivation in using the package. Some individuals stated that they would have benefited from more human interaction.

Interviewees cited a wide variety of things that aided them in maintaining motivation while using OBO. The support worker was the most commonly cited aid. Participants said that the support worker helped to reassure them about various issues relating to the package and helped them move past difficulties and continue through the package: "I think if it hadn't been for (the support worker) at that point. I would have thought that it's not it er is obviously not working, it's not worth persevering with". [P1]

Knowing that there was someone there who knew that they were using the package and asking how things were going helped some individuals to maintain their motivation: "you've still got that person in the background who's wanting you know, wanting to know what's happening so that's motivation". [P1]

Other aids highlighted by interviewees included: telling family members about their use of the package, the study newsletter, completing the follow-up questionnaires, thinking about the "Letter from the future" and actually completing the sessions. Planning and goal setting was identified as important by some individuals, with some participants expressing difficulties in completing the sessions, acknowledging that planning their next session was really needed to make any progress: "...you've got to just try and be determined haven't you, set yourself the goals to do it. (hesitation) mmm you know work towards achieving those goals". [P2]

Finally, some participants were motivated to use the package because they saw it as an opportunity for change and to address their eating problems: "I just thought, if I don't take this opportunity and sort of, em, give it my best shot, mm huh then I was going to go back to how I was". [P5]

Participant Engagement

This theme relates to participant's experiences of taking part in the online research study. Many participants saw taking part in the study as a significant opportunity to gain access to treatment with many participants portraying a significant desire for change in their lives: "I thought er this is great, an opportunity you know to take part in the study, to do something that I er I wouldn't normally be able to access". [P1]

Participants also expressed the fact that they wanted to make a contribution to the research and aid the development of treatment for the future, to help others with similar problems: "...you're kind of involved in a research study which means that you're kind of doing something which might benefit other people as well". [P5]

The online nature of the research process was generally acceptable to participants with many participants saying that the website was easy to use and that completing questionnaires online was convenient and economical: "I think in this day and age online is a good thing". [P7]

Privacy and Secrecy

The concepts of privacy and secrecy were pertinent throughout the dataset and related to many of the themes. As outlined, the desire for privacy with regard to their eating problems affected help-seeking behavior and the preference for the self-help approach. The private nature of OBO was cited as a significant advantage of this intervention. However, the desire for their eating problems to be kept hidden meant that some individuals had limited external support when their support sessions ceased with some saying that they had no one else to talk to about their problems. This had an impact on their motivation to address their problems. Additionally, as previously described, maintaining secrecy with regard to their eating problems had an impact on participants' use of the online sessions and accompanying materials because they did not want others to see them: "I don't actually have anybody else, there's nobody else in my life day to day that I've told about this". [P6]

This desire for privacy and secrecy was acknowledged as damaging by one participant. This participant said that having used the intervention, she is now more open to talking about her eating problems: "And there's something about you know it being hidden and awful that it's not helpful isn't it...It just feeds itself". [P1]

Progress and Recovery

Many participants said that they had made noticed improvements in their bulimic symptoms and behavior with regard to food and eating: "I'm nearly recovered now em...I think it was sometime

in June that I last made myself sick and even that was a one off". [P8]

Others had identified more subtle differences in their behavior by applying some of the skills they had learned from the package. For example, one participant found the Rules of Healthy Eating particularly helpful: "...actually saying, 'that's your meal over' kind of thing yeah and I've done that quite a few times without really thinking of it". [P9]

However, there was some indication that the package perhaps could not completely solve eating problems with some suggesting that more time and support are required to address these problems. Related to this point, some participants mentioned other helpful forms of support that they had accessed for their mental health problems including high intensity specialist psychotherapeutic help such as a therapist, Cognitive Analytic Therapy, and a social worker. "..it's not just something that you can get over in 8 weeks". [P7]

Finally, one participant said that the package helped her to gain a more internal locus of control with regard to her eating disorder and her recovery and made her take a more autonomous approach to addressing her problems: "I realised throughout my recovery that there is no one else that can help me but myself that was just, just an eye opening moment..." [P9]

Discussion

Summary and Discussion of Key Findings

This qualitative study is a significant contribution to the growing body of literature regarding participants' experiences of using the online self-help package "Overcoming Bulimia Online" for bulimia nervosa. These participants were recruited from the wider community rather than a specialist eating disorders setting or student setting, and all participants were supported by a nonspecialist for the first time in a BN research study of this kind. This method has previously been successful in a study of cCBT for depression [13]. Participants expressed a range of attitudes towards help-seeking, and in particular, their negative past experiences support statements made in previous studies. As in Sanchez-Ortiz et al's [9] study, participants indicated that many previous sources of support had been unsuccessful including: counseling, self-help books, CBT, and GP visits. There was a general sense of negativity regarding accessing traditional treatments, an awareness of long waiting lists, and one participant reported being turned away or "fobbed off" by a GP. This experience supports the statistics from a *beat* (UK eating disorders voluntary sector organisation) survey [14] in which only 15% of respondents felt that their eating problems were understood by their GP. Thus taken together the findings highlight the fact that perhaps there needs to be more information and training available to primary care workers.

This lack of faith in traditional treatments and participants' negative past experiences influenced the decision to try the online guided CBT self-help approach. Additionally, the private, accessible, and convenient nature of online CBT was attractive to interviewees. These reasons for choosing self-help have been cited in previous studies of computerized CBT [8,15,16] and seem to represent a key attraction of the self-help approach for those who do not wish to pursue traditional routes of treatment.

Similar to the findings of previous studies of the OBO package [8,9], the support that participants received from their support worker while using OBO was commented on frequently and in the present study was seen as a core element of the intervention for the majority of participants, particularly in relation to maintaining motivation. This is important as the support was not provided by a clinician, as in previous studies, but was still considered as highly helpful and desirable by the majority of participants. This highlights the fact that support can be effectively delivered by trained nonspecialist health care workers, overcoming the problem of the lack of CBT therapists in many areas. This supports the view that low intensity CBT can be delivered by individuals who have had self-help training rather than specialist CBT training, therefore increasing access to psychological therapies and reducing the cost of psychological interventions for bulimia nervosa.

Another significant finding was that some participants expressed the desire for follow-up support sessions following the block of 8 support sessions they received while working through the package. There was a sense that some participants felt abandoned when the support ceased, and follow-up sessions may be needed to help keep participants on track with their progress, particularly for those with little or no social support. This suggestion for how to improve the support provided is valuable and could be easily implemented in future delivery of the intervention.

Motivation was widely cited as a problem in using the online package with many individuals saying that it was often difficult to maintain drive and determination to engage with the package. The support worker, goal setting, and seeing themselves making progress were all cited as increasing motivation. However, factors such as a lack of progress with the sessions, the feeling that the intervention was not working, and the fear of gaining weight all hindered motivation to engage with the intervention. The latter factor is significant as it is something that can be tackled within self-help packages. It is possible that an additional session early on in the package addressing body image and weight concern may specifically help to alleviate such concerns and may prevent such worries from interfering with the individual's engagement with the later sessions and their implementation. The site was overall considered to be user-friendly, despite some reported minor technical difficulties with navigation.

The findings showed that the desire for privacy and secrecy had a detrimental impact on participants' help-seeking, their use of the intervention, and the support they could seek from family and friends. This finding highlights the often cited problems faced by individuals with eating disorders such as shame, embarrassment, and the attached stigma. This may be a reason that individuals with eating problems often do not seek the help they need. This shame and embarrassment may also mean that individuals are less willing to seek support from significant others such as carers or health workers when they do take the step to seek help and engage with an intervention. This supports the need for community-based initiatives in order to identify

and engage such individuals who do not wish to access help through traditional channels. Future research should examine these feelings of shame and investigate how they can be overcome to allow individuals to seek the help they need, whether it is through traditional (NHS) or nontraditional routes including direct self-referral or interventions offered via the voluntary sector such as *beat*.

Overall participants expressed positive attitudes towards taking part in the study and saw it as a valuable opportunity to not only help themselves but also to contribute to the research and possibly help others with similar problems. As in the two previous qualitative studies of OBO [8,9], some participants outlined the fact that they had noticed improvements in their eating problems. This supports the use of online CBT as a first step in the treatment of even chronic bulimic type eating problems, as recommended in the NHS National Institute for Health and Clinical Excellence guidelines [1].

The online nature of the research project was popular with many participants saying that they felt it was easy to participate and that completing the questionnaires was convenient and efficient. This suggests that this online format could be an efficient, effective, and acceptable method of recruitment, screening, and data collection, especially in a group of individuals who perhaps want the entire research process and intervention to remain remote and anonymous.

Limitations

The study had some limitations. First, the sample size was relatively small, and the randomization for the interviews took place before the end of recruitment meaning those who entered later could not be interviewed. However, as it is a qualitative study in which a small sample size is expected, the important issue is one of saturation rather than generalizability, and saturation was achieved. Also, the sample consisted of purely community-based users who were actively seeking help and chose to take part in a study on cCBT and therefore may have more positive attitudes towards the approach than those seeking specialist therapies. This may affect the relevance of the findings in a clinical setting. The aim of the study was, however, to investigate experiences of community-based users and their engagement with an online self-help package with remote support. This delivery model may serve as a valuable support option for individuals in community and voluntary sector setting. Future research should assess uptake and efficacy of cCBT in clinical practice in order to gain a more complete knowledge of the potential of cCBT as a treatment for BN.

Despite these limitations, this study adds to the growing body of research in this area, first, because the participants had been recruited directly from the community through various avenues. The sample had a wider age range than previous qualitative studies, and the mean age was also higher in this study than previous studies, providing the opinions of an older group of user with the majority having chronic eating problems. Additionally, community-based recruitment meant that a variety of individuals entered the study, which means that the results may be more relevant to the general population than in previous studies of young people and students. The study provided further qualitative data from individuals who may not normally access treatment due to the lack of services available to them or the desire for privacy in relation to their eating disorders.

Conclusions and Implications

This study provides valuable information regarding participants' views of taking part in an online research study and may influence the implementation of future community based projects, within the area of mental health. The information regarding the factors that aid and hinder participants' motivation in using online self-help could also be used in the development of future online self-help strategies in order to maximize adherence and health and mental health-related outcomes. Finally, the study was the first of its kind to include individuals who had been involved in a guided self-help intervention for BN that used solely nonspecialist support workers and has established the acceptability of this type of support.

In the NHS, this intervention could be tested and, if successful, applied by staff members including assistant psychologists, nursing staff, or other members of staff trained in the delivery of self-help materials. This approach would save resources, reduce waiting lists, and therefore increase access to psychological therapies to those who are suitable for and willing to use cCBT resources. The guided self-help model therefore has the potential to improve productivity and facilitate early intervention. The implementation of cCBT will also aid the achievement of the NHS targets as it allows the treatment of high numbers of patients with minimal therapeutic input. However, the results of a recent survey of Scottish NHS boards indicated that NHS policy and infrastructure may not be fully in place to support to implementation of cCBT in many areas [17]. Overall, most of the health boards possess the required software to use cCBT programs. However, the majority of NHS health boards reported that they lack dedicated computers for patient use, hence access to cCBT at NHS sites is limited. Additionally, local policy in the majority of boards prevent staff from routinely contacting patients via email, Skype, or instant messenger, making the delivery of short efficient support sessions difficult.

For cCBT to be successfully delivered in the treatment of BN within a guided support model, as recommended by national guidelines, dedicated patient computers should be provided to allow access to online interventions. Additionally, IT policy should allow staff to support patients in convenient ways such as via email or live chat, and procedures need to be developed to allow this within confidential and appropriate clinical support. These measures would increase access to reputable, evidence-based, and expertly supported online packages such as OBO; this would serve as an alternative to poor quality mental health websites and potentially harmful "pro-ED" websites that many individuals, in the absence of specialist treatment, with eating problems encounter online [18].

Acknowledgments

This study was funded by the Helen Hay Pollock Bequest and was sponsored by the University of Glasgow and was approved by the College of Medical Veterinary and Life Science Ethics Committee, University of Glasgow (Ref. FM03508).

Conflicts of Interest

Dr. Chris Williams is one of the developers of the OBO package and holds IPR in this and a range of other free and licenced self-help computerized and book resources.

References

1. National Collaborating Centre for Mental Health. Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa, and Related Eating Disorders. London: British Psychological Society and Royal College of Psychiatrists; 2004.
2. Bara-Carril N, Williams CJ, Pombo-Carril MG, Reid Y, Murray K, Aubin S, et al. A preliminary investigation into the feasibility and efficacy of a CD-ROM-based cognitive-behavioral self-help intervention for bulimia nervosa. Int J Eat Disord 2004 May;35(4):538-548. [doi: 10.1002/eat.10267] [Medline: 15101069]
3. Murray K, Schmidt U, Pombo-Carril M, Grover M, Alenya J, Treasure J, et al. Does therapist guidance improve uptake, adherence and outcome from a CD-ROM based cognitive-behavioral intervention for the treatment of bulimia nervosa? Computers in Human Behavior 2007 Jan 2007;23(1):850-859. [doi: 10.1016/j.chb.2004.11.014]
4. Schmidt U, Andiappan M, Grover M, Robinson S, Perkins S, Dugmore O, et al. Randomised controlled trial of CD-ROM-based cognitive-behavioral self-care for bulimia nervosa. Br J Psychiatry 2008 Dec;193(6):493-500 [FREE Full text] [doi: 10.1192/bjp.bp.107.046607] [Medline: 19043154]
5. Pretorius N, Arcelus J, Beecham J, Dawson H, Doherty F, Eisler I, et al. Cognitive-behavioral therapy for adolescents with bulimic symptomatology: the acceptability and effectiveness of internet-based delivery. Behav Res Ther 2009 Sep 2009;47(9):729-736. [doi: 10.1016/j.brat.2009.05.006] [Medline: 19515360]
6. Sánchez-Ortiz VC, Munro C, Stahl D, House J, Startup H, Treasure J, et al. A randomized controlled trial of internet-based cognitive-behavioral therapy for bulimia nervosa or related disorders in a student population. Psychol Med 2011 Feb;41(2):407-417. [doi: 10.1017/S0033291710000711] [Medline: 20406523]
7. Williams CJ, Aubin SD, Cottrell D, Harkin PJR. Overcoming bulimia: A self-help package (computer software). Leeds: University of Leeds; 1998.
8. Pretorius N, Rowlands L, Ringwood S, Schmidt U. Young people's perceptions of and reasons for accessing a web-based cognitive behavioral intervention for bulimia nervosa. Eur Eat Disord Rev 2010 May;18(3):197-206. [doi: 10.1002/erv.985] [Medline: 20443203]
9. Sánchez-Ortiz VC, House J, Munro C, Treasure J, Startup H, Williams C, et al. "A computer isn't gonna judge you": a qualitative study of users' views of an internet-based cognitive behavioral guided self-care treatment package for bulimia nervosa and related disorders. Eat Weight Disord 2011 Jun;16(2):e93-e101. [Medline: 21989103]
10. McClay CA, Ewan L, Schmidt U, Williams C. Attitudes towards self-help for eating disorders. In: Proceedings of the 39th British Association for Behavioral and Cognitive Psychotherapies' Annual Conference. 2011 Presented at: British Association for Behavioral and Cognitive Psychotherapies' Annual Conference; July 20-24, 2011; Guildford, UK.
11. Fairburn CG, Cooper Z, O'Connor M. Eating Disorders Examination. In: Fairburn CG, editor. Cognitive Behavior Therapy and Eating Disorders. New York: The Guilford Press; 2008.
12. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
13. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
14. beat. beat. Choice or Chance? Ending the Information Lottery URL: http://library.nhsggc.org.uk/mediaAssets/CYPSS/Choice%20or%20Chance.pdf [accessed 2012-02-02] [WebCite Cache ID 659CiEDwT]
15. Carrard I, Rouget P, Fernández-Aranda F, Volkart AC, Damoiseau M, Lam T. Evaluation and deployment of evidence based patient self-management support program for Bulimia Nervosa. Int J Med Inform 2006 Jan;75(1):101-109. [doi: 10.1016/j.ijmedinf.2005.07.031] [Medline: 16115793]
16. Murray K, Pombo-Carril MG, Bara-Carril N, Grover M, Reid Y, Langham C, et al. Factors determining uptake of a CD-ROM-based CBT self-help treatment for bulimia: patient characteristics and subjective appraisals of self-help treatment. Eur Eat Disorders Rev 2003 May 2003;11(3):243-260. [doi: 10.1002/erv.519]
17. Kenicer D, McClay CA, Williams C. A national survey of health service infrastructure and policy impacts on access to computerised CBT in Scotland. BMC Med Inform Decis Mak 2012;12:102 [FREE Full text] [doi: 10.1186/1472-6947-12-102] [Medline: 22958309]

18. Peebles R, Wilson JL, Litt IF, Hardy KK, Lock JD, Mann JR, et al. Disordered eating in a digital age: eating behaviors, health, and quality of life in users of websites with pro-eating disorder content. J Med Internet Res 2012;14(5):e148 [FREE Full text] [doi: 10.2196/jmir.2023] [Medline: 23099628]

Abbreviations

BN: bulimia nervosa
CBT: cognitive behavioral therapy
cCBT: computerized cognitive behavioral therapy
ED: eating disorder
EDNOS: eating disorder not otherwise specified
OBO: Overcoming Bulimia Online

Edited by G Eysenbach; submitted 28.02.12; peer-reviewed by L Palmili; comments to author 06.10.12; revised version received 16.11.12; accepted 19.12.12; published 15.03.13.

Please cite as:
McClay CA, Waters L, McHale C, Schmidt U, Williams C
Online Cognitive Behavioral Therapy for Bulimic Type Disorders, Delivered in the Community by a Nonclinician: Qualitative Study
J Med Internet Res 2013;15(3):e46
URL: http://www.jmir.org/2013/3/e46/
doi:10.2196/jmir.2083
PMID:23502689

Original Paper

Utilizing Social Media to Study Information-Seeking and Ethical Issues in Gene Therapy

Julie M Robillard[1], PhD; Louise Whiteley[2], PhD; Thomas Wade Johnson[1], B.Sc; Jonathan Lim[3], B.Sc; Wyeth W Wasserman[3], PhD; Judy Illes[1], PhD

[1]National Core for Neuroethics, Division of Neurology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
[2]Medical Museion & Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark
[3]Centre for Molecular Medicine and Therapeutics, Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Judy Illes, PhD
National Core for Neuroethics
Division of Neurology, Department of Medicine
University of British Columbia
2211 Wesbrook Mall, Koerner Pavilion, Room S124
Vancouver, BC, V6T 2B5
Canada
Phone: 1 604 822 7920
Fax: 1 604 827 5229
Email: jilles@mail.ubc.ca

Abstract

Background: The field of gene therapy is rapidly evolving, and while hopes of treating disorders of the central nervous system and ethical concerns have been articulated within the academic community, little is known about views and opinions of different stakeholder groups.

Objective: To address this gap, we utilized social media to investigate the kind of information public users are seeking about gene therapy and the hopes, concerns, and attitudes they express.

Methods: We conducted a content analysis of questions containing the keywords “gene therapy” from the Q&A site “Yahoo! Answers” for the 5-year period between 2006 and 2010. From the pool of questions retrieved (N=903), we identified those containing at least one theme related to ethics, environment, economics, law, or society (n=173) and then characterized the content of relevant answers (n=399) through emergent coding.

Results: The results show that users seek a wide range of information regarding gene therapy, with requests for scientific information and ethical issues at the forefront of enquiry. The question sample reveals high expectations for gene therapy that range from cures for genetic and nongenetic diseases to pre- and postnatal enhancement of physiological attributes. Ethics questions are commonly expressed as fears about the impact of gene therapy on self and society. The answer sample echoes these concerns but further suggests that the acceptability of gene therapy varies depending on the specific application.

Conclusions: Overall, the findings highlight the powerful role of social media as a rich resource for research into attitudes toward biomedicine and as a platform for knowledge exchange and public engagement for topics relating to health and disease.

(J Med Internet Res 2013;15(3):e44) doi:10.2196/jmir.2313

KEYWORDS

gene therapy; social media; content analysis; ethics; public opinion

Introduction

The field of gene transfer, often referred to as “gene therapy”, is rapidly evolving and generating hope for the treatment of a large variety of diseases and disorders [1,2]. Research developments and clinical trials are often featured prominently in traditional news media and on the Internet [3], and both media domains contribute to public expectations and health decision making. Existing alongside promises about the medical benefits of genetic research, which are often emphasized in news media [3], are ethical concerns about laboratory and clinical research and its translation into clinical settings [4,5]. Within the

bioethics community, topics such as the risk-benefit tradeoff in human studies, potential inadvertent transmission of germline changes, the blurring of the distinction between research and treatment with attendant issues surrounding informed consent, and the possible use of gene therapy for nontherapeutic applications have been debated [4,5]. While these issues are covered prominently in the traditional news media and online when an event such as the death of a research participant occurs [6,7], little is known overall about how prospective patients and the broader public think about these concerns and indeed the extent to which they are concerned at all.

Our work draws on past studies that have polled public opinion regarding genetics research. In general, these survey studies have focused on testing existing basic scientific knowledge, perceived acceptability of treatments, and new scientific developments such as cloning. The results from these studies reveal ambivalence about gene therapy, and that acceptability is linked to the potential for treating serious diseases [8,9]. In a meta-analysis of survey research on various aspects of genetics, Singer showed that just over half of the respondents would be willing to undergo gene therapy, but that this measure of acceptability climbs to nearly 90% in the context of curing a fatal genetic disease in children or fetuses. Strong predictors of these attitudes are the degree of religious belief or practice [10]. These survey studies are all structured around fixed questions and thus do not capture emergent opinions or reflect participant-driven concerns. This is a particular concern for affective variables such as those related to ethics content, for which close-ended questions have low validity and risk creating framing effects, for instance by implying that an issue ought to be of ethical concern by asking if it is [11,12].

As people increasingly communicate through various forms of online media, it has become possible to use websites and online applications to assess freely initiated opinions and attitudes on a large variety of health-related topics [13,14]. Online social media hold particular potential for both the identification of attitudes and priorities when considering health interventions [15] and as a global, widely used, and accessible platform for engagement. As research investigating public attitudes and interactions through the lens of social media grows, various frameworks are emerging, such as narrative analysis of social media content [16]. Another such framework is infodemiology, the science of distribution and determinants of information in electronic media, including but not limited to the Internet and mobile applications [17,18]. The goal of infodemiology is ultimately to use the knowledge gained to inform public health and public policy. Examples of research using Internet parameters to survey health include the tracking of flu-related searches on an Internet search engine [19] and the examination of vaccine criticism on webpages [20]. The multidisciplinary field of infodemiology is emerging as a lens through which we can observe the health-seeking behaviors of people involved in social media and their attitudes towards health and illness [17].

Elucidating public perceptions of scientific research and clinical trials is crucial to a full understanding of contemporary biomedicine, which both influences and is influenced by public understandings of health, and which interacts with public opinion through mechanisms such as funding structures, patient advocacy, and protest, lobbying, and debate [21,22]. This broader understanding of the social context of research is also crucial for effective science communication. In recent decades, there has been much discussion of the failures of top-down dissemination of scientific knowledge to improve public trust in science, increase scientific literacy, or produce more engaged scientific citizens [23,24]. Effective communication relies on the diverse forms of knowledge, expertise, and attitudes that different public audiences bring to the conversation [23,24] and on taking seriously its multidirectional nature [25]. Though social and online media introduce concerns about accuracy, trust, and expertise [26-29], they offer new possibilities for communication and for breaking down traditional barriers between expert scientists and public audiences. Research into how social media are used is crucial to grounding future efforts to utilize these platforms to promote public engagement with biomedical research and its clinical application [30].

The proposed methodological shift from dissemination to dialogue has been motivated and accompanied by a normative argument about the political and ethical desirability of fully engaging public audiences in scientific research [23,25,31,32]. Arguments for reciprocal public engagement are particularly pressing for biomedical research, which can have a profound impact both on people's health and on their sense of self and social relations and for clinical domains where decisions about treatment are rarely black and white. The present study delivers insights into current concerns and information-seeking practices surrounding gene therapy among users of a highly trafficked social media website. In doing so, the study adds to knowledge about the specific challenges of communicating genetics research [33,34] and contributes to a growing body of knowledge about the possibilities of social media for public engagement [27,29].

Methods

Study Design

We conducted a content analysis of questions containing the keywords "gene therapy" from the Q&A site *Yahoo! Answers* for the 5-year period between 2006 and 2010.

Sample and Data Mining

We mined the online Q&A website Yahoo! Answers to obtain the sample of question and answers for this study. Yahoo! Answers is a website belonging in the social media family. It constitutes a social software that specifically supports interactive dialogue and user-generated content, blurring boundaries between media producer and consumer. Launched in 2005, Yahoo! Answers is a free, community-driven "knowledge market". On the site, users can both submit questions to be answered and answer questions posed by other users, with a points system being used to encourage participation. While points have no value, they serve as an indicator of how active a user is, and reaching point thresholds (levels) can give a user more site access. For each question, a "Best Answer" is selected either by the asker or through votes by the community. Yahoo! Answers was chosen over other similar Q&A sites based on two criteria: (1) it was the Q&A site gathering the largest traffic (as measured by unique monthly visitors measured by analytics

provider Compete) at the time of the study, and (2) returned the largest number of matches for a search of the keywords "gene therapy". According to December 2010 estimates by Quantcast, an audience measurement provider, Yahoo! Answers traffic is made up of similar proportions of male (47%) and female (53%) users, with the most present age groups being 25-24 (23%), 35-44 (22%), and 18-24 (19%).

We used a customized automatic program retrieval method to search for "gene therapy" on the Yahoo! Answers result pages in the 5-year time period between January 1, 2006, and December 31, 2010. Data fields for questions, answers, and users from each page were parsed and stored. Duplicates and irrelevant retrievals were manually removed from the database.

Coding and Intercoder Reproducibility

Coding the Questions

The first phase of the analysis considered the questions. The entire sample was coded by one investigator (JR), using a coding guide developed by 2 coders (JR and LW) from a pilot analysis of a random sample of 10% of the data. A second coder (TJ) analyzed 20% of the final sample to test for reproducibility, tested via percentage intercoder reliability. Reproducibility was initially 93%, and remaining disagreements were settled through discussion to achieve consensus.

We used an emergent coding strategy where categories were established after an initial, preliminary examination of 10% of the sample by two independent coders. The coding structure was developed to capture the salient thematic features of our sample as identified in our preliminary analysis. The final coding guide comprised the following major themes: (1) type of question (eg, request for scientific information, opinion gathering), (2) application of gene therapy (eg, disease treatment, enhancement), and (3) ethical, environmental, economic, legal, and social implications (eg, effects on self, impact on society) of gene therapy. We further coded for subthemes within each major theme. For an example of the coding strategy used for the question sample, see Table 1.

Table 1. Example of coding strategy (questions).

Theme	Subtheme[a]
Question features	
	Question ID
	Year
	User name
Theme 1 Type of question	
	Science information
	Effectiveness
	Progress
	Risks
	Ethical, environmental, economic, legal, social issues
	Education
	Careers
	Opinion
	Polemic
	Other (describe)
Theme 2 Applications of gene therapy	
	Disease treatment
	Disease cure
	Disease prevention
	Enhancement
	In utero
	Sexual orientation
	Other social modification (describe)
Theme 3 Ethical, environmental, economic, legal, social issues	
	Resource allocation
	Effects on self
	Change to society
	Discrimination
	Nature
	God and religion
	Evolution

[a] Multiple subthemes within a theme may apply to a single question.

Coding the Answers

We analyzed answers given by users to the subset of questions that contained a theme related to ethics, environment, economics, law, or society (ie, those that received a code under Theme 3) of the coding guide for questions. This focus derived from our specific interest in the views of public users on ethical and societal implications of gene therapy, shaping the development of a second, answers-specific coding guide based on an initial analysis of 10% of the answers. Intercoder reproducibility was assessed in the same manner as for analysis of the questions. The thematic categories in the answer coding guide were (1) attitudes towards applications of gene therapy (eg, explicitly for or against enhancement), and (2) ethical, environmental, economic, legal, and social implications of gene therapy (eg, risk-benefit balance, diversity). We further coded for subthemes within the main themes using the same methods applied to the questions.

Questions and answers were the unit of analysis to which individual codes were applied. We used a rich coding strategy, allowing multiple categorizations of individual questions and answers [35]. The analyses of both questions and answers could be considered a form of mixed-methods research, utilizing an initial qualitative and theory-driven evaluation of a sample to produce a coding guide amenable to further quantitative content analysis. This approach is appropriate when the goal is to produce a representative picture of a large population but without wanting to determine the coding schema in advance [36].

Statistical Analysis

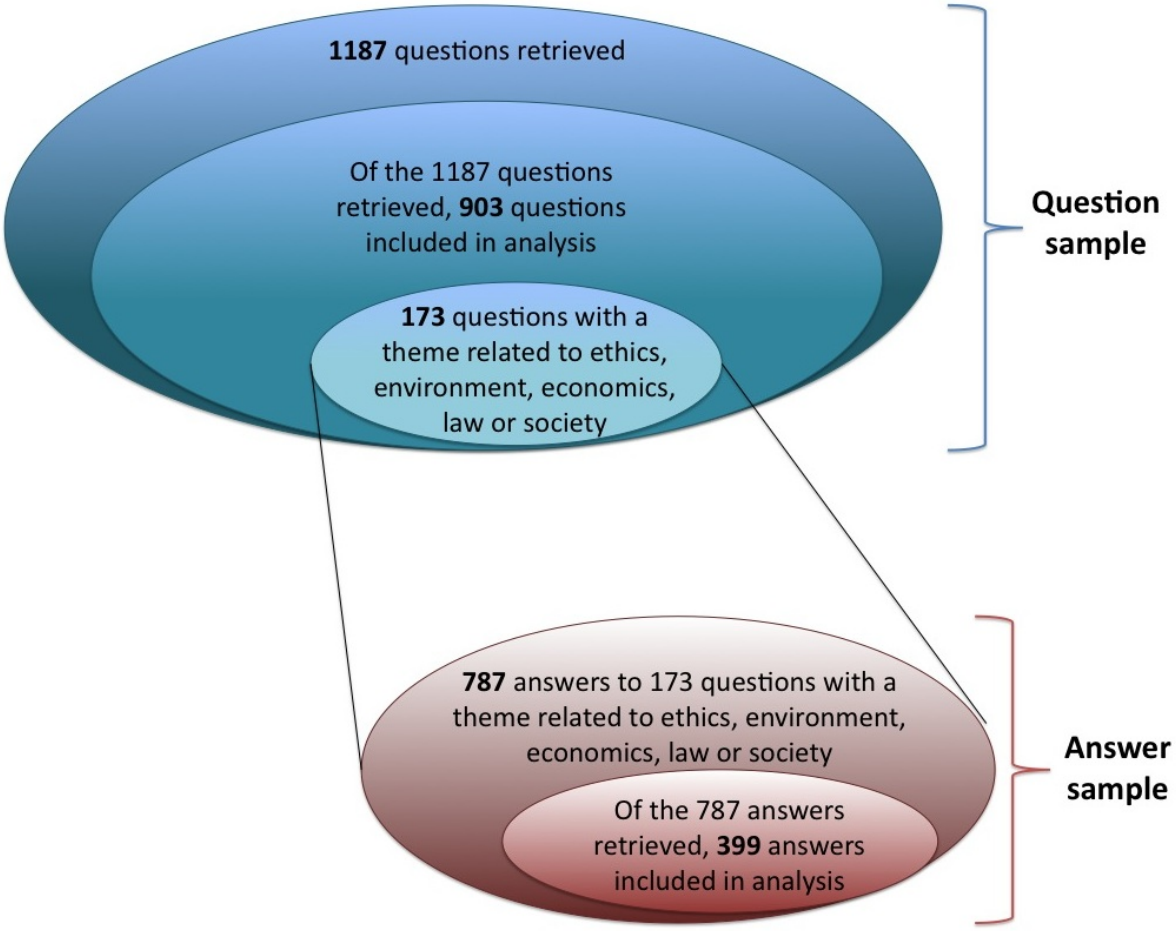
We used descriptive statistics to quantitatively characterize the composition of the both the question and answer samples generated by the coding guide.

Results

Final Sample

The initial search for questions submitted to Yahoo! Answers between 2006 and 2010 and containing the keywords “gene therapy” yielded 1187 entries (see Figure 1 for a schematic of the sampling procedure). Duplicates and questions that did not discuss gene therapy were removed. The resulting final sample for the questions contained 903 entries. From this question sample, we retrieved 173 questions containing a theme related to ethics, environment, economics, law, or society for analysis. The initial answer sample from these 173 questions contained 787 entries. Following removal of answers that did not discuss gene therapy, the resulting answer sample contained 399 entries. We chose quotations that are representative of each analytic category to illustrate and discuss individual themes.

Figure 1. Data sample. Diagram showing the samples for questions and answers and the relationships between data sets.



Questions

Types of Questions

The sample of questions regarding gene therapy was diverse (see Table 2 for a breakdown and examples). Nearly half of the questions (49%) included a request for information about the science of gene therapy. These questions related to areas such as the methods for carrying out gene therapy and the process by which cells can express new genes. Other questions focused on possible ethical, environmental, economic, legal, and societal implications of gene therapy (19%), including references to god and religion, societal outcomes of gene therapy, and impact of gene therapy on the self. A similar proportion of questions (18%) was about the progress of gene therapy and probed whether gene therapy was currently available for different diseases. Questions directly probing the opinion of other users on various aspects of gene therapy made up 13% of the sample. Smaller subsets of the questions included requests for information about careers in gene therapy (3%), information about risks (3%), and information regarding the effectiveness of gene therapy, in which users often expressed knowledge about the availability of gene therapy but were unclear on efficacy (3%). A small proportion of questions (1%) was polemic in tone.

Table 2. Examples of types of questions.

Subtheme	Frequency (%)	Examples of questions	Year
Science	49	How do they [providers of gene therapy] change the genetic makeup in all the millions of cells?	2006
Education	44	Could ny1 giv me a short paragraph on 'Somatic cell gene therapy' for a gene therapy essay on cystic fibrosis	2007
GE^3LS	19	Is it ok to build the perfect or elite human or is it only ok to fix genetic diseases such as alzimers?	2006
Progress	18	Is there any gene therapy available for parkinson’s disease?	2010
Opinion	13	If we have a gene therapy injectible, which would make you illness free, and comes free, would you go for it?	2007
Careers	3	hi, […] i really am interested in gene therapy. what would i have to take for […] post grad for a career in gene therapy?? i want to work in a lab and do research etc.... that kind of job. so what career options would be open for me?	2009
Risks	3	What are the risks involved in using gene therapy?	2006
Effectiveness	3	To what extend is gene therapy effective in treating cancerous diseases?	2008
Polemic	1	Modern Day Liberalism. Mental Illness or Mental Deficiency? I'm not sure witch one. I know the main syptom is intellectual laziness and the inability to see reality and thier surroundings. If its an illness than it should be able to be cured. If its a deficiancy than its genetic and cannot be cured without gene therapy. What do you think.	2006

Applications of Gene Therapy

All questions were coded for mentions of specific applications of gene therapy (Figure 2). Applications focused largely on disease treatment or cure (27%). A fifth of all questions (20%) mentioned a specific disease or condition; 39 such diseases and conditions were recorded from the sample ranging from benign (acne) to severe (Alzheimer disease) (see full list in Multimedia Appendix 1). Of the questions mentioning a specific disease, 25% were about cystic fibrosis, 15% were about various forms of cancer, and 7% were about diabetes, with other diseases represented by smaller portions.

Other questions mentioned nontherapeutic applications of gene therapy, such as enhancement performed in children or adults or *in utero*. Figure 2 illustrates the subthemes within enhancement, with the most common being enhancement of physical appearance. Other forms of enhancement included increasing lifespan and unusual types of enhancement, such as gaining superhuman characteristics (eg, wings, chloroplasts) (Table 3).

Figure 2. Uses of gene therapy. A) Codes for uses of gene therapy encountered in the question sample. B) Codes for the types of enhancement encountered in the question sample.

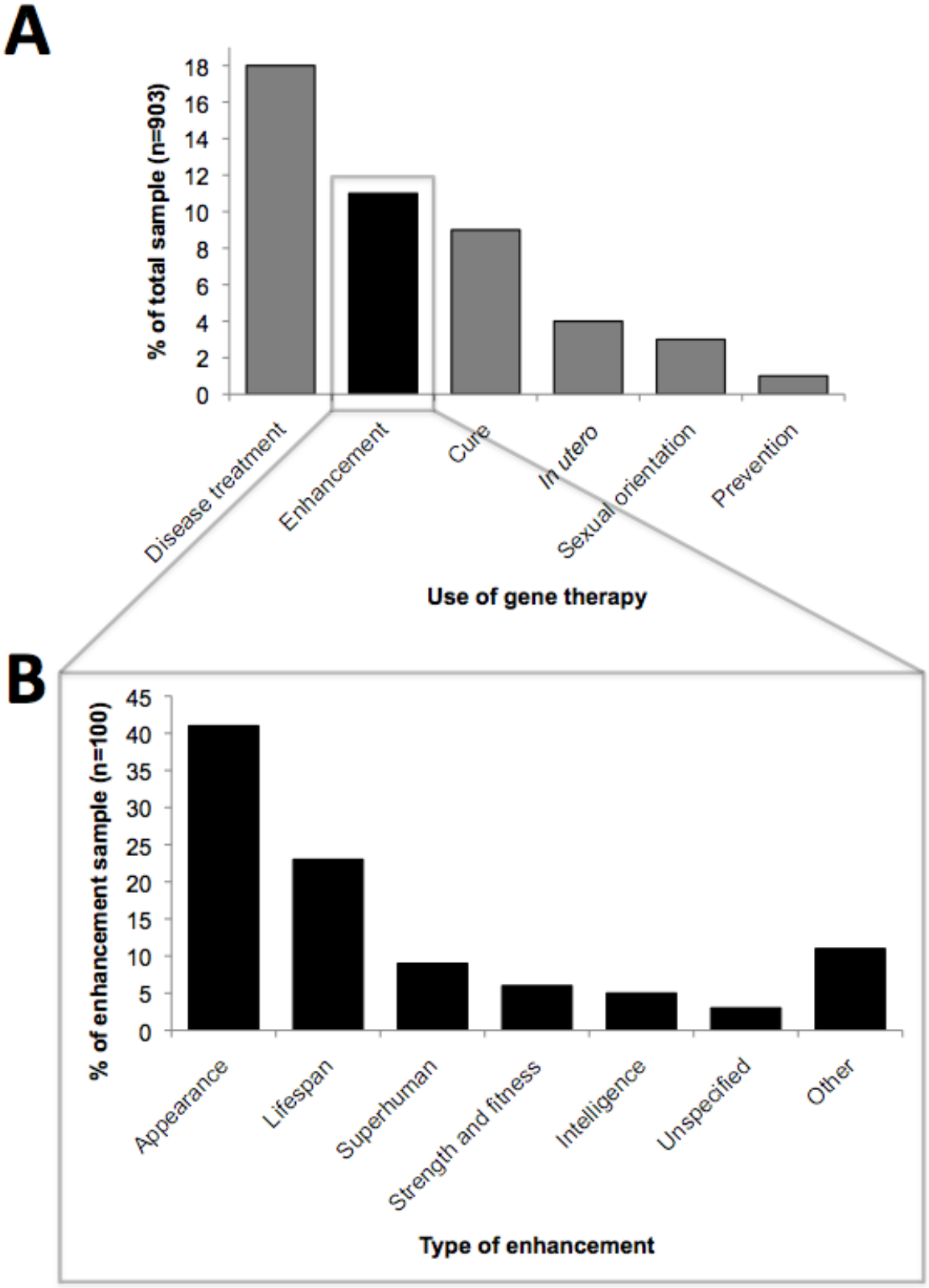


Table 3. Examples of types of enhancement.

Theme	Example	Year
Appearance	Would you have gene therapy if it really could make you look 18 again?	2007
Lifespan	A genetic researcher stated that […] gene therapy will drastically increase human life spans. Assume he is correct, how would life change?	2007
Superhuman	With today's advances in gene therapy, can people ever expect to have wings?	2006
Strength and fitness	Is there a gene therapy that can make you run faster?	2010
Intelligence	In the future, will it be possible to make humans more intelligent...? With gene therapy? I've heard about this, they're already thinking about trying it with mentally retarded children, and in the future they may well be able to do it with adults as well.	2007
Unspecified	Now that we are in the age of gene therapy and genetic engineering, would you gene boost yourself?	2007

Ethical, Environmental, Economic, Legal, and Social Implications

We further coded all questions referring to ethical, environmental, economic, legal, and societal implications of gene therapy (19% of total sample, see above section on Types of Questions). The major subthemes (Table 4) were the impact of gene therapy on society (5% of total sample), on the self (4%), and god and religion (4%). Other subthemes were resource allocation (2%), evolution (1%), nature (1%), and discrimination/equity (1%).

Table 4. Examples of ethics, environment, economic, and legal themes in questions.

Theme	Example	Year
Change to society	[…] With gene therapy and other genetic research going on out there I'm sure scientists will know how to turn off the "aging gene or genes" sooner or later. […] I'm sure there are benefits and unforseen side effects on the human body and in civilization in general.	2007
Effects on self	Do you think the tendency toward fundamentalism is genetic? If so, could gene therapy wipe this disease out […]?	2007
God and religion	Do atheists need genetic therapy? There has been a discovery of a "faith" gene in scientific research that theorizes faith could be genetic. Could this be the "mark" that is talked about in the Bible?	2010
Resource allocation	[…] I have to have 3 bioethical issues to discuss, and I can't find anything other than, death from trials of gene therapy, and that only the rich can benefit. […]	2009
Discrimination and equity	[…] But if scientists were to take [gene therapy] a step further and maybe alter genes to have perfect vision, better immune system, change genes to make a child taller, etc. If this was done successfully and it was applied to many child or embryos, would it be right to say that child that did not have gene therapy to get superior genes be inferior in a way? Then could they be considered to be a sub-class of humans […]?	2008
Against nature	If you were pregnant and found out your baby was going to be retarded, […] and the doctor told you that with experimental gene therapy it could be fixed while still in the womb by manipulating a single gene, would you do it? Or would you let nature run its course?	2009
Evolution	How cloning, gene therapy, and other technologies affect evolution?	2009

Answers

In the second phase of analysis, we studied the answers given to the subset of questions referring to ethical, environmental, economic, legal, and societal implications of gene therapy (n=173; 19% of total sample). Below we first characterize the number of answers generated by these questions and then report the coding results under the major themes of the analysis: (1) attitudes towards applications of gene therapy, and (2) ethical, environmental, economic, legal, and social implications of gene therapy.

Answer Statistics and Response-Generating Topics

Questions containing a theme related to ethical, environmental, economic, legal, or societal implications of gene therapy (n=173) had a mean of 4.5 answers each (range: 0 to 25 answers/questions). Nearly two thirds of the questions (63%) had 1-3 answers. We retrieved the questions that generated the top 1% of number of answers/question (≥ 17 answers/question). Ten questions from our sample (10/173) met this criteria. Of these 10 questions, 7 were about gene therapy to modify sexual orientation; the other 3 were about gene therapy for longevity, gene therapy for enhancement in adults, and gene therapy for enhancement *in utero*.

Attitudes Towards Applications of Gene Therapy

Users expressed general attitudes about gene therapy in 65 of the answers (Table 4), of which 75% answers were in favor of (for) gene therapy. Out of the 38 answers in which attitudes were instead expressed specifically regarding nontherapeutic

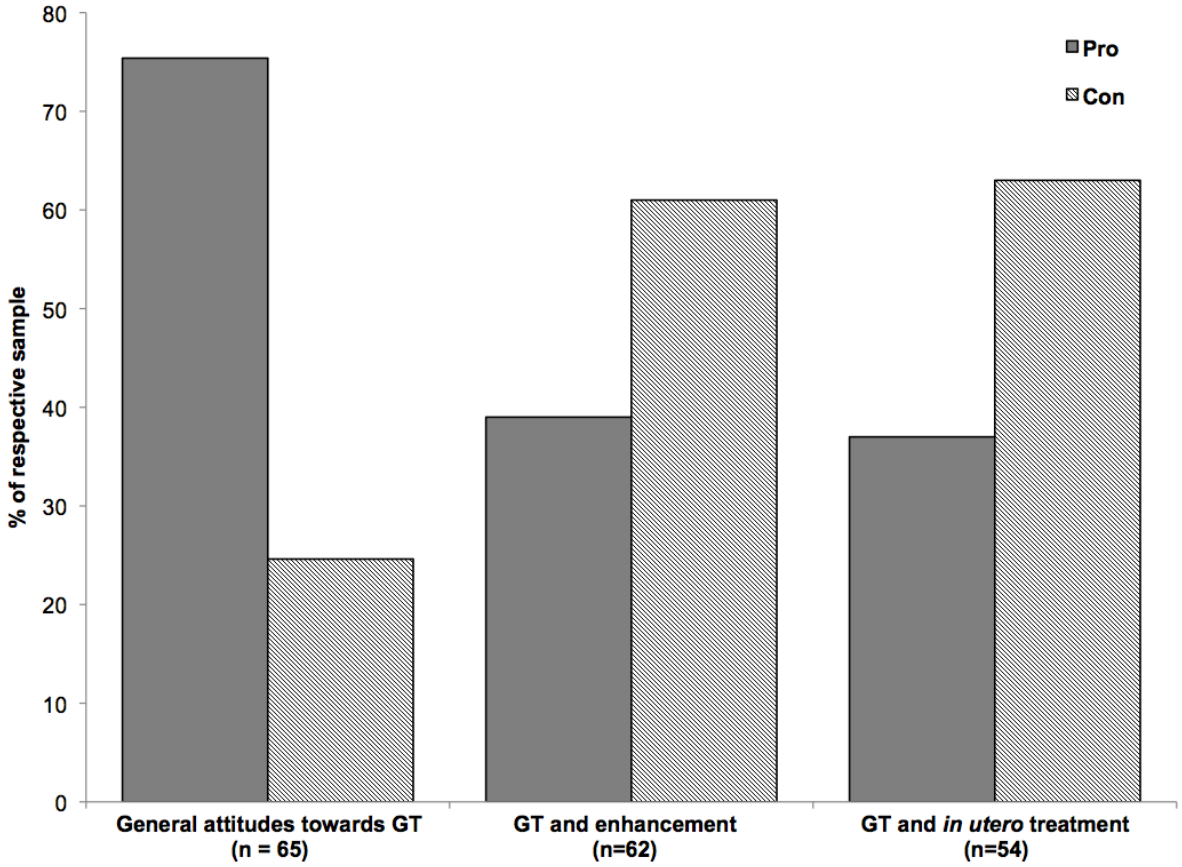
enhancement (eg, a use of gene therapy not aimed at treating or curing diseases, but rather at enhancing human characteristics such as appearance or intelligence), only 39% of users were in favor. Various arguments were stated both for and against gene therapy, which we identified through the positive or negative valence of the answers. Examples of these arguments (Table 5) included making the most of emerging technologies and abiding by religious rules. The proportions for acceptability were similar in the case of enhancement *in utero* (Figure 3).

In addition to these attitudes about applications of gene therapy, answers revealed attitudes regarding ethical aspects of gene therapy. In this context, a majority of answers (>50%) suggested that gene therapy was against nature, that it was not against religion, that it held the potential to control evolution, and that it would lead to discrimination and inequity as well as to uneven resource allocation.

Table 5. Examples of attitudes towards gene therapy and its applications.

Theme	Example	Year
For gene therapy	Support it [gene therapy], it could save lots of lives and maybe even yours one day.	2009
Against gene therapy	[…] The last thing we need to do is to genetically alter natural life.	2009
For enhancement	Cobble a gene together and create a new characteristic for man. That is the future. I hope I will be able to get some cool characteristic like a sonar.	2006
Against enhancement	We wish to go back to heaven where we once belong, so preserving/cure human health is allowed, but enhancing it may mostly seen unethical.	2007

Figure 3. Attitudes towards gene therapy. General attitudes towards applications of gene therapy encountered in the answer sample.



Ethical, Environmental, Economic, Legal, and Social Implications of Gene Therapy

While some ethical issues that arose in the question sample were also present in the answer sample (eg, impact on self and on society), new ethical issues emerged in the answers. We describe here a few qualitative examples from these themes that appeared in only a few instances. While in the questions users asked about risks of gene therapy, in the answers users responded by demonstrating reasoning about the dependence of the risk-benefit calculation on the availability of other options (1%): "[…] Although, cancer is generally treatable whereas X-SCID is fatal so it's not necessarily the end of the world if someone gets cancer since there are options" (2008). Other new concepts included conformity and diversity (2%), as some users saw the advent of gene therapy to be a threat to genetic diversity:

"Imagine if everyone could create the kid of their dreams. There would be nothing on this planet but perfect-looking human beings. And imperfect on the inside…" (2007). As well, the answers sample discussed issues around freedom of choice (2%), as some users felt gene therapy might become mandatory for certain conditions: "[…] Fixing birth defects is cool, but government interference in our baby kids is just less freedom" (2009).

Discussion

Principal Results

This content analysis of questions and answers about gene therapy from a major online social media platform provides new insights into public discourse on gene therapy. The results show that (1) social media users are seeking a large variety of types of information regarding gene therapy and, after requests for scientific information, the cluster of ethical, economic, environmental legal, and social issues are at the forefront of the discussion, (2) questions about gene therapy reflect high expectations that range from cures for a large number of diseases—both genetic and nongenetic—to enhancing various physiological features before and after birth, (3) fears primarily concern changes to self and society, and (4) the acceptability of gene therapy varies depending on specific applications.

Comparison With Prior Work

These results are consistent with those of an international survey about gene therapy in which 75% of responders supported the personal use of gene therapy, but significantly fewer supported specific applications involving nontherapeutic enhancement [37]. Those data and others suggest that acceptability of human genetic manipulation is weighed according to the perceived benefits and risks of the intervention [9]. Another previous study, examining prospective attitudes to gene therapy among patients uncovered significant concerns about effects on personal identity [38], again echoing the findings of the present study. Studies looking at genetic testing and engineering report that religious practice and beliefs are predictors of attitudes towards these technologies [10], consistent with the prominent religion theme in the present sample.

Better information on public attitudes is crucial to informing debate about the meaning, directions, and applications of biomedical research, augmenting academic discussion with public and stakeholder voices [24,31,34,39,40]. This study joins a growing body of research utilizing the potential of social media to capture such voices without the constraints of close-ended questionnaire research, or the problems of access and resources that are attendant to in-person ethnographic or sociological research [14,15]. Research harnessing the multidirectional features of online communications may take various forms, such as the tracking of the distribution and determinants of information online (infodemiology) [17-19], narrative analyses of social media content [16], computer-assisted data crawling [41], and the assessment of the impact of using social media platforms such as blogs on experiences of illness [42]. These emerging frameworks and methods are aimed at both better understanding, and ultimately improving, health and health policy, and they constitute a reflection of the growing role of social media in health communication [30].

By studying user-generated content, we were able to demonstrate high levels of interest in gene therapy—and in its potential environmental, economic, legal, and social implications—without intervening with a clipboard or audio recorder. Although the disadvantage of this hands-off approach is that we were not able to probe users' perspectives directly, we were able to sample a diverse range of apparent motivations for asking and answering questions. These included a demonstrated interest in particular diseases, dissatisfaction with physical appearance, concerns about procreation, educational needs, and, interestingly, a desire to generate debate outside of institutionalized, top-down science communication frameworks. We uncovered concerns around nontherapeutic enhancement, including discussion about impact on selfhood and authenticity—a topic deserving further research. The spontaneous nature of communication on social media thus gives insight into public interests and attitudes, and the diversity of information-seeking practices. Its participatory format also suggests possibilities for enriching public engagement. For instance, there was a high level of demand for scientific information. This finding supports arguments made in the context of Wikipedia and blogging for scientific voices to join in online discussion and information curation, rather than just posting static texts on institutional websites [29]. Overall, this study emphasizes the value of taking social media seriously—even Q&A websites, which are often criticized for disseminating inaccurate information. In particular, the prominence of spontaneous debate of ethical issues suggests that such platforms are a good setting for public engagement, augmenting more formal, top-down practices such as the consensus conference or public panel debate [43].

Despite its promise, online public engagement with health information and biomedical research raises concerns. Scientists keen to garner public trust and support, health care providers worried about poor decision making, or educators trying to promote evidence-based learning, all bemoan the "wild west" qualities of the Internet [44]. An increasing amount of research has highlighted shortcomings of online health information, ranging from uneven quality of medical information to the potential for harm and the risks of overconsumption of health information [44]. More work needs to be done to assess the current state of online information and how it can be made more accurate, relevant, and open. This study argues for attending to what people seek to know in this endeavor and to the active nature of intra-user information sharing over traditional source-to-receiver model of information transmission. Our observation that users often reasoned according to the social context and possible alternatives of particular applications, rather than giving a universal risk-benefit analysis, emphasizes the importance of elucidating context in public engagement activities.

Analysis of traditional media forms has often focused on the way in which research is framed and the potential effects on public opinion and health decision making [45]. For instance, news reports frequently frame stories about genetics in a way that emphasizes the potential benefits of genetic research [3];

as another example, deep brain stimulation is often reported in the form of miracle stories that lack discussion of ethical issues [46]. As Johansson (2011) points out, however, optimism and fears may be represented in different media information channels, resulting in an array of perspectives that may not be fully integrated [47]. When discussing the impact of a particular media form, it is thus important to remember that it is only one among many. For instance, we speculate that important media contexts for gene therapy include the positive framings of future treatments in the news media, dystopian fictional representations of cloning, the wide circulation of religious concepts in secular contexts, and a broader culture that increasingly views the physical body as a malleable platform for shaping the self [22]. Q&A websites themselves seem to be characterized by a diverse range of framings of gene therapy and situate ethics discussion within more pragmatic information-seeking and in relation to a range of positions and styles of reasoning.

Limitations

Despite the attractions of using social media as a research tool, we also appreciate the limitations of this study. Research using social media lends itself to selection bias, as it can be difficult to establish whether the study population represents the sampling population [48]. Our answer sample is also susceptible to response bias, especially when the question contains an explicit opinion that may direct and frame the discussion. While these biases inevitably limit the generalizability of our results and call for replication and convergent evidence, this study provides specific insight into the information-seeking patterns and the attitudes of the large Yahoo! Answers community.

Another limitation relates to our sample: it derives from a single social media platform, and while Yahoo! Answers boasts high traffic and relatively broad demographics, it may not represent public attitudes as a whole, including those of nonInternet users. Nonetheless, the consistency of the present findings with related studies on attitudes to gene therapy suggests that Yahoo! Answers attracts a sufficiently broad segment of the population to act as a proxy for public opinion at this level of analysis. We recognize that in aiming for a generalizable sample of public attitudes, we are not attending to the more specific interests and expertise that particular user-groups such as patients, religious groups, or health professionals might bring to this topic. The absence of reliable demographic information on Yahoo! Answers also poses a limitation. While it protects the anonymity of users, it makes it impossible to verify the authenticity of content. As well, Yahoo! Answers users are not required to volunteer sociodemographic details such as age or gender, and our study design does not involve contacting the users in any way, thus limiting our ability to undertake statistical analyses by age, gender, or other demographic characteristics to establish correlates of attitudinal findings. Finally, we cannot confirm that Yahoo! Answers users possess a clear understanding of what constitutes gene therapy. However, as our interest is in interests and opinions rather than in the accuracy of public knowledge, it is appropriate that we examine what users understand as gene therapy.

Conclusions

Overall, we find a rich discussion of gene therapy and associated ethical issues on a social media platform, which represents a spontaneous form of public engagement but also highlights a need for improved communication about gene therapy. The present work and future studies in this area are critical to inform research and medical communities of the current state of information-seeking and discussion regarding fast-paced advances in their fields and highlight the need for evidence-based and reciprocal communication between the academy and diverse publics.

Acknowledgments

Judy Illes is Canada Research Chair in Neuroethics. Support for this work was generously provided by Genome British Columbia AGCP-CanEuCre-01 (Wasserman, Illes), CIHR CNE #85117, the British Columbia Knowledge Development Fund, and the Canadian Foundation for Innovation (Illes).

Thomas Wade Johnson is currently affiliated with the Experimental Imaging Centre, Department of Radiology, Faculty of Medicine, University of Calgary, Calgary, AB, Canada.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Diseases and conditions mentioned in question sample.

[PDF File (Adobe PDF File), 72KB - jmir_v15i3e44_app1.pdf]

References

1. Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, et al. Effect of gene therapy on visual function in Leber's congenital amaurosis. N Engl J Med 2008 May 22;358(21):2231-2239. [doi: 10.1056/NEJMoa0802268] [Medline: 18441371]
2. Wilson, JM. The History and Promise of Gene Therapy. Genetic Engineering & Biotechnology News 2011 Oct;31(17):62-65.

3. Petersen A. Biofantasies: genetics and medicine in the print news media. Soc Sci Med 2001 Apr;52(8):1255-1268. [Medline: 11281408]
4. King NM, Cohen-Haguenauer O. En route to ethical recommendations for gene transfer clinical trials. Mol Ther 2008 Mar;16(3):432-438. [doi: 10.1038/mt.2008.13] [Medline: 18297094]
5. Rabino I. Gene therapy: ethical issues. Theor Med Bioeth 2003;24(1):31-58. [Medline: 12735489]
6. Couzin J, Kaiser J. Gene therapy. As Gelsinger case ends, gene therapy suffers another blow. Science 2005 Feb 18;307(5712):1028. [doi: 10.1126/science.307.5712.1028b] [Medline: 15718439]
7. Savulescu J. Harm, ethics committees and the gene therapy death. J Med Ethics 2001 Jun;27(3):148-150 [FREE Full text] [Medline: 11417019]
8. Calnan M, Montaner D, Horne R. How acceptable are innovative health-care technologies? A survey of public beliefs and attitudes in England and Wales. Soc Sci Med 2005 May;60(9):1937-1948. [doi: 10.1016/j.socscimed.2004.08.058] [Medline: 15743645]
9. Macer DR. Public acceptance of human gene therapy and perceptions of human genetic manipulation. Hum Gene Ther 1992 Oct;3(5):511-518. [doi: 10.1089/hum.1992.3.5-511] [Medline: 1420449]
10. Singer E, Corning A. The polls--trends: Genetic testing, engineering, and therapy. Public Opinion Quarterly 1998;62(4):633.
11. Barr M, Rose D. The great ambivalence: factors likely to affect service user and public acceptability of the pharmacogenomics of antidepressant medication. Sociol Health Illn 2008 Sep;30(6):944-958. [doi: 10.1111/j.1467-9566.2008.01116.x] [Medline: 18761513]
12. Shaughnessy J. Research Methods in Psychology, 7th ed. New York: Mcgraw-Hill Education; 2005.
13. O'Connor B, Balasubramanyan R, Routledge BR, Smith NA. From tweets to polls: Linking text sentiment to public opinion time series. In: Proceedings of the International AAAI Conference on Weblogs and Social Media. 2010 Presented at: International AAAI Conference on Weblogs and Social Media; May 23-26, 2010; Washington, DC p. 122-129.
14. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. Am J Infect Control 2010 Apr;38(3):182-188. [doi: 10.1016/j.ajic.2009.11.004] [Medline: 20347636]
15. Kozinets RV. Netnography: Doing Ethnographic Research Online. California: Sage Publications Ltd; 2009.
16. Chou WY, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. J Med Internet Res 2011;13(1):e7 [FREE Full text] [doi: 10.2196/jmir.1569] [Medline: 21247864]
17. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
18. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
19. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc 2006:244-248 [FREE Full text] [Medline: 17238340]
20. Zimmerman RK, Wolfe RM, Fox DE, Fox JR, Nowalk MP, Troy JA, et al. Vaccine criticism on the World Wide Web. J Med Internet Res 2005;7(2):e17 [FREE Full text] [doi: 10.2196/jmir.7.2.e17] [Medline: 15998608]
21. Bucchi M. Science in society: an introduction to social studies of science. New York, NY: Routledge; 2004.
22. Rose N. The politics of life itself: biomedicine, power, and subjectivity in the twenty-first century. Princeton, NJ: Princeton University Press; 2007.
23. Wilsdon J, Willis R. See-Through Science. London, UK: Demos; 2004.
24. Gregory J, Miller S. Science in public: communication, culture, and credibility. Cambridge, Mass: Perseus Publishing; 2000.
25. Racine E, Bar-Ilan O, Illes J. fMRI in the public eye. Nat Rev Neurosci 2005 Feb;6(2):159-164 [FREE Full text] [doi: 10.1038/nrn1609] [Medline: 15685221]
26. Bucchi M. Handbook of public communication of science and technology. New York: Routledge; 2008.
27. Mandavilli A. Peer review: Trial by Twitter. Nature 2011 Jan 20;469(7330):286-287. [doi: 10.1038/469286a] [Medline: 21248816]
28. Rodder S, Franzen M, Weingart P. The Sciences' Media Connection. Public Communication and Its Repercussions. London, UK: Springer; 2011.
29. Wolinsky H. More than a blog. EMBO reports 2011;12(11):1102-1105.
30. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
31. Pickersgill MD. Research, engagement and public bioethics: promoting socially robust science. J Med Ethics 2011 Nov;37(11):698-701 [FREE Full text] [doi: 10.1136/jme.2010.041954] [Medline: 21673017]
32. Taylor PL. Rules of engagement. Nature 2007 Nov 8;450(7167):163-164. [doi: 10.1038/450163a] [Medline: 17994064]
33. Einsiedel EF, Geransar R. Framing genetic risk: trust and credibility markers in online direct-to-consumer advertising for genetic testing. New Genetics and Society 2009 Dec;28(4):339-362.
34. Einsiedel EF, Jones M, Brierley M. Cultures, Contexts and Commitments in the Governance of Controversial Technologies: US. UK and Canadian Publics and Xenotransplantation Policy Development. Science and Public Policy 2011;38(8):619-628.

35. Denzin NK, Lincoln YS. The SAGE Handbook of Qualitative Research. California: Sage Publications; 2005.
36. Tashakkori A, Creswell JW. Editorial: The New Era of Mixed Methods. Journal of Mixed Methods Research 2007;1(1):3-7.
37. Macer DR, Akiyama S, Alora AT, Asada Y, Azariah J, Azariah H, et al. International perceptions and approval of gene therapy. Hum Gene Ther 1995 Jun;6(6):791-803. [doi: 10.1089/hum.1995.6.6-791] [Medline: 7548279]
38. Scully JL, Rippberger C, Rehmann-Sutter C. Non-professionals' evaluations of gene therapy ethics. Soc Sci Med 2004 Apr;58(7):1415-1425. [doi: 10.1016/S0277-9536(03)00336-8] [Medline: 14759686]
39. Bubela T, Hagen G, Einsiedel E. Synthetic biology confronts publics and policy makers: challenges for communication, regulation and commercialization. Trends Biotechnol 2012 Mar;30(3):132-137. [doi: 10.1016/j.tibtech.2011.10.003] [Medline: 22119159]
40. Greenberg AJ, McCormick J, Tapia CJ, Windebank AJ. Translating gene transfer: a stalled effort. Clin Transl Sci 2011 Aug;4(4):279-281 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00297.x] [Medline: 21884516]
41. Doing-Harris KM, Zeng-Treitler Q. Computer-Assisted Update of a Consumer Health Vocabulary Through Mining of Social Network Data. J Med Internet Res 2011;13(2):e37 [FREE Full text] [doi: 10.2196/jmir.1636] [Medline: 21586386]
42. Ressler PK, Bradshaw YS, Gualtieri L, Chui KK. Communicating the experience of chronic pain and illness through blogging. J Med Internet Res 2012;14(5):e143 [FREE Full text] [doi: 10.2196/jmir.2002] [Medline: 23092747]
43. Horst M. In Search of Dialogue: Staging Science Communication in Consensus Conferences. In: Cheng D, editor. Communicating Science in Social Contexts: New models, new practices. Netherlands: Springer; 2008.
44. Benigeri M, Pluye P. Shortcomings of health information on the Internet. Health Promot Int 2003 Dec;18(4):381-386 [FREE Full text] [Medline: 14695369]
45. Nisbet MC. Public Opinion About Stem Cell Research and Human Cloning. Public Opin Q 2004;68(1):131-154.
46. Racine E, Waldman S, Palmour N, Risse D, Illes J. "Currents of hope": neurostimulation techniques in U.S. and U.K. print media. Camb Q Healthc Ethics 2007;16(3):312-316. [Medline: 17695623]
47. Johansson V, Garwicz M, Kanje M, Röcklinsberg H, Schouenborg J, Tingström A. Beyond Blind Optimism and Unfounded Fears: Deep Brain Stimulation for Treatment Resistant Depression. Neuroethics 2011 May:1--15.
48. Janssens ACJW, Kraft P. Research conducted using data obtained through online communities: ethical implications of methodological limitations. PLoS Med 2012;9(10):e1001328 [FREE Full text] [doi: 10.1371/journal.pmed.1001328] [Medline: 23109913]

Edited by G Eysenbach; submitted 14.08.12; peer-reviewed by C Fry, K McArthur, A Townsend; comments to author 07.10.12; revised version received 15.11.12; accepted 29.12.12; published 04.03.13.

Please cite as:
Robillard JM, Whiteley L, Johnson TW, Lim J, Wasserman WW, Illes J
Utilizing Social Media to Study Information-Seeking and Ethical Issues in Gene Therapy
J Med Internet Res 2013;15(3):e44
URL: http://www.jmir.org/2013/3/e44/
doi:10.2196/jmir.2313
PMID:23470490

Original Paper

Using Information Technology and Social Networking for Recruitment of Research Participants: Experience From an Exploratory Study of Pediatric Klinefelter Syndrome

Sharron Close[1], MS, CPNP-PC, PhD; Arlene Smaldone[2], CPNP-PC, DNSc; Ilene Fennoy[3], MPH, MD; Nancy Reame[2], MSN, PhD; Margaret Grey[1], RN, DrPH

[1]Yale University, School of Nursing, New Haven, CT, United States

[2]Columbia University, School of Nursing, New York, NY, United States

[3]Columbia University Medical Center, Division of Pediatric Endocrinology, New York, NY, United States

Corresponding Author:
Sharron Close, MS, CPNP-PC, PhD
Yale University
School of Nursing
100 Church Street South Ste 132
New Haven, CT, 06536
United States
Phone: 1 914 420 4508
Fax: 1 914 238 5922
Email: sharron.close@yale.edu

Abstract

Background: Recruiting pediatric samples for research may be challenging due to parental mistrust of the research process, privacy concerns, and family time constraints. Recruitment of children with chronic and genetic conditions may further complicate the enrollment process.

Objective: In this paper, we describe the methodological challenges of recruiting children for research and provide an exemplar of how the use of information technology (IT) strategies with social networking may improve access to difficult-to-reach pediatric research participants.

Methods: We conducted a cross-sectional descriptive study of boys between the ages of 8 and 18 years with Klinefelter syndrome. This study presented unique challenges for recruitment of pediatric participants. These challenges are illustrated by the report of recruitment activities developed for the study. We reviewed the literature to explore the issues of recruiting children for research using conventional and IT approaches. Success rates of conventional recruitment approaches, such as brochures, flyers in medical offices, and physician referrals, are compared with IT-based outreach. The IT approaches included teleconferencing via a Klinefelter syndrome support group, services of a Web-based commercial recruitment-matching company, and the development of a university-affiliated research recruitment website with the use of paid advertising on a social networking website (Facebook).

Results: Over a 3-month period, dissemination of over 150 recruitment brochures and flyers placed in a large urban hospital and hospital-affiliated clinical offices, with 850 letters to physicians and patients were not successful. Within the same period, face-to-face recruitment in the clinical setting yielded 4 (9%) participants. Using Web-based and social networking approaches, 39 (91%) agreed to participate in the study. With these approaches, 5 (12%) were recruited from the national Klinefelter syndrome advocacy group, 8 (19%) from local and teleconference support groups, 10 (23%) from a Web-based research recruitment program, and 16 (37%) from the university-affiliated recruitment website. For the initial 6 months, the university website was viewed approximately 2 to 3 times per day on average. An advertisement placed on a social networking site for 1 week increased website viewing to approximately 63 visits per day. Out of 112 families approached using all of these methods, 43 (38%) agreed to participate. Families who declined cited either travel distance to the study site (15, 22%) or unwillingness to disclose the Klinefelter syndrome diagnosis to their sons (54, 78%) as the reasons for nonparticipation.

Conclusions: Use of Web-based technologies enhances the recruitment of difficult-to-reach populations. Of the many approaches employed in this study, the university-affiliated recruitment website supported by a Facebook advertisement appeared to be the most successful. Research grant budgets should include expenses for website registration and maintenance fees as well as online

advertisements on social networking websites. Tracking of recruitment referral sources may be helpful in planning future recruitment campaigns.

(J Med Internet Res 2013;15(3):e48) doi:10.2196/jmir.2286

KEYWORDS

patient recruitment; research subject recruitment; health information technology; social networking; Klinefelter syndrome

Introduction

Recruiting children for research can present many challenges due to parental mistrust of the research process, privacy concerns, and family time constraints [1,2]. Children with chronic and genetic conditions may further complicate the recruitment process [3]. For studies conducted in the United States, additional challenges exist including regulations and guidelines that direct how researchers contact and enroll participants for studies.

Prior to 1996, medical treatment of children was based on clinical trials and the testing of products and medications that were conducted in adults [4]. Although many treatments were effective for adults, some were shown to be ineffective or harmful to children [4]. In 1996, members of a joint workshop of the American Academy of Pediatrics and the National Institute of Child Health and Development issued a consensus statement calling for children to receive adequately tested treatments recommending efforts to include children in research [4]. In1998, the National Institutes of Health (NIH) established a policy with guidelines requiring that children must be included in all human subjects research that is conducted or supported by the NIH unless there was a scientific or ethical reason for not doing so [5]. The goal of this policy was to increase child participation in research for the purpose of generating data specific to the treatment of children [4,5]. However, inadequate representation of children in selected areas of low-prevalence diseases, orphan conditions (such as Klinefelter syndrome), and genetic conditions persist [6]. More innovative approaches to the enrollment of pediatric volunteers in clinical studies are needed.

The general public increasingly uses information technology (IT) as a source of health information. Approximately 80% of the American public search for health information using Internet sources [7]. Researchers are now turning to the Internet as a tool to recruit target study populations [8-14] for Internet-based interventions in conditions such as hypertension [15], diabetes [16-18], smoking cessation [8,12,19], human immunodeficiency virus risk management [13], and depression [14].

The Internet offers many opportunities for informing potential research participants about a study. These include email [8,20,21], discussion boards, blogs [8], search engines [20,21], study websites [20-22], and Web-based platforms for matching researchers with participants. Each form of Web-based communication provides opportunities and challenges for subject recruitment.

On the one hand, exposure of information to vast numbers of Internet users creates an enormous opportunity for visibility and communication with potential research participants. At the same time, the recruitment process can be sabotaged by problems on the Internet, such as emails sent to spam folders [8-10], discussion board and blog administrators blocking content associated with the researcher [9], and poor choice or lack of adequate keywords on study websites that diminish search engine exposure [23]. A key limitation of the use of Internet-based recruitment activities include the inability to reach socioeconomically or educationally disadvantaged groups as well as culturally diverse populations who may lack access to the Internet or familiarity with its use [10]. In the same vein, potential recruits to a study may not be receptive to unsolicited emails, or may not trust the legitimacy of the sender [8]. Clinical investigators encounter special challenges when attempting to recruit children as research volunteers, especially those who have low-prevalence diseases or genetic conditions [6,24].

The purpose of this paper is to describe methodological challenges associated with the recruitment of children as volunteers in research and to discuss how IT may improve access and enrollment of children in research. A case study of our experience in recruiting boys with Klinefelter syndrome for an exploratory cross-sectional study is used to illustrate specific challenges encountered with a difficult-to-recruit pediatric research population and how IT was used to support the enrollment of participants in the study.

Research Recruitment Challenges

Successful enrollment of clinical research participants is both a science and an art. A number of factors, including patients, health care professionals and researchers, structural and organizational entities, and history, interact to compromise or undermine successful enrollment of patient volunteers into a clinical study [1]. Patients may have limited access to research information and might not fully understand the role of clinical research in the advancement of knowledge for drug and behavioral therapy development [1,21]. They may also worry about or mistrust researchers and their institutions due to lack of understanding about the research process or associated risks and benefits [1,21,25]. Patient characteristics, such as culture, language, and religion [1], may further reduce the chance of successful enrollment. Health care providers may play an important role in gaining access to potential participants, but also may represent barriers to such access [1,2,21].

Health care providers in nonacademic settings may have a limited understanding or interest in clinical trials or may have misgivings about academic institutions [1,26]. Community health care providers may also have concerns about losing control over their patient's care, or losing the patient to another provider [1,26]. Full-time clinicians are frequently pressed for time in caring for patients and may be concerned about the additional administrative workload and lack of administrative

support for research activities [1,2,26]. This concern may lead to financial disincentives for clinical providers to become involved in informing their patients about research enrollment opportunities [26]. Researchers themselves sometimes fail to recognize how they may contribute to recruitment and enrollment problems in their own studies. Lack of training and proficiency in communication for the conduct of research with low-literacy populations may lead to misunderstandings between the researcher and potential participant and, in turn, lower response rates of participation [1].

Other barriers facing researchers include lack of attention to the mistrust of the population to be recruited, failure to demonstrate cultural sensitivity, and lack of training in understanding health care disparities in underserved populations [1]. These barriers are often unrecognized by researchers and get in the way when attempting to gather the desired sample. Structural and organizational factors may also be associated with the desire or ability of people to volunteer for research [1]. Researchers need to consider logistic arrangements to facilitate patient participation, such as creating convenient times and locations for study participation.

Communities may also be sensitive about allowing researchers entry into their environment, especially when they perceive that their participation in the scientific efforts does not result in any return or reward at the community level [1]. This concern makes it very difficult for researchers to re-enter the same community or for the community to be approached by other researchers.

The history of disreputably negative research practices persists in the minds of the public and these perceptions may influence the attitudes of potential research volunteers. The awareness of inhumane treatment by Nazi researchers during World War II and the infamous Tuskegee Syphilis Project conducted by the US Public Health Service from 1932 to 1972 [27] may promote overall fear and mistrust about the research process in the minds of many potential research participants.

Levels of Protection that Challenge Research Recruitment

In the United States, several guidelines offer protection to the public with regard to personal health care and participation in research. Public Law 104-191, also known as the Health Insurance Portability and Accountability Act (HIPAA), was enacted to protect the privacy and personal health information of the public [28]. The HIPAA requirements also guide researchers on how to protect the privacy of research participants. Although all health care providers and researchers are required to obey these laws, many members of the public may be apprehensive of the attendant side effects of disseminating private medical information by researchers. Levels of protection, designed to benefit the public, also may impede progress in the timing and accomplishment of recruitment.

Although pediatric researchers are charged with the responsibility of recruiting children for research, several challenges exist in such efforts. Because parents are legally responsible for their children, it is the parent who must be approached for permission for their child to participate. Parents' willingness to have their child participate in a study may be influenced by their perception of benefits, risks, and barriers to participation [2]. The child must also assent to the activities of the research project. The child's willingness to participate in the project may depend upon his/her developmental status and any vulnerability related to illness, chronic condition, or communication disabilities. Children may view research participation as a positive experience, including a wish to help others, reward incentives, and the desire to have a fun experience [2]. These positive motivations may be offset by anticipated unpleasantries, such as blood tests, disagreeable medication regimens, or interruptions in their daily lives [29]. The child-recruit is embedded within a family with complex daily schedules often including parental work, school schedules, and sport practices or other extracurricular activities. All members of the family, including the child's siblings, influence the busy family schedule. Researchers must anticipate and accommodate time commitments of the family as well as considerations for transportation and commute time. Finally, the parents and the child-recruit must be prepared to agree about certain participation risks and unpleasantries such as completing multiple forms and surveys, or medical examinations, including blood collection.

Methods

Case Illustration: A Study of Boys With Klinefelter Syndrome

The exemplar case illustrates our recent experience with recruiting boys with Klinefelter syndrome for participation in a cross-sectional study. Traditional approaches to recruitment fell short of obtaining the desired sample and expanding the approach with IT resulted in a significant gain in enrollment.

Klinefelter syndrome is a genetic condition caused by the presence of an extra X chromosome (karyotype 47, XXY). This condition occurs in an estimated 1 in 450-500 male births [30,31]. Although it is not rare, it is extremely underdiagnosed. Approximately 64% of affected males are not aware of the diagnosis, and of the 36% who are aware, only 10% are diagnosed in childhood [32]. Klinefelter syndrome in adults is associated with androgen deficiency, gynoid distribution of body fat, gynecomastia, small testes, and azoospermia [33,34]. Individuals diagnosed with Klinefelter syndrome during adulthood report childhood developmental delay; speech, language and learning problems; and psychological issues including depression, shyness, aggression, and social interaction difficulties [35,36]. Klinefelter syndrome poses increased health risks throughout the life span, including increased risk for cardiovascular disease, diabetes, and osteoporosis [37]. Diagnosis of Klinefelter syndrome during childhood may represent an opportunity to address both physical and psychosocial health challenges.

Klinefelter syndrome is a misunderstood condition owing to a paucity of research in children, lack of clear clinical guidelines for treatment during life stages, and unfortunate conclusion errors made by early researchers that suggested men with Klinefelter syndrome were at increased risk for criminal behavior [38-40]. As a result, Klinefelter syndrome families may struggle with inadequate information, lack of support,

perceived stigma, and uncertainties about their son's health [41]. Current research focused on boys with Klinefelter syndrome report fairly small sample sizes, ranging from groups of less than 20 [42-44] to the largest reported cohort of 93 [45]. Misunderstandings about Klinefelter syndrome may contribute to reluctance on the part of many men and families of young sons with Klinefelter syndrome to discuss or disclose information about their diagnosis to others [24].

We conducted an exploratory descriptive study to better understand phenotype, biomarkers, and psychosocial health parameters of boys with Klinefelter syndrome between the ages of 8 and 18 years [46]. The study protocol included a physical examination, blood collection for reproductive and cardiovascular biomarkers, and psychosocial health measurements including quality of life, self-esteem, self-concept, and risk for depression. The Columbia University Institutional Review Board approved the protocol for this study. For this exploratory study, sample size was based on a moderate correlation of at least 0.40 between the clinical characteristics and psychosocial variables as observed in studies of health-related quality of life and polycystic ovary syndrome [47,48]. For a correlation of 0.40 with alpha=.05, a total of 46 subjects were required for a minimum power of 80%. No previous studies with a Klinefelter syndrome population studied the relationship between clinical characteristics and psychosocial health. Recruitment was planned with traditional approaches, including contacting patients in a local pediatric endocrine practice; sending letters to pediatricians, pediatric endocrinologists, geneticists, and genetic counselors; and the use of recruitment flyers and brochures placed strategically throughout the medical center. After sending 850 letters, placing 150 brochures and fliers, and approaching 23 families during clinical visits, only 4 boys were recruited in a 3-month period. It became readily apparent that the traditional approach would fail to achieve the minimum sample size of 46 according to our sample size calculation. Thus, a more innovative approach was devised using IT and social networking.

New recruitment strategies included the development of a study website, in-person information sessions, Web-links, teleconferences, and email access to members of a national and several regional Klinefelter syndrome support organizations, as well as registration with a computer platform clinical recruitment-matching service. Each strategy is briefly described subsequently.

Klinefelter Syndrome Study Recruitment Website

A study information and recruitment website [49] was created using the keywords *Klinefelter syndrome*, *KS*, *boys with KS*, and *KS phenotype* to increase the likelihood that people searching the Internet for information on Klinefelter syndrome might find the website when conducting searches. The website pages provided information regarding the study and its eligibility requirements, study procedures, and how to contact the researcher for further information or enrollment. The study website home page screenshot is shown in Multimedia Appendix 1.

Patient Advocacy Associations

We contacted a national Klinefelter syndrome advocacy association, Knowledge Support & Action (KS&A) [50], who agreed to place information about our study with the study website link on their website. A screenshot of KS&A home page is provided in Multimedia Appendix 2. Regional Klinefelter syndrome support groups with links to the national organization then invited us to give live presentations about our study at their meetings and also agreed to send emails about the presentation and the study to their members. One of the regional groups, the Klinefelter Syndrome Global Support Group (screenshot is shown in Multimedia Appendix 3), offered a monthly parent teleconference. Over a 3-month period, we were able to explain the purpose of the study and to respond to questions regarding our protocol.

Web-Based Clinical Recruitment-Matching Service

RecruitSource is a search engine and computer platform for matching clinical research participants with researchers [51]. A screenshot of the RecruitSource home page can be seen in Multimedia Appendix 4. Researchers can register details about their study and provide eligibility requirements for matching with potential participants.

Patients who might be interested in research participation register their health information via PrivateAccess [52] as shown in Multimedia Appendix 5. This website is a secure Internet registry that enables them to control who can and cannot see all or selected parts of their personal health information. This IT-based platform prescreens the potential participants who give advance privacy directives about their health information and are asked whether they wish to be contacted by a researcher. The incentive for people using this registry is that they can share their personal health information with properly authenticated doctors, researchers, or family members on a secure Internet platform. All contact information is coded and encrypted for privacy. The potential participant gives specific permission to be contacted by the researcher. Once the patient is registered, the researcher receives information about participants who have expressed an interest in being contacted for possible inclusion in the study. This service is provided at no cost to the researcher if the RecruitSource Web link is accessed via a patient advocacy association. In this case, the study was linked to the KS&A organization, a national advocacy association for Klinefelter syndrome [50].

Social Networking

Social networking is often defined by Web-based platforms, such as Facebook and others. Social networking, however, may also include face-to-face and teleconference transactions with groups, audiences, researcher-participant, and participant-participant networking. Participant-participant networking is the central component to the recruitment strategy known as *snowballing* [53]. We used all these networking processes in our Klinefelter syndrome study. The interlinking of IT-based and face-to-face networking provided an opportunity for multiple modes of information exposure about the study. Midway into recruitment, we decided to conduct a short trial of a Facebook advertisement (ad) as shown by the screenshot

in Multimedia Appendix 6. Because we had not anticipated this strategy a priori, funding for advertising was limited. Nevertheless, we wished to observe how a 1-week social networking ad might impact exposure to the study website.

Results

Of 112 families approached, 43 (38%) agreed to participate. The most frequent reasons for families declining participation was nondisclosure of the diagnosis to their sons (54/112, 78%) and geographic distance from the study site (15/112, 22%). Most parents who had not disclosed the diagnosis to their sons feared that their sons would learn of the diagnosis through participation.

Recruitment approaches for the participants in the Klinefelter syndrome study are summarized in Table 1. Recruitment using IT and social networking yielded a greater number of participants (39/43, 91%) compared to use of traditional approaches (4/43, 9%).

Of the 69 families who declined, over one-fifth (15/69, 22%) came from direct clinical contact; almost twice that number (29/69, 42%) declined during support group presentations, and one-sixth (10/69, 16%) declined during the national KS&A meeting. The most frequent reason for decline was parents not wanting their boys to learn of their diagnosis (n=54, 78%) and travel distance to study site (15/69, 22%).

In an effort to boost activity from general Web users, we placed an ad on Facebook. The ad ran for 1 week in June of 2010, targeting a general audience. Impressions are the raw number of times an ad is shown to different Facebook users. The Facebook ad was shown a total of 2,522,169 times. Social impressions reflect the number of times the ad was shown with social context who visited the study webpage. There were 2835 social impressions for this ad resulting in 509 clicks directly to the study website's home page. At a total cost of $311 for the week's ad, this represents the researcher's cost of $0.61 per visit. Prior to placing the ad on Facebook, the study website received 2 to 3 visits per day. During the week of Facebook advertising, website visits climbed to an average of 63 visits per day. The Klinefelter study website activity increase in response to the Facebook ad can be seen in Figure 1.

Because multiple techniques were employed to attract this difficult-to-reach population, it is difficult to attribute any one recruitment approach to increasing the number of participants in this study. Figure 2 shows a timeline of the 1-year recruitment process.

Table 1. Number of participants using traditional and information technology with social networking recruitment approaches to the Klinefelter syndrome study (N=43).

Recruitment approach	n	(%)
Traditional	4	9
Physician letters	0	0
Patient letters	0	0
Brochures	0	0
Clinic referral	4	9
Information technology and social networking	39	91
Advocacy group	5	12
Support groups and teleconference	8	19
RecruitSource	10	23
Study website with Facebook ad	16	37

Figure 1. Response (visits per day to the Klinefelter syndrome study website) during Facebook advertisement period showing increase in activity during Facebook advertising.

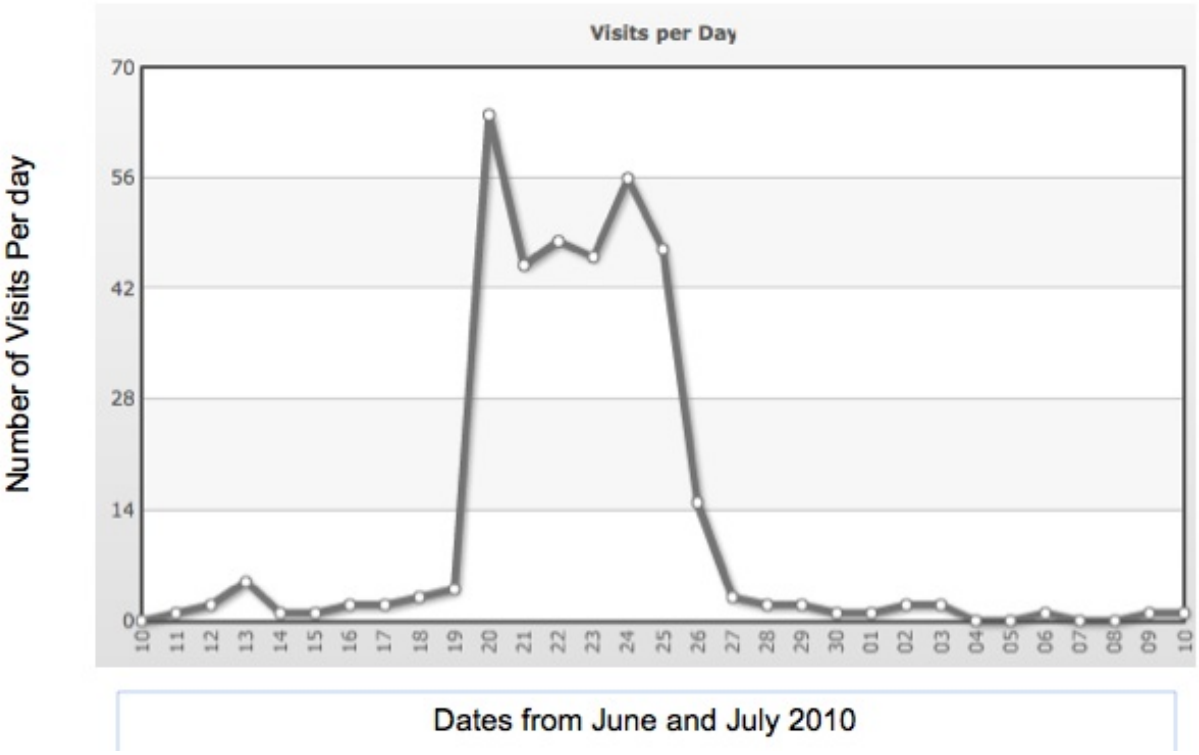
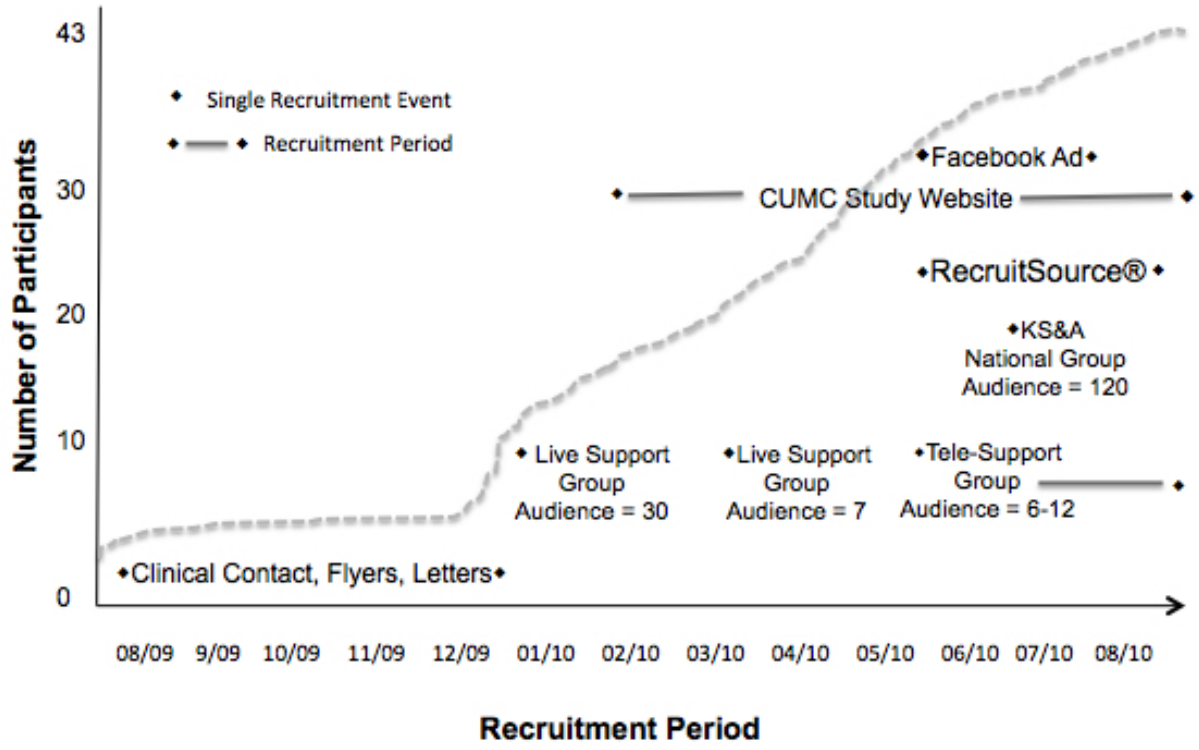


Figure 2. Recruitment to the Klinefelter syndrome study by source timeline.



Discussion

The Internet represents an increasingly valuable resource for researchers, especially for those who wish to understand the social and cultural context of the populations they are attempting to reach [54]. Since the advent of IT as a common mode of communication, researchers have learned many lessons about the pearls and pitfalls of using this recruitment approach. In difficult-to-reach and vulnerable populations, such as Klinefelter syndrome families, our experience has led us to understand

better that more than 1 recruitment technique may be required to inform potential participants and to foster trust in them. We believe that the construction and launch of the study website served these 2 important purposes. We were, however, unable to attribute increased participation due to any 1 technique, including placement of the Facebook ad. Important lessons were learned during this challenging recruitment process, such as the need to track how participants make decisions about whether to participate. We were unable to track which potential research candidates came from the study website activity while the Facebook ad ran because we did not have access to the server log. Anecdotally, several families reported that they chose to participate only after being exposed to the study information from multiple sources. Some families reported that friends or other family members who saw the Facebook ad contacted them to let them know about the study. Once informed, these families conducted either a general Internet search, visited the website directly, or visited the KS&A website for more information. Most importantly, we discovered that we need to track sources of recruitment more carefully in the future by surveying participants about how they found out about the study and also by looking at server logs whenever possible. It would also be helpful to track website visits by Internet protocol (IP) addresses to examine how many potential candidates are first-time or repeat visitors. Because we were unable to attribute which of our recruitment responses came from the Facebook ad, we are unable to estimate the cost per participant. This information would have been very helpful in planning cost allocation for a future study.

Since the advent of IT and social networking in the scientific community, there has been a steady evolution of its use for recruitment and Internet-based interventions. Even within the past 5 to 7 years, much has been learned about the limitations of Internet-based approaches and how such problems might be mitigated.

Although early experience with the use of IT-based recruitment for clinical research, as reported by Koo and Skinner [8], was disappointing, others have offered solutions to optimize challenges that make this form of recruitment difficult. Murray et al [20] solved issues related to mass emailing and spam management by providing recipients with the option to unsubscribe in order to decline further contact by researchers. They were also able to demonstrate the benefits of advertising their study on the home page of a well-known and trusted charity. Our Klinefelter syndrome study recruitment was greatly enhanced by our exposure with the KS&A national advocacy association and with support groups. Ip et al [9] addressed IT recruitment challenges by developing a guide describing a 12-step process to improve visibility and popularity of recruitment messages. The goal of this guide was to increase the interest of potential participants and to offer researchers ways in which to anticipate and respond when IT communication difficulties arise. Recent work with Ramo and Prochaska [11] demonstrated the value of Facebook advertising as an effective mechanism to reach young adults in clinical research. However, reaching a target group under the age of 18 years imposes special issues. For example, although children can be attracted to recruitment advertising for research, they would still be required to obtain parental consent for participation. Although social networking may interest a child about a research project, additional means of informing and developing trust with a parent are still necessary. Sullivan et al [13] and Graham et al [12] each illustrated how changing the composition of banner advertising may improve communication to desired target groups. In the case of pediatric research, such customization may promote discourse between parents and children. The recruitment process, as described by Patel et al [21], is explained as a dialog or discourse that takes place between the investigator and the potential research participant. In the case of minor children, discourse needs to be promoted between investigator, parent, and child if pediatric recruitment is to be successful.

Our recent experience in recruiting boys for the Klinefelter syndrome study can be described as a multilayered strategy of communication using IT. The process of communication began with a traditional print exposure that proved to be ineffective. Adding the various IT communication approaches, including the study website, a computer-based research recruitment website, social networking on Facebook, exposure via support groups online, and by teleconference, offered parents multiple exposures to study information. Although our original sample size calculation called for 46 participants based upon an effect size of 0.40, the effect size from our study proved to be larger (–0.47). We believe that the overall number of participants (43 boys) did not negatively affect the study.

Conclusions

Recruiting boys for a study on Klinefelter syndrome proved to be a challenging endeavor that was best accomplished using IT-based techniques. Important lessons were learned as we dealt with early recruitment challenges. The first lesson is that multiple exposures to the study information and personal contact with the researcher may be helpful in fostering parental trust. Parents must believe that the study, the institution, and the researcher are trustworthy before they will agree to have their child take part. These acts of communication, presented in multiple ways, were central to the success of our recruitment effort and are distinct advantages offered by IT-based strategies. Expenses related to website creation, registration of a domain name, website maintenance, and planning for social networking advertising were not initially anticipated by us, but should be considered by future researchers in the planning process when study budgets are developed. A limitation of our reported recruitment observations is that an in-depth recruitment analysis was not conducted to determine how multiple recruitment exposures occurred. Future IT-based recruitment efforts should preplan the collection of profile data, including IP addresses and tracking of how, when, and how many times a recruitment website was visited. This type of data may assist in the planning of customized approaches for the creation of more effective social networking banner advertising. Nevertheless, the observations from this study may advance the understanding of how difficult-to-recruit participants, like children, might be reached and have parental communication needs met with a view to obtaining their consent to participate in a study. It is noteworthy to mention that there has been inadequate representation of children in Klinefelter syndrome research and in other genetic conditions.

Researchers need to expand their knowledge of how potential recruits might be encouraged to participate in studies by understanding the utility of traditional approaches versus IT and other social networking approaches for recruitment. By offering multiple opportunities for exposure, parents have the opportunity to digest and think about the idea of having their child participate in a study. Because IT and social networking have become well-accepted modes of communication, these tools enable the researcher to layer the recruitment message in order to optimize the likelihood that recruitment efforts will be successful.

Acknowledgments

This research was supported by The Pediatric Endocrinology Nursing Society, the National Association of Pediatric Nurse Practitioners, the Alpha Zeta Chapter of Sigma Theta Tau, Columbia University's CTSA grant No. UL1 RR024156 from NCATS-NCRR/NIH and Yale School of Nursing T32 Post-Doctoral Fellowship Training Grant, No. 5T32NR008346-08. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of Columbia University Klinefelter syndrome study website.

[JPG File, 89KB - jmir_v15i3e48_app1.jpg]

Multimedia Appendix 2

Screenshot of Knowledge Support & Action (KS&A) Klinefelter syndrome advocacy association website.

[JPG File, 357KB - jmir_v15i3e48_app2.jpg]

Multimedia Appendix 3

Screenshot of Klinefelter Syndrome Global Support Group website.

[JPG File, 226KB - jmir_v15i3e48_app3.jpg]

Multimedia Appendix 4

Screenshot of RecruitSource website.

[JPG File, 321KB - jmir_v15i3e48_app4.jpg]

Multimedia Appendix 5

Screenshot of PrivateAccess website.

[JPG File, 449KB - jmir_v15i3e48_app5.jpg]

Multimedia Appendix 6

Screenshot of Klinefelter syndrome study Facebook advertisement.

[JPG File, 94KB - jmir_v15i3e48_app6.jpg]

References

1. Baquet CR, Henderson K, Commiskey P, Morrow JN. Clinical trials: the art of enrollment. Semin Oncol Nurs 2008 Nov;24(4):262-269 [FREE Full text] [doi: 10.1016/j.soncn.2008.08.006] [Medline: 19000600]
2. Caldwell PH, Murphy SB, Butow PN, Craig JC. Clinical trials in children. Lancet 2004;364(9436):803-811. [doi: 10.1016/S0140-6736(04)16942-0] [Medline: 15337409]
3. Gallo AM, Hadley EK, Angst DB, Knafl KA, Smith CA. Parents' concerns about issues related to their children's genetic conditions. J Spec Pediatr Nurs 2008 Jan;13(1):4-14. [doi: 10.1111/j.1744-6155.2008.00129.x] [Medline: 18096008]
4. National Institutes of Health. 1998 Mar 06. NIH policy and guidelines on the inclusion of children as participants in research involving human subjects URL: http://grants.nih.gov/grants/guide/notice-files/not98-024.html [accessed 2012-06-01] [WebCite Cache ID 686KeGP0F]

5. US Department of Health and Human Services, National Institutes of Health. 1999 Mar 17. Grants and funding: inclusion of children as participants in research involving human subjects URL: http://grants.nih.gov/grants/funding/children/pol_children_qa.htm [accessed 2012-06-01] [WebCite Cache ID 686KpGZ0Q]
6. Martinez-Castaldi C, Silverstein M, Bauchner H. Child versus adult research: the gap in high-quality study design. Pediatrics 2008 Jul;122(1):52-57 [FREE Full text] [doi: 10.1542/peds.2007-2849] [Medline: 18595986]
7. Fox S, Fallows D. Health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Washington, DC: Pew Internet & American Life Project; 2003 Jul 16. URL: http://www.pewinternet.org/Reports/2003/Internet-Health-Resources/1-Summary-of-Findings/Findings.aspx [accessed 2012-06-01] [WebCite Cache ID 686JPg5cU]
8. Koo M, Skinner H. Challenges of internet recruitment: a case study with disappointing results. J Med Internet Res 2005;7(1):e6 [FREE Full text] [doi: 10.2196/jmir.7.1.e6] [Medline: 15829478]
9. Ip EJ, Barnett MJ, Tenerowicz MJ, Perry PJ. The touro 12-step: a systematic guide to optimizing survey research with online discussion boards. J Med Internet Res 2010;12(2):e16 [FREE Full text] [doi: 10.2196/jmir.1314] [Medline: 20507843]
10. Suarez-Balcazar Y, Balcazar FE, Taylor-Ritzler T. Using the Internet to conduct research with culturally diverse populations: challenges and opportunities. Cultur Divers Ethnic Minor Psychol 2009 Jan;15(1):96-104. [doi: 10.1037/a0013179] [Medline: 19209984]
11. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
12. Graham AL, Fang Y, Moreno JL, Streiff SL, Villegas J, Muñoz RF, et al. Online advertising to reach and recruit Latino smokers to an internet cessation program: impact and costs. J Med Internet Res 2012;14(4):e116 [FREE Full text] [doi: 10.2196/jmir.2162] [Medline: 22954502]
13. 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]
14. 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]
15. Green BB, Anderson ML, Ralston JD, Catz S, Fishman PA, Cook AJ. Patient ability and willingness to participate in a web-based intervention to improve hypertension control. J Med Internet Res 2011;13(1):e1 [FREE Full text] [doi: 10.2196/jmir.1625] [Medline: 21371993]
16. Glasgow RE, Strycker LA, Kurz D, Faber A, Bell H, Dickman JM, et al. Recruitment for an internet-based diabetes self-management program: scientific and ethical implications. Ann Behav Med 2010 Aug;40(1):40-48. [doi: 10.1007/s12160-010-9189-1] [Medline: 20411443]
17. Grey M, Whittemore R, Liberti L, Delamater A, Murphy K, Faulkner MS. A comparison of two internet programs for adolescents with type 1 diabetes: design and methods. Contemp Clin Trials 2012 Jul;33(4):769-776. [doi: 10.1016/j.cct.2012.03.012] [Medline: 22484337]
18. Whittemore R, Jaser SS, Jeon S, Liberti L, Delamater A, Murphy K, et al. An internet coping skills training program for youth with type 1 diabetes: six-month outcomes. Nurs Res 2012;61(6):395-404. [doi: 10.1097/NNR.0b013e3182690a29] [Medline: 22960587]
19. Sadasivam RS, Delaughter K, Crenshaw K, Sobko HJ, Williams JH, Coley HL, et al. Development of an interactive, Web-delivered system to increase provider-patient engagement in smoking cessation. J Med Internet Res 2011;13(4):e87 [FREE Full text] [doi: 10.2196/jmir.1721] [Medline: 22011394]
20. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009;11(2):e9 [FREE Full text] [doi: 10.2196/jmir.1052] [Medline: 19403465]
21. Patel MX, Doku V, Tennakoon L. Challenges in recruitment of research participants. Advances in Psychiatric Treatment 2003;9:229-238. [doi: 10.1192/apt.9.3.229]
22. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. J Med Internet Res 2007;9(2):e10 [FREE Full text] [doi: 10.2196/jmir.9.2.e10] [Medline: 17513281]
23. Lowery S. Web-Source. 2012. Website optimization URL: http://www.web-source.net/ [accessed 2012-06-01] [WebCite Cache ID 686KSwFu4]
24. Herlihy AS, Halliday JL, Gillam LH. Ethical issues in recruiting prenatally diagnosed adults for research: Klinefelter syndrome as an example. Public Health Genomics 2012;15(1):31-33. [doi: 10.1159/000328845] [Medline: 21720152]
25. Kalfoglou AL, Boenning DA, Woolley M. Public Confidence and Involvement in Clinical Research: Symposium Summary, Clinical Roundtable, September 2000. Washington, DC: The National Academies Press; 2001.
26. Sullivan-Bolyai S, Bova C, Deatrick JA, Knafl K, Grey M, Leung K, et al. Barriers and strategies for recruiting study participants in clinical settings. West J Nurs Res 2007 Jun;29(4):486-500. [doi: 10.1177/0193945907299658] [Medline: 17538128]

27. Gamble VN. Under the shadow of Tuskegee: African Americans and health care. Am J Public Health 1997 Nov;87(11):1773-1778. [Medline: 9366634]
28. Office of the Federal Register National Archives and Records Administration. Code of Federal Regulations. Washington, DC: US Government Printing Office; 2002 Oct 01. 45 parts 1 to 199: public welfare URL: http://www.gpo.gov/fdsys/search/pagedetails.action?packageId=CFR-2002-title45-vol1 [accessed 2012-12-10] [WebCite Cache ID 6CoNNaL0x]
29. Tercyak KP, Johnson SB, Kirkpatrick KA, Silverstein JH. Offering a randomized trial of intensive therapy for IDDM to adolescents. Reasons for refusal, patient characteristics, and recruiter effects. Diabetes Care 1998 Feb;21(2):213-215. [Medline: 9539984]
30. Herlihy AS, Halliday JL, Cock ML, McLachlan RI. The prevalence and diagnosis rates of Klinefelter syndrome: an Australian comparison. Med J Aust 2011 Jan 3;194(1):24-28. [Medline: 21449864]
31. Bojesen A, Juul S, Gravholt CH. Prenatal and postnatal prevalence of Klinefelter syndrome: a national registry study. J Clin Endocrinol Metab 2003 Feb;88(2):622-626 [FREE Full text] [Medline: 12574191]
32. Abramsky L, Chapple J. 47,XXY (Klinefelter syndrome) and 47,XYY: estimated rates of and indication for postnatal diagnosis with implications for prenatal counselling. Prenat Diagn 1997 Apr;17(4):363-368. [Medline: 9160389]
33. Lanfranco F, Kamischke A, Zitzmann M, Nieschlag E. Klinefelter's syndrome. Lancet 2004;364(9430):273-283. [doi: 10.1016/S0140-6736(04)16678-6] [Medline: 15262106]
34. Bojesen A, Gravholt CH. Klinefelter syndrome in clinical practice. Nat Clin Pract Urol 2007 Apr;4(4):192-204. [doi: 10.1038/ncpuro0775] [Medline: 17415352]
35. Ross JL, Roeltgen DP, Stefanatos G, Benecke R, Zeger MP, Kushner H, et al. Cognitive and motor development during childhood in boys with Klinefelter syndrome. Am J Med Genet A 2008 Mar 15;146A(6):708-719. [doi: 10.1002/ajmg.a.32232] [Medline: 18266239]
36. Rovet J, Netley C, Keenan M, Bailey J, Stewart D. The psychoeducational profile of boys with Klinefelter syndrome. J Learn Disabil 1996 Mar;29(2):180-196. [Medline: 8820203]
37. Bojesen A, Juul S, Birkebaek NH, Gravholt CH. Morbidity in Klinefelter syndrome: a Danish register study based on hospital discharge diagnoses. J Clin Endocrinol Metab 2006 Apr;91(4):1254-1260 [FREE Full text] [doi: 10.1210/jc.2005-0697] [Medline: 16394093]
38. Baker D. Chromosome errors and antisocial behavior. CRC Crit Rev Clin Lab Sci 1972 Jan;3(1):41-101. [Medline: 4590115]
39. Schröder J, de la Chapelle A, Hakola P, Virkkunen M. The frequency of XYY and XXY men among criminal offenders. Acta Psychiatr Scand 1981 Mar;63(3):272-276. [Medline: 7234480]
40. Witkin HA, Mednick SA, Schulsinger F, Bakkestrom E, Christiansen KO, Goodenough DR, et al. Criminality in XYY and XXY men. Science 1976 Aug 13;193(4253):547-555. [Medline: 959813]
41. Cover V. Living with Klinefelter Syndrome (47,XXY) Trisomy X (47,XXX) and 47,XYY: A guide for families and individuals affected by X and Y chromosome variations. Canada: Virginia Isaacs Cover; 2012.
42. Ross JL, Samango-Sprouse C, Lahlou N, Kowal K, Elder FF, Zinn A. Early androgen deficiency in infants and young boys with 47,XXY Klinefelter syndrome. Horm Res 2005;64(1):39-45. [doi: 10.1159/000087313] [Medline: 16088206]
43. Wikström AM, Dunkel L, Wickman S, Norjavaara E, Ankarberg-Lindgren C, Raivio T. Are adolescent boys with Klinefelter syndrome androgen deficient? A longitudinal study of Finnish 47,XXY boys. Pediatr Res 2006 Jun;59(6):854-859. [doi: 10.1203/01.pdr.0000219386.31398.c3] [Medline: 16641204]
44. Wikström AM, Painter JN, Raivio T, Aittomäki K, Dunkel L. Genetic features of the X chromosome affect pubertal development and testicular degeneration in adolescent boys with Klinefelter syndrome. Clin Endocrinol (Oxf) 2006 Jul;65(1):92-97. [doi: 10.1111/j.1365-2265.2006.02554.x] [Medline: 16817826]
45. Ross JL, Zeger MP, Kushner H, Zinn AR, Roeltgen DP. An extra X or Y chromosome: contrasting the cognitive and motor phenotypes in childhood in boys with 47,XYY syndrome or 47,XXY Klinefelter syndrome. Dev Disabil Res Rev 2009 Jul;15(4):309-317 [FREE Full text] [doi: 10.1002/ddrr.85] [Medline: 20014371]
46. Close S. An Exploratory Study of Physical Phenotype, Biomarkers and Psychosocial Health Parameters in Boys with Klinefelter Syndrome [dissertation]. New York: Columbia University; 2011.
47. Hahn S, Janssen OE, Tan S, Pleger K, Mann K, Schedlowski M, et al. Clinical and psychological correlates of quality-of-life in polycystic ovary syndrome. Eur J Endocrinol 2005 Dec;153(6):853-860 [FREE Full text] [doi: 10.1530/eje.1.02024] [Medline: 16322391]
48. McCook JG, Reame NE, Thatcher SS. Health-related quality of life issues in women with polycystic ovary syndrome. J Obstet Gynecol Neonatal Nurs 2005;34(1):12-20. [doi: 10.1177/0884217504272945] [Medline: 15673641]
49. Close S. Columbia University Medical Center. 2009. 47, XXY Klinefelter Syndrome Study for Boys URL: http://www.47xxyklinefelteryouth.com/ [accessed 2012-12-10] [WebCite Cache ID 6CoNRtZXs]
50. Genetic.org. Knowledge, support & action (KS&A) URL: http://www.genetic.org/ [accessed 2012-12-10] [WebCite Cache ID 6CoNUenfW]
51. RecruitSource. URL: http://www.recruitsource.com/Pages/Home.aspx [accessed 2012-12-10] [WebCite Cache ID 6CoOHsdsK]
52. PrivateAccess. URL: http://www.privateaccess.com/Pages/Home.aspx [accessed 2012-12-10] [WebCite Cache ID 6CoOJs5ex]

53. Browne K. Snowball sampling: using social networks to research non‐heterosexual women. International Journal of Social Research Methodology 2005 Feb 2005;8(1):47-60. [doi: 10.1080/1364557032000081663]
54. 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]

Abbreviations

IP: Internet protocol
IT: information technology
KS: Klinefelter syndrome

Edited by G Eysenbach; submitted 29.07.12; peer-reviewed by M Koo, I Adeleke, S Williams; comments to author 05.10.12; revised version received 16.12.12; accepted 13.01.13; published 19.03.13.

Please cite as:
Close S, Smaldone A, Fennoy I, Reame N, Grey M
Using Information Technology and Social Networking for Recruitment of Research Participants: Experience From an Exploratory Study of Pediatric Klinefelter Syndrome
J Med Internet Res 2013;15(3):e48
URL: http://www.jmir.org/2013/3/e48/
doi:10.2196/jmir.2286
PMID:23512442

Original Paper

An Assessment of Incentive Versus Survey Length Trade-offs in a Web Survey of Radiologists

Jeanette Y Ziegenfuss[1], PhD; Blake D Niederhauser[2], MD; David Kallmes[2], MD; Timothy J Beebe[3], PhD

[1]HealthPartners, Bloomington, MN, United States

[2]Mayo Clinic, Department of Radiology, Rochester, MN, United States

[3]Mayo Clinic, Department of Health Sciences Research, Division of Health Care Policy & Research, Rochester, MN, United States

Corresponding Author:
Jeanette Y Ziegenfuss, PhD
HealthPartners
8170 33rd Ave S
Bloomington, MN, 55425
United States
Phone: 1 952 967 5020
Fax: 1 952 967 5022
Email: Jeanette.Y.Ziegenfuss@HealthPartners.com

Abstract

Background: It is generally understood that shorter Web surveys and use of incentives result in higher response rates in Web surveys directed to health care providers. Less is known about potential respondent preference for reduced burden as compared to increased reward.

Objective: To help elicit preference for minimized burden compared to reward for completion of a survey, we observed physician preferences for shorter Web surveys compared to incentives as well as incentive preference (small guaranteed incentive compared to larger lottery incentive) accompanying an electronic request to complete a survey.

Methods: This was an observational study that accompanied a large Web survey study of radiology staff, fellows, and residents at select academic medical centers in the United States. With the request to complete the survey, potential respondents were offered three options: (1) a 10-minute Web survey with the chance to win an iPad, (2) a 10-minute Web survey with a guaranteed nominal incentive ($5 amazon.com gift card), or (3) a shorter (5-7 minute) Web survey with no incentive. A total of 254 individuals responded to the Web survey request.

Results: Overwhelmingly, individuals chose a longer survey accompanied by an incentive compared to a shorter survey with no incentive (85% compared to 15%, $P<.001$). Of those opting for an incentive, a small, but not significant majority chose the chance to win an iPad over a guaranteed $5 gift card (56% compared to 44%).

Conclusions: When given the choice, radiologists preferred a reward (either guaranteed or based on a lottery) to a less burdensome survey, indicating that researchers should focus more attention at increasing perceived benefits of completing a Web survey compared to decreasing perceived burden.

(J Med Internet Res 2013;15(3):e49) doi:10.2196/jmir.2322

KEYWORDS

survey methods; Internet methods; physician surveys

Introduction

There is growing literature on methods to increase survey response rates in the general population [1] and also specifically among health professionals [2,3]. With respect to survey length and incentives, it is generally understood that physicians are more likely to respond to a request to complete a survey when presented with a shorter survey and when the request is accompanied with an incentive [2-5].

To our knowledge, there has been no study that has shown that longer surveys produce higher likelihood of response in a physician population. Some, however, have shown that shorter surveys perform better [2], and we know that with issue salience, length is associated with burden, which in turn is associated with increased nonresponse [3]. Specifically, in a comparison

of surveys of various word lengths, Jepson and colleagues concluded that 1000 words may be a threshold of where physician likelihood of response falls off [6].

There is some evidence in physician populations that small token incentives given to all (guaranteed incentives) result in higher response rates than giving respondents the opportunity to be chosen for a larger incentive (lottery incentives). In a national survey of US nurse practitioners and physician assistants, a guaranteed $5 incentive resulted in a response rate 19.5 percentage points higher than that with a lottery incentive of the chance to win $100 [7]. Similarly a survey of emergency physicians had much greater response rates with a $2 guaranteed incentive compared to a $250 lottery incentive (56% compared to 44% respectively) [8]. An Australian study simultaneously compared the impact of survey length and lottery incentives (compared to no incentives); however, this was a general population survey. They found that a lottery resulted in higher response rates to initial mailings and that a shortened survey did not have any impact. Moreover, there was no interaction effect between a shortened survey combined with a lottery incentive [9].

However, the above research—and much of the literature on which we base decisions about how to balance minimizing perceived burden and maximizing benefit—is based on experimental designs where individuals are not given the choice between burden and reward. As such, preferences cannot be observed. Instead, individuals are randomized to one of many survey conditions with the condition producing the highest response rate being the one that is considered the most effective. Any one individual does not make a choice of what length of survey or incentive they would prefer. While the use of a randomized study design is to be lauded for its ability to isolate the impact of any given experimental manipulation, it falls short with respect to knowing what a respondent would prefer if given the *choice* of competing survey conditions thought to either minimize the burden directly (ie, shortening the length) or indirectly (ie, offering incentives). Moreover, much of what is known has been ascertained in the paper survey context; little is known about such choices when Web-based surveys are considered.

Here we present results from a study where potential respondents were able to choose among three different Web survey conditions in order to determine relative preference for burden compared to reward. Specifically, in a population of radiologists, respondents could choose to respond to a short survey with no incentive, a longer survey with a token incentive, or the same longer survey with the opportunity to win an iPad.

Methods

This study was approved by the Institutional Review Board and determined to be exempt from informed consent requirements due to minimal risk. In late March and early April 2012, we sent an electronic request to radiology staff, residents, and fellows at 16 academic health centers in the United States asking them to participate in a survey about electronic learning resources available to them. The substantive findings of this study will be reported elsewhere (in preparation). The method of distributing the survey request differed somewhat between our home institution and other institutions. At the Mayo Clinic, the survey request was sent one time via personalized email by author BDN to all radiology residents, fellows, and staff. For the remaining institutions, our residency program director sent individualized email requests to each institution's residency program director with a request to forward the survey links to the same population, although the forwarded surveys were not requested to be personalized. The request was sent to 209 individuals at our home institution; however, as we were unable to determine how many individuals the request was forwarded to at other institutions, we do not know how many individuals in total received the request to participate in the survey. In all, more than 1 response was obtained from 9 of the 15 institutions for which the survey requests were to be sent by program directors.

The email request described three options for completing the survey: (1) a 10-minute survey with the chance to win an iPad, (2) a 10-minute survey with a guaranteed nominal incentive ($5 amazon.com gift card), or (3) a shorter (5-7 minute) survey with no incentive. Each option was represented by a distinct link to complete the Web survey. Within each version of the survey, slight variations were present between questions for radiologists and trainees to make them applicable, such as substituting "resident" for "staff" and "your program" for "your institution", and the third option consisted of fewer, but otherwise identical, questions.

In order to determine the extent to which different types of providers may show variable preference for length and/or incentive type, we present observed preference by key demographic variables (career stage, institutional setting, and region of the United States) and whether the individual reported using an iPad or tablet for learning or referencing radiology content. Chi-square goodness of fit tests were used to determine significance of differences. All reported differences are significant unless stated otherwise.

Results

At our home institution (Mayo Clinic), the overall response rate was 50.2% (105 responses out of 209 requests). At the other institutions, we had an additional 149 responses for a total of 254 respondents. Because the number of survey requests sent is known only for our home institution, response rates could not be calculated overall.

A large majority (85%) of respondents preferred the combination of an incentive with a longer survey to a shorter survey without any incentive. Of those respondents choosing an incentive and longer survey, 56% chose the iPad lottery incentive and 44% the guaranteed token incentive (not significant). Overall, the iPad lottery was preferred by almost half (47%) of the respondents. See Table 1.

Table 1. Survey condition choice (length and incentive) overall and by practice characteristics among responding radiology residents, fellows, and staff.

Characteristics		10 minute		5-7 minute	
		$5 gift card	iPad lottery	No incentive	*P* value
Total (n=254)		38%	47%	15%	<.001
Training					<.001
	Resident or fellow (n=136)	43%	52%	4%	
	Staff: <3 years (n=6)	50%	50%	0%	
	Staff: 3-10 years (n=21)	19%	48%	33%	
	Staff: 10+ years (n=88)	34%	39%	27%	
	Retired (n=2)	0%	50%	50%	
Setting					.427
	University (n=221)	38%	46%	16%	
	Non-university (n=28)	39%	54%	7%	
	Region				.503
	Northeast (New England, etc) (n=18)	39%	44%	17%	
	Southeast (Tennessee, Carolinas, Florida, etc) (n=57)	30%	58%	12%	
	Midwest (Ohio, Illinois, Minnesota, etc) (n=173)	40%	43%	16%	
Use of iPad/tablet for learning, referencing, or studying radiology?					.725
	Yes (n=124)	36%	47%	17%	
	No (n=117)	41%	44%	15%	

Considering stage of training, those that were later in their career were more likely to opt for the shorter survey than were their counterparts earlier in their careers. When choosing a longer survey, these individuals more advanced in their careers were more likely to opt for the guaranteed incentive. There were no differences in preference by university as compared to nonuniversity setting or regional setting of the institution. There was also no observed difference in preference by if the individual reported present use of an iPad or tablet for studying/referencing radiology. See Table 1.

The last question on the survey itself asked respondents to disclose their address if they indeed wanted the incentive sent to them (as a matter of course with the guaranteed incentive or if they were the selected individual with the lottery incentive). Those that opted for the guaranteed incentive were somewhat more likely to not give their follow-up information. More than 8 out of 10 in the lottery condition gave their follow-up information. See Table 2.

Table 2. Disclosure of contact information for gift receipt by incentive condition among responding radiology residents, fellows, and staff.

	10 minute survey		
	$5 gift card	iPad lottery	*P* value
Yes, please send me a gift or enter me in the lottery	69%	84%	<.001
No thank you	22%	4%	
Missing	9%	12%	

Discussion

Overall, the iPad incentive was the preferred choice among responding radiologists, even though it was accompanied by a Web survey that was estimated to be 3-5 minutes longer than a no-incentive condition. In fact, only 15% of respondents opted for the shorter survey option. This suggests that, among those that decide to respond to a survey, reducing burden (or perceived burden) may be less important than presenting the opportunity for individual reward via an incentive (either guaranteed or lottery-based). This pattern did not change with respect to most available practice characteristics, with the exception of career stage; those further into their career were more likely to opt for the shorter survey than their earlier counterparts. As these

individuals may have more competing demands, this observation is intuitively appealing.

Observed preference for both survey length and incentive type contradict our hypotheses based on the extant literature showing that shorter surveys and guaranteed small incentives produce higher response rates. We recognize that the experimental design is different enough from past efforts so as not to cause concern about the defensibility of our a priori hypotheses or the validity of our findings. Rather, we think that there may be something about the ability to choose in and of itself that is worthy of future investigation.

In our study, we did not test the impact of choice on response rate. However, our findings lead to the question of whether some types of choice may actually benefit the survey's outcome by empowering the potential respondent or some other yet unidentified mechanism. It is possible that offering the choice between burden and reward could alternatively decrease overall response rate as has been seen in studies of general populations that offer the choice of mode [10,11]. Either way, in this study the somewhat counterintuitive findings about preference for more survey burden in exchange for an incentive may be useful in designing future studies that could disentangle the impact of the choice in of itself from the survey condition. Ultimately, choice may be a mechanism to elicit more self-reported data from a majority of respondents.

It is also important to note that among those individuals that opted for a longer survey with an incentive, those opting for the lottery incentive were more likely to provide contact information for receipt of incentive. Fully one of five that opted for the guaranteed incentive did not choose to accept it. This could be because after completing the survey, they did not think that they needed the token reward or, perhaps, because they did not want to disclose their identity. We cannot determine the underlying mechanism from the present study design. Perhaps an investigation that is more qualitative in nature would help elucidate what the underlying mechanism(s) might be, an approach and line of inquiry recently suggested by others [3].

Our finding that among responders the iPad lottery was more preferable than the small guaranteed incentive could have important financial implications. Interestingly, if we assign the value of the iPad at $600, at the number of people who opted for this choice (120), the value was actuarially equivalent to the guaranteed incentive of $5. Of course, because the choice was offered, this could not have been predicted a priori, but it does present another potential avenue for exploration.

There are a number of important limitations to the present study that we should acknowledge to inform future work. This study was limited to radiologists at academic health centers. Due to the potentially unique composition of this population compared to other physician populations, it will be important to replicate this study design. Moreover, we do not know the denominator at institutions other than our home institution, preventing us from calculating a response rate. Incidentally, the response rate at the other institutions was likely lower than at our own (149 responses from 15 institutions) and likely due to the divergent contact protocol. However, neither the inability to calculate the response rate nor the likely lower response rate inhibits our ability to observe preference among *responders*, the primary purpose of this study. It is also important to note that all potential respondents were sent only one request to participate in the survey, whereas most physician surveys have at least two contacts. Thus it is possible that the choices we observed may be unique to early responders.

Another important limitation relates to the choices that we offered. We defined 5-7 minutes as "short" and a 10-minute survey as "long". It is possible that the perceived burden of a survey at these two lengths was not differentiated by the providers who chose to respond. Our survey lengths were driven by the needs of the overarching study, not externally validated definitions of short and long, suggesting the need for research along these lines.

Despite these limitations, this study represents a unique contribution to our understanding of surveying this policy-relevant population. Moreover, it suggests that findings garnered from experimental designs may mask important revealed preferences by individuals when they are given the choice to weigh perceived burden and reward to survey completion.

Acknowledgments

JYZ analyzed the data, wrote the manuscript, and contextualized the findings. TJB edited the manuscript and contributed to the interpretation of findings. DK and BN conceived of the experiment and collected data. All authors approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Edwards PJ, Roberts IG, Clarke MJ, DiGuiseppi C, Pratap S, Wentz R, et al. Methods to increase response rates to postal questionnaires. Cochrane Database Syst Rev 2007;2(2):MR000008 [FREE Full text] [doi: 10.1002/14651858.MR000008.pub3] [Medline: 17443629]
2. VanGeest JB, Johnson TP, Welch VL. Methodologies for improving response rates in surveys of physicians: a systematic review. Eval Health Prof 2007 Dec;30(4):303-321. [doi: 10.1177/0163278707307899] [Medline: 17986667]

3. Klabunde CN, Willis GB, McLeod CC, Dillman DA, Johnson TP, Greene SM, et al. Improving the quality of surveys of physicians and medical groups: a research agenda. Eval Health Prof 2012 Dec;35(4):477-506. [doi: 10.1177/0163278712458283] [Medline: 22947596]
4. Thorpe C, Ryan B, McLean SL, Burt A, Stewart M, Brown JB, et al. How to obtain excellent response rates when surveying physicians. Fam Pract 2009 Feb;26(1):65-68 [FREE Full text] [doi: 10.1093/fampra/cmn097] [Medline: 19074758]
5. Flannigan T, McFarlane E, Cook S. Conducting survey research among physicians and other medical professionals: A review of current literature. In: Proceedings of the AAPOR Conference. American: American Statistical Association; 2008 Presented at: American Association for Public Opinion Research; May 2008; New Orleans, LA p. 4136-4147.
6. Jepson C, Asch DA, Hershey JC, Ubel PA. In a mailed physician survey, questionnaire length had a threshold effect on response rate. J Clin Epidemiol 2005 Jan;58(1):103-105. [doi: 10.1016/j.jclinepi.2004.06.004] [Medline: 15649678]
7. Ulrich CM, Danis M, Koziol D, Garrett-Mayer E, Hubbard R, Grady C. Does it pay to pay? A randomized trial of prepaid financial incentives and lottery incentives in surveys of nonphysician healthcare professionals. Nurs Res 2005;54(3):178-183. [Medline: 15897793]
8. Tamayo-Sarver JH, Baker DW. Comparison of responses to a US 2 dollar bill versus a chance to win 250 US dollars in a mail survey of emergency physicians. Acad Emerg Med 2004 Aug;11(8):888-891. [Medline: 15289200]
9. Kalantar JS, Talley NJ. The effects of lottery incentive and length of questionnaire on health survey response rates: a randomized study. J Clin Epidemiol 1999 Nov;52(11):1117-1122. [Medline: 10527007]
10. Dillman D, Smyth JD, Christian LM. Internet, mail, and mixed-mode surveys: the tailored design method. Hoboken, N.J: Wiley & Sons; 2009.
11. Ziegenfuss JY, Beebe TJ, Rey E, Schleck C, Locke GR, Talley NJ. Internet option in a mail survey: more harm than good? Epidemiology 2010 Jul;21(4):585-586. [doi: 10.1097/EDE.0b013e3181e09657] [Medline: 20539115]

Edited by G Eysenbach; submitted 24.08.12; peer-reviewed by B Messer, Z Yan, D Parakkal; comments to author 08.10.12; revised version received 21.12.12; accepted 13.01.13; published 20.03.13.

Please cite as:
Ziegenfuss JY, Niederhauser BD, Kallmes D, Beebe TJ
An Assessment of Incentive Versus Survey Length Trade-offs in a Web Survey of Radiologists
J Med Internet Res 2013;15(3):e49
URL: http://www.jmir.org/2013/3/e49/
doi:10.2196/jmir.2322
PMID:23514869

Original Paper

Patient Experiences With Full Electronic Access to Health Records and Clinical Notes Through the My HealtheVet Personal Health Record Pilot: Qualitative Study

Susan S Woods[1,2], MD, MPH; Erin Schwartz[1], PhD; Anais Tuepker[1], PhD, MPH; Nancy A Press[3], PhD; Kim M Nazi[4], PhD; Carolyn L Turvey[5], PhD; W. Paul Nichol[6], MD

[1]Portland VA Medical Center, Health Services Research & Development, Portland, OR, United States

[2]Oregon Health & Science University, Department of Medical Informatics & Clinical Epidemiology, Portland, OR, United States

[3]Schools of Nursing and Medicine, Oregon Health & Science University, Portland, OR, United States

[4]Veterans and Consumers Health Informatics Office, Office of Informatics & Analytics, Veterans Health Administration, Washington, DC, United States

[5]Center for Comprehensive Access & Delivery Research and Evaluation, Iowa City VA Health Care System, Iowa City, IA, United States

[6]VA Puget Sound Health Care System, Seattle, WA, United States

Corresponding Author:
Susan S Woods, MD, MPH
Portland VA Medical Center
Health Services Research & Development
3710 S.W. US Veterans Hospital Road
P3HSRD
Portland, OR, 97239
United States
Phone: 1 5035044205
Fax: 1 5032735367
Email: susan.woods@va.gov

Abstract

Background: Full sharing of the electronic health record with patients has been identified as an important opportunity to engage patients in their health and health care. The My Health*e*Vet Pilot, the initial personal health record of the US Department of Veterans Affairs, allowed patients and their delegates to view and download content in their electronic health record, including clinical notes, laboratory tests, and imaging reports.

Objective: A qualitative study with purposeful sampling sought to examine patients' views and experiences with reading their health records, including their clinical notes, online.

Methods: Five focus group sessions were conducted with patients and family members who enrolled in the My Health*e*Vet Pilot at the Portland Veterans Administration Medical Center, Oregon. A total of 30 patients enrolled in the My Health*e*Vet Pilot, and 6 family members who had accessed and viewed their electronic health records participated in the sessions.

Results: Four themes characterized patient experiences with reading the full complement of their health information. Patients felt that seeing their records positively affected communication with providers and the health system, enhanced knowledge of their health and improved self-care, and allowed for greater participation in the quality of their care such as follow-up of abnormal test results or decision-making on when to seek care. While some patients felt that seeing previously undisclosed information, derogatory language, or inconsistencies in their notes caused challenges, they overwhelmingly felt that having more, rather than less, of their health record information provided benefits.

Conclusions: Patients and their delegates had predominantly positive experiences with health record transparency and the open sharing of notes and test results. Viewing their records appears to empower patients and enhance their contributions to care, calling into question common provider concerns about the effect of full record access on patient well-being. While shared records may or may not impact overall clinic workload, it is likely to change providers' work, necessitating new types of skills to communicate and partner with patients.

(J Med Internet Res 2013;15(3):e65) doi:10.2196/jmir.2356

KEYWORDS

personal health records; eHealth; patient access to records; veterans; patient participation

Introduction

> *As younger generations embrace technology, one of the oldest tools in medicine, the doctor's note, is in its infancy of reform* [TW Feeley and KI Shine, Annals of Internal Medicine, 2011]

Forty years have passed since Shenkin and Warner proposed that patients routinely be given "complete and unexpurgated copy of all medical records, both inpatient and outpatient" [1]. They forecast that record sharing would enhance patient autonomy, improve patient-physician relationships, and serve as an educational tool. In the ensuing years, the few practices that have opened notes and test results to patients have confirmed such predictions [2-4]. A recent study, referred to as *Open Notes*, provided patients at three large US health systems access to primary care notes online [5]; the great majority of patients reported greater understanding of their medical issues and recall of their treatment plans [6]. Despite this, complete sharing of health records with patients remains an uncommon and controversial practice.

US law established the right of every patient to review their medical records or request amendments [7]. Few people obtain copies of their records due to a lack of awareness of this option and a cumbersome process [8]. Yet many adults want full access to their records [9-11], believing access will help with self-care [12]. Several large health systems offer patients a personal health record (PHR) to securely retrieve test results, make appointments, refill medications, and email providers [13-15]. Patients using PHRs have been shown to be more engaged in their health and have greater satisfaction with care [16-17]. Further, PHRs may be associated with improved health outcomes [18-20].

However, PHRs do not typically provide full access to clinical notes and test results. Physician barriers to sharing clinical notes with patients have been described [6,11,21]. Reluctance stems from concerns about patient harm or confusion, burden on clinical work, and questioning of physician performance. While not all patients will choose to view their records, and health literacy is likely to affect information accessibility, patient support for full sharing of records continues to escalate [22-24].

At the US Department of Veterans Affairs (VA), the initial PHR prototype offered military veterans a "virtual window" into their health record. The My Health*e*Vet Pilot afforded a unique opportunity for patients to access their complete health records, including primary care and specialty notes, discharge summaries, and laboratory and imaging results. The purpose of this study was to understand, using qualitative methods, the experience of patients who read their records using a PHR. We sought to determine if veterans who accessed their health data and notes felt that such access had an impact on their care or their relationships with their provider(s), and if they believed that access was associated with any unintended consequences.

Methods

My Health*e*Vet Pilot Program

Between 2000 and 2010, nine VA facilities in Oregon, Florida, New York, and Washington, DC, recruited 7464 patients to enroll in the My Health*e*Vet Pilot. An enrolled patient completing in-person identity proofing could access clinic notes, hospital discharge notes, problem lists, vital signs, medications, allergies, appointments, and laboratory and imaging test results. Users could also manually enter personal data (eg, blood pressure, blood sugar, weight), access educational content, and authorize others to use the PHR on their behalf. Secure email with providers was not yet available through the PHR prototype. Figure 1 shows a screen shot of the portal landing page. Users could access their records until July 2010, after which time the Pilot was discontinued.

Study Design and Setting

To explore patient perceptions of having full electronic access to their health records, we conducted a qualitative study using focus group interviews. A semistructured discussion guide (Table 1) was designed to elicit feedback about how participants accessed their information, whether and how communication with providers was affected, and emotional and behavioral experiences resulting from seeing their clinical information and providers' notes. The study was approved by the Institutional Review Board of the Portland VA Medical Center.

Sampling and Recruitment

This study was conducted at the Portland, Oregon VA Medical Center. This facility achieved the highest enrollment in the My Health*e*Vet Pilot program, with 72% (5361) of enrollees among nine Pilot sites. Since the primary goal was to interview patients with a recent experience viewing their health records, we chose to recruit those who accessed the PHR during the 18-month interval before recruitment. Portland VA patients were eligible if they logged in and accessed any part of their record between January 1, 2008, and June 30, 2009. Excluding patients who did not use the PHR or accessed records only before 2008, a total of 697 patients met the criteria for eligibility.

We used a purposive sampling strategy to recruit patients for 5 focus groups. For 4 groups, we targeted those who accessed their records 10 or more times during the 18-month interval, with the rationale that this level of use would ensure participants had gone beyond a trial of PHR use. For a fifth group, we recruited patients who accessed their record 2-3 times during the interval. This allowed us to examine the experiences of patients accessing their records less frequently.

Figure 1. My Health*e*Vet Pilot landing page screenshot.

Recruitment letters were mailed to a random sample of 126 eligible patients; 45 patients and 2 family delegates responded, of which 40 expressed interested in the study. The principal investigator telephoned the 40 responders, providing more information about the study and inviting each to attend a focus group. A total of 30 patients and 6 family members attended a session. Groups averaged 7 participants. Patient age ranged from 49-82 years and 4 patients (11%) were women. Five of the delegates were women. Participants signed informed consent for the study and for audio recording the session and were given US $25 reimbursement for travel.

Data Analysis

Focus groups were conducted between November 2009 and January 2011 and moderated by an experienced facilitator (NP or SW). Interviews lasted between 60 and 90 minutes. Sessions were audio recorded and transcribed, then coded using a conventional content analysis approach in which themes of interest emerge inductively during analysis after all data are collected [25]. Two researchers independently coded the transcripts using QSR NVivo 9 software. The team met regularly to iteratively reach consensus on code definitions, identify themes, and resolve any coding discrepancies. Final intercoder agreement on individual codes ranged from 89% to 100%. Common broad themes identified included perceived benefits to self-care and perceived benefits to participation in care, as well as positive and negative experiences with the My Health*e*Vet Pilot program. More granular analysis of these themes is the focus of this paper.

Table 1. Focus group interview guide.

Item #	Questions
1.	Relate your experience looking at your medical records prior to the My Health*e*Vet Pilot.
2.	Talk about your experiences reading your medical records online and what type of information that you reviewed.
3.	How did you make sense of the medical records you viewed? What did you do when you read information that you didn't understand?
4.	In your opinion, what has been positive about viewing your VA medical record?
5.	In viewing your records, was there anything that you didn't anticipate or that surprised you?
6.	Have there been any negative issues with viewing the records? Have you experienced any stress or anxiousness from viewing any part of the records?
7.	Did you talk to your doctor or provider about viewing the record? Can you discuss a particular experience related to this?
8.	How has viewing your medical record impacted your relationship with your provider(s)? Have you noticed any changes in how your doctor or provider writes their notes?
9.	After seeing your medical records, did you request any changes to the content of the records?
10.	Did you find, as a result of viewing your records, that you changed (increased or decreased) the amount of times you called or visited the VA? If these changed, why do you think so?
11.	What are your feelings about continuing access to medical records in the future?
12.	Having been through the experience of having access to your medical records online, what would you want to tell others?
13.	What expectations did you have for My Health*e*Vet Pilot Program? What did you think or hope would happen, reading your medical records online?

Results

Four broad themes characterized patients' experiences with reading the full complement of their health information (summarized in Table 2). Three themes related to patients' perceived benefits of electronic record access, and one theme focused on their concerns. First, patients reported that seeing their records had a positive effect on care communication between visits as well as during encounters. Second, access was felt to improve patients' knowledge about their own health and prompted greater desire for self-care. Third, patients reported that health record access improved participation in their care in a variety of ways. Last, patients described challenges resulting from viewing clinical notes. Because there is a great deal of interest in the field regarding the potential for patient distress created by access to notes, our coding was highly sensitive to this category of comments. We analyzed the minor theme of patient difficulty in order to contribute to this discussion. Each theme is explored further, with examples of patients' statements illustrating subthemes.

The analysis did not find any recurring thematic differences between the experiences of patients having higher frequency of PHR usage (Groups 2-5) compared to lower usage (Group 1). We analyzed the frequency of code occurrences, or how often a passage was assigned a particular code, and the coverage of codes, or percentage of total text to which a code was applied. For all groups, the thematic content was found to be similar. While there were more numerous, less detailed utterances by patients in Group 1, the perceived experiences of full record access were broadly shared across the focus groups. Therefore, our findings are pertinent to all focus groups.

Perceived Enhanced Communication With Providers and Health Care Team

One benefit frequently described by patients was that access to health record information served to facilitate communication about their care. Patients reported better recall of appointments and care issues, felt more prepared for in-person visits, and found a greater ability to communicate with providers inside and outside the VA.

Communication Supplement

Access to the record was considered to be a valuable supplement to communicating in-person with providers. Several patients reported feeling less reliant on providers and staff to relay pertinent information during or between visits which, in turn, allowed them to avoid situations such as remembering in-person discussions or waiting for a phone call to be returned: "Then eVet came along and the write-ups were real great, and I didn't have to wait 6 months to talk, or a year to talk to the doctor, find out what happened almost a year ago." (FG3) and:

> *Often I get very stressed at a doctor's appointment, don't remember half of what's going on and I could go on to eVet and get my information and go, "ok, we're not in sync with this"…that helped a great deal* [Focus Group Two, FG2]

> *If they tell you something you don't understand or you forget, because maybe it's bad news or something, you go home and you really don't remember. Somebody will say, "What did they tell you?" Well, I don't know, but if you go on HealtheVet, you can find it.* [FG4]

Appointment Recall

A commonly cited benefit was assistance remembering appointments and scheduling follow-up. Patients felt shared

records offered advantages over mailed letters that may get lost or misplaced: "I like the appointments. You know, sometimes you forget. You go in there every couple of days and 'Oh yeah, I got one coming up'." (FG4)

Preparing for Encounters

Access to notes was seen as a way to help prepare for clinic visits. Patients described how knowledge of record content allowed for a better understanding of what questions to ask and consequently, to improve the visit by leading to a more efficient encounter:

> *I can go in and ask more intelligent questions and we don't have to spend as much time with them explaining everything to me. And then, with my stress level up at the doctor's office, I don't hear half of it and then we may have to do it again and again and so, it helps us to have better communication* [FG5]

> *It kind of better prepares me for the upcoming appointment, because I've got the data in my own hand. So, without starting out all over again, him repeating a whole lot of history, we can start a conversation at the treatment level we're at right now.* [FG5]

Sharing Health Information

Patients described how electronic access to records helped them coordinate care between VA and nonVA providers. They focused on ease of access and how the information offered important data that otherwise might not have been readily available:

> *I got on it because I have a civilian doctor as well as VA. Because I live so far from VA, so if I have a problem I have a local doctor. It's a hundred miles up here, so that's the reason I want the records because he, my doctor, needs to see those records too.* [FG1]

Perceived Improved Patient Knowledge and Self-Care

For many patients, access to their records increased perceived knowledge about their medical conditions and fostered a greater sense of control of their care. Several commented that seeing what was written about them prompted more efforts at self-care. Patients collectively and repeatedly discussed using the Internet to help understand medical information in their record.

Insight Into Health Conditions

Patients shared that access to clinical notes gave them what they perceived to be greater insight about their health conditions and treatment plans: "It was nice to be able to see what diagnoses you had, what your conditions were because sometimes, they're just more than you can remember." (FG2) and "Being a diabetic and having liver disease, it was a lot of information, a lot of instruction from doctors. I was always looking back to see what they said, instructions and everything." (FG5)

Insight Into Provider Assessments

Patients also discussed experiencing benefit by gaining insight into their providers' perspectives that came from reading clinical notes:

> *We talked about it when I was in there, but I don't think I gathered it all up. To see it down in writing, I think helped me. He [the doctor] was concerned, and he put it down there. I think just seeing it in writing actually opened my eyes a little bit more.* [FG2]

> *Doctors aren't real gabby and they never tell you everything. Even if you ask questions, they'll sort of slide around them. They don't have the time, you know. I found stuff out that I was just amazed at, truly, about myself.* [FG4]

Personal Control for Self-Care

Patients described wanting to be more responsible and take more control of their health issues as a direct result of reviewing their records, which was perceived as a positive influence:

> *It personally helped me assume the role of taking care of my own health, which my wife, a nurse, said, "I'm not taking care of you anymore. You've got to take care of you". So, all of a sudden, I had access to the information so I could do that. That was very positive.* [FG3]

> *I found the cholesterol and all that other stuff, and it made me start thinking about my lifestyle and how I needed to do a little bit more on my own and not depend on the doctor to hand me pills and stuff. It encouraged me…I knew more. I understood more.* [FG4]

Self-Directed Internet Research

Patients in every group reported using online search tools to learn more about diagnoses, tests, and abbreviations. Patients felt that terminology could be confusing, yet they valued access to all the information and tried to understand medical jargon on their own before querying providers.

> *Well, you could just pop over to Google or go to the library in there, a dictionary in there, you could pop over and check it out and see what it's saying instead of sitting there sweating it out trying to figure out what it is.* [FG4]

> *Well, sometimes I can figure out a problem myself either by my own online research or by just thinking about it and saying, "Well, I'm going to try this and fix it without the doctors involved.* [FG2]

Perceived Greater Patient Contribution to Care

Patients brought their own perspectives about health care quality while talking about their experiences reading their records. Several patients shared how reading notes led to awareness of a service that was needed, or not needed, from their standpoint. They demonstrated how the information prompted more active patient participation in discussions about their health issues and their care.

Monitoring and Reminder Assistant

A number of patients described how access to their information, in their view, served as a way to monitor needed follow-up, ensuring that appropriate evaluation or treatment was completed: "I found out I was anemic. No one ever told me. Then when I

asked the doctor about it, I was on iron pretty quick after that." (FG4) and:

> *I had an ultrasound on my liver and I saw the results online. It said, "Re-do in six months". Well, six months came around and nothing happened. So I called the doctor and say, "Well, it said here six months, re-do" and he said, "Well, let me look at your records". He says, "Oh yeah, they did say that". So, if I hadn't reminded him, I probably wouldn't have got it* [FG5]

Engaged to Discuss Health and Health Care

Patients talked about how reading their notes led to more dialogue with providers about what was written, including offering their opinion about whether or not the information was consistent with what had been discussed during visits:

> *My Oncologist was a pretty up-front guy. But I got on HealtheVet I found out he wasn't as up front as I thought he was…with his comments, what he had written. So, when I went to see him the next time, I said, "I'd like to know, what you think and what you know, and what you're predicting. So, rather than just write it in there, tell me and then write it".* [FG4]

Participatory Care and Shared Decision-Making

Patients offered stories about how access to their records allowed them to play a more active role with their providers as an advocate about their care and treatment:

> *It just probably made me healthier than I might have been without having the information available, to either talk to the doctor, you know, just something as simple as changing a medication for something. You know, going, "Hey, look, the thing you got me on ain't really working that great. Let's try something different. What do you think about this?"* [FG3]

Perceived Challenges From Reading Notes and Electronic Documentation

Patients were repeatedly probed about stress or harm related to reading notes. Three people responded directly that reading notes caused initial stress. In some instances, patients expressed discomfort about the language in notes, errors or inconsistencies in note content, or strain on patient-provider dialogue. At the same time, several patients voiced contrary views, opening up discussions about pros and cons of having their information available. A number of patients were frustrated by technical problems that became more prevalent prior to the closure of the Pilot.

Disclosure of Information

Statements illustrated challenges stemming from viewing newly revealed information that had not previously been disclosed to patients. One participant, a wife of a patient, expressed stress upon seeing an operative report; when asked if reading such notes was harmful, she denied harm had ensued: "I would rather not have known. There was a lot of little things they wrote, you know, step-by-step what had happened in his operation." (FG1)

A second participant described concern about the potential for negative consequences from access to all of his health information:

> *I think that's like a latent danger in all this information and having access to it. Meaning, I have access to my records and if there's something that I don't quite understand, I go to an outside source for feedback and even that might come across clear as mud. So now I'm sort of left with dis-information because it's not serving me. That's one of the dangers I see for me personally.* [FG1]

A few patients shed light on the challenge that greater information access provides data that are both helpful as well as worrisome: "I think being able to see the tests results has created, sometimes a little more stress because 'Uh-oh, look at this.' But overall, less worry, less stress because now I have facts." (FG5) and:

> *Last year I had some mental health issues. I was able to see some things that I didn't like what they said, one of them being about causing my own illness. Well, I read more about it online because I'm going, what does he actually mean by that? It explained a lot to me and calmed me down a lot. It really helped having the records and reading what they said.* [FG2]

Other participants denied stressful experiences from access to their information, feeling that having all of the health record available offered greater benefit than having partial information:

> *I think this is a way to encourage them to write the whole truth and all of the truth, you know, to say it all the way they feel it. The better for the patient. "Well, we don't want to tell the person this because it may make them upset"…I say that's a lot of bull. I want to know.* [FG2]

> *Just knowing....is better than not knowing, I think, in most things in life. Because you can imagine a lot of stuff in your health world, when you think, "Is this a pimple or am I dying of cancer?" You go through that whole thing. So, just knowing, just being able to review that and say, "Okay, I'm not dying of cancer. That's a pimple"…gave me peace of mind.* [FG3]

One patient reported that their provider did not agree with this level of patient access: "My doctor, she didn't like the fact that I could see the progress notes, saying there should always be a place where the doctors can write private notes to each other." (FG4)

Language in Documentation

A small number of patients mentioned what they believed to be offensive language in notes. Some reported talking to the note writers about their observations:

> *There were a couple of cases where someone had written something derogatory about me. I was able to see it and find out the person fabricated what they were saying and I could bring that to the doctor's and admin's attention without having to plead my case. It was written in black and white.* [FG5]

> *I went to optometry, and the guy said, "These [glasses] aren't VA, I can't deal with them." Well, I left. Then I was looking through my notes on something else, and I ran across this guy. And "hostile" was the word he used [in the note]. I felt there was a wall there after that.* [FG1]

Inconsistencies in Content

Several patients perceived inconsistencies in notes, citing that information given verbally at visits was not equivalent to that written in notes. Some patients spoke with providers about these discrepancies; many did not. A few formally requested a change in the record. One patient discussed seeking an alternate provider:

> *One time I found something totally wrong in his notes. It was like not even me. When I told the doctor, he corrected it. I went back in after the visit and read it, and he had, he'd gone in and edited his notes. It was a real mistake, which could have cost me down the road.* [FG3]

> *And sometimes, you want to change doctors as a result of what the doctor said. Not because he's describing it wrong, but he's describing it incompetently because you know yourself better than he does in many cases* [FG2]

Observations on Electronic Records

Several patients commented about electronic records and the ability to access information remotely. Some expressed a preference for notes that were thorough and without redundancies; one man criticized the presence of "boilerplate" elements in notes. A few patients voiced frustration over technical issues with the Pilot, particularly in 2010 when notes were not updating in a timely manner. All patients were disappointed with its impending closure:

> *But the Pilot is part of the whole system and if we can make things better by using the pilot and making our suggestions, it makes it better for all the veterans who've had bad experiences and that's a good start for the VA.* [FG2]

> *Communication is the key and if you can't communicate with your physician, either electronically or verbally, and this particular vehicle that we had given to us, was to me personally one of the best communication tools that the VA's ever come out with.* [FG3]

Table 2. Summary of themes on patient experiences with full record access.

Theme 1: Perceived enhanced communication with providers and health care teams	
	Supplements in-person communication
	Improved recall of appointments
	More prepared for encounters with providers
	Greater ability to share information with non-VA providers
Theme 2: Perceived improved patient knowledge and self-care	
	Improved understanding of health issues
	Greater insight into provider assessments and recommendations
	Improved sense of control of health issues
	Prompt to use the Internet to understand information
Theme 3: Perceived greater patient participation in care	
	Prompt to remind health care team for appropriate care or follow-up
	More engaged to discuss health and health care issues
	More able to participate in decisions if care is needed or not
Theme 4: Perceived challenges from viewing records and electronic documentation	
	Stress related to information not routinely disclosed
	Concern about language in notes
	Inconsistencies or errors in documentation
	Observations on electronic records and PHR technical problems

Discussion

Our findings support prior qualitative research that shows full health record access is empowering for patients and caregivers [26]. Patients' perspectives provide insight into how shared notes can foster active patient participation in their care. In all focus groups, participants put knowledge from their records to use by learning more about their health issues, gaining more knowledge about their providers' views, and advocating for themselves in discussions about their care. Reading health information in an unpressured manner allowed patients time to contemplate its content and meaning. Records were also a starting place for online research. As a result, patients felt more prepared for clinic visits but sometimes were also less likely to

call the clinic or request an appointment. Of particular interest were stories of patients serving as their own "clinical reminders", making an effort to improve the quality of their care by ensuring follow-up care was provided.

While participants reported that viewing their records was positive, a number described some difficulty upon seeing clinical notes. Predominant issues identified were the use of derogatory terms, stress that initially emerged from reading detailed personal information, and challenging conversations with members of their health care teams. At the same time, most participants, including many who cited these concerns, believed it was important and valuable to have all of their health record data. Of note, study participants viewing their records a few times expressed similar themes to those logging in more frequently, suggesting that patient use of PHRs at any given time is driven by dynamic factors such as personal health needs, rather than by initial positive or negative experiences accessing records.

While our study did not identify appreciable harm, small but significant concerns about negative consequences of sharing records remain. In some instances, patient-provider communications about shared records created discomfort. It appears that all parties needed to adjust to a new dynamic, with patients having higher expectations of disclosure of information. Such issues were described in the *Open Notes* study [6], pointing to a need for professional education on clinical documentation and communication that optimizes patient participation and shared decision-making.

Some limitations of the study must be noted. Patients were recruited from a single medical center and viewed their records during a specific time interval, and therefore may not be generalizable to the VA patient population or My Health*e*Vet Pilot users overall. Study responders could have been more satisfied with the PHR or had more positive experiences overall with reading their health records. Second, many patients enrolled in the Pilot program yet did not view their records during the 18 months before the study, so they were not eligible for recruitment. Pilot enrollment occurred early in the 10 years it was active, and the study was conducted towards the end of this period. There were likely many patients who enrolled but chose not to access their records, had technical problems using the site, or viewed their records years before but did not to do so again. Further, release of the national My Health*e*Vet portal in 2003 likely caused Pilot enrollees to stop using the PHR prototype. Since our goal was to learn from patients who were able to recall using the Pilot, it was important to recruit those who accessed their health records relatively recently.

Study Implications

To our knowledge, this is the first qualitative study of veteran patients' experience viewing electronic records that included clinical notes and test results. Our findings have important implications for the development of electronic health records and PHRs. While patients by and large welcome full record access, clinicians reveal protective postures and worry about patient distress and confusion, resulting in more work for staff [19,21]. Our study, focusing on the patient experience, suggests that the first two of these expectations are overestimated. Concern about workload is likely more complex. Patients' accounts suggest that sharing all records reduce workload in some areas, for example, fewer visits or decreasing requests for copies of records. At the same time, participants' experiences also challenge traditional roles for patients and physicians. While patient-provider conversations may prove uncomfortable, they also demonstrate greater patient participation in care and contribution to care delivery. As evidence shows that activated patients achieve higher levels of self-care and satisfaction [27], sharing all clinical notes with patients and their delegates could serve as a fundamental component for the meaningful use of electronic records and health information exchange.

In this era of greater transparency and technology designed to optimize the user experience, new skills will be needed to achieve shared care planning and decision-making. Ultimately, patient access to all health record notes may translate into care that is more effective and more satisfying—for both patients and for health professionals.

Acknowledgments

The authors would like to thank Belinda Dalton and Dr. David Douglas for their help identifying eligible My Health*e*Vet Pilot enrollees at Portland VA Medical Center.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

SSW, NAP, and WPN participated in the conception and design of the study. SSW, ES, and AT contributed to the analysis and interpretation of data. SSW, ES, AT, NAP, KMN, and CLT contributed to drafting the article and/or revising it critically for important intellectual content. SSW, ES, and WPN conducted the final approval of the version to be published.

Conflicts of Interest

None declared.

References

1. Shenkin BN, Warner DC. Sounding board. Giving the patient his medical record: a proposal to improve the system. N Engl J Med 1973 Sep 27;289(13):688-692. [doi: 10.1056/NEJM197309272891311] [Medline: 4727972]

2. Baldry M, Cheal C, Fisher B, Gillett M, Huet V. Giving patients their own records in general practice: experience of patients and staff. Br Med J (Clin Res Ed) 1986 Mar 1;292(6520):596-598 [FREE Full text] [Medline: 3081187]
3. Fischbach RL, Sionelo-Bayog A, Needle A, Delbanco TL. The patient and practitioner as co-authors of the medical record. Patient Couns Health Educ 1980;2(1):1-5. [Medline: 10294999]
4. Fisher B, Britten N. Patient access to records: expectations of hospital doctors and experiences of cancer patients. Br J Gen Pract 1993 Feb;43(367):52-56 [FREE Full text] [Medline: 8466775]
5. Delbanco T, Walker J, Darer JD, Elmore JG, Feldman HJ, Leveille SG, et al. Open notes: doctors and patients signing on. Ann Intern Med 2010 Jul 20;153(2):121-125. [doi: 10.1059/0003-4819-153-2-201007200-00008] [Medline: 20643992]
6. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 2;157(7):461-470. [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]
7. Office for Civil Rights, HSS. Public Law, 104-191. 1996. Health Insurance Portability and Accountability Act of 1996 (HIPAA) URL: http://www.gpo.gov/fdsys/pkg/PLAW-104publ191/html/PLAW-104publ191.htm [accessed 2013-03-03] [WebCite Cache ID 6EqszCFss]
8. Fowles JB, Kind AC, Craft C, Kind EA, Mandel JL, Adlis S. Patients' interest in reading their medical record: relation with clinical and sociodemographic characteristics and patients' approach to health care. Arch Intern Med 2004 Apr 12;164(7):793-800. [doi: 10.1001/archinte.164.7.793] [Medline: 15078650]
9. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
10. Walker J, Ahern DK, Le LX, Delbanco T. Insights for internists: "I want the computer to know who I am". J Gen Intern Med 2009 Jun;24(6):727-732 [FREE Full text] [doi: 10.1007/s11606-009-0973-1] [Medline: 19412641]
11. 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]
12. Winkelman WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. J Am Med Inform Assoc 2005;12(3):306-314 [FREE Full text] [doi: 10.1197/jamia.M1712] [Medline: 15684128]
13. Ralston JD, Martin DP, Anderson ML, Fishman PA, Conrad DA, Larson EB, et al. Group health cooperative's transformation toward patient-centered access. Med Care Res Rev 2009 Dec;66(6):703-724. [doi: 10.1177/1077558709338486] [Medline: 19549993]
14. Kaiser. In: Liang L, editor. Connected for health: using electronic health records to transform care delivery. San Francisco: Jossey-Bass; 2010.
15. Reti SR, Feldman HJ, Ross SE, Safran C. Improving personal health records for patient-centered care. J Am Med Inform Assoc 2010;17(2):192-195 [FREE Full text] [doi: 10.1136/jamia.2009.000927] [Medline: 20190063]
16. Ralston JD, Hirsch IB, Hoath J, Mullen M, Cheadle A, Goldberg HI. Web-based collaborative care for type 2 diabetes: a pilot randomized trial. Diabetes Care 2009 Feb;32(2):234-239 [FREE Full text] [doi: 10.2337/dc08-1220] [Medline: 19017773]
17. Nazi KM. Veterans' voices: use of the American Customer Satisfaction Index (ACSI) Survey to identify My HealtheVet personal health record users' characteristics, needs, and preferences. J Am Med Inform Assoc 2010;17(2):203-211 [FREE Full text] [doi: 10.1136/jamia.2009.000240] [Medline: 20190065]
18. Zhou YY, Garrido T, Chin HL, Wiesenthal AM, Liang LL. Patient access to an electronic health record with secure messaging: impact on primary care utilization. Am J Manag Care 2007 Jul;13(7):418-424 [FREE Full text] [Medline: 17620037]
19. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc 2004;11(5):410-417 [FREE Full text] [doi: 10.1197/jamia.M1479] [Medline: 15187074]
20. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, et al. Effectiveness of home blood pressure monitoring, Web communication, and pharmacist care on hypertension control: a randomized controlled trial. JAMA 2008 Jun 25;299(24):2857-2867 [FREE Full text] [doi: 10.1001/jama.299.24.2857] [Medline: 18577730]
21. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819. [doi: 10.1059/0003-4819-155-12-201112200-00003] [Medline: 22184688]
22. Bhavnani V, Fisher B, Winfield M, Seed P. How patients use access to their electronic GP record--a quantitative study. Fam Pract 2011 Apr;28(2):188-194 [FREE Full text] [doi: 10.1093/fampra/cmq092] [Medline: 21084568]
23. Greenhalgh T, Wood GW, Bratan T, Stramer K, Hinder S. Patients' attitudes to the summary care record and HealthSpace: qualitative study. BMJ 2008 Jun 7;336(7656):1290-1295 [FREE Full text] [doi: 10.1136/bmj.a114] [Medline: 18511764]

24. Zulman DM, Nazi KM, Turvey CL, Wagner TH, Woods SS, An LC. Patient interest in sharing personal health record information: a web-based survey. Ann Intern Med 2011 Dec 20;155(12):805-810. [doi: 10.1059/0003-4819-155-12-201112200-00002] [Medline: 22184687]
25. 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]
26. Fisher B, Bhavnani V, Winfield M. How patients use access to their full health records: a qualitative study of patients in general practice. J R Soc Med 2009 Dec;102(12):539-544 [FREE Full text] [doi: 10.1258/jrsm.2009.090328] [Medline: 19966130]
27. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med 2012 May;27(5):520-526. [doi: 10.1007/s11606-011-1931-2] [Medline: 22127797]

Abbreviations

PHR: personal health record
VA: US Department of Veterans Affairs

Edited by G Eysenbach; submitted 18.09.12; peer-reviewed by S Ross, J Walker; comments to author 09.11.12; revised version received 24.01.13; accepted 06.02.13; published 27.03.13.

Please cite as:
Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, Nichol WP
Patient Experiences With Full Electronic Access to Health Records and Clinical Notes Through the My HealtheVet Personal Health Record Pilot: Qualitative Study
J Med Internet Res 2013;15(3):e65
URL: http://www.jmir.org/2013/3/e65/
doi:10.2196/jmir.2356
PMID:23535584

Original Paper

Comparison of Web-Based and Paper-Based Administration of ADHD Questionnaires for Adults

Oliver Hirsch[1], PhD; Franziska Hauschild[1], MD; Martin H Schmidt[2], MSc; Erika Baum[1], MD; Hanna Christiansen[2], PhD

[1]Faculty of Medicine, Department of General Practice/Family Medicine, Philipps University Marburg, Marburg, Germany

[2]Faculty of Psychology, Department of Clinical Psychology, Philipps University Marburg, Marburg, Germany

Corresponding Author:
Oliver Hirsch, PhD
Faculty of Medicine
Department of General Practice/Family Medicine
Philipps University Marburg
Karl-von-Frisch-Str.4
Marburg, 35043
Germany
Phone: 49 64212826520
Fax: 49 64212865121
Email: oliver.hirsch@staff.uni-marburg.de

Abstract

Background: Satisfactory psychometric properties in offline questionnaires do not guarantee the same outcome in Web-based versions. Any construct that is measured online should be compared to a paper-based assessment so that the appropriateness of online questionnaire data can be tested. Little research has been done in this area regarding Attention-Deficit/Hyperactivity Disorder (ADHD) in adults.

Objective: The objective was to simultaneously collect paper-based and Web-based ADHD questionnaire data in adults not diagnosed with ADHD in order to compare the two data sources regarding their equivalence in raw scores, in measures of reliability, and in factorial structures.

Methods: Data from the German versions of the Connors Adult ADHD Rating Scales (CAARS-S), the Wender Utah Rating Scale (WURS-k), and the ADHD Self Rating Scale (ADHS-SB) were collected via online and paper questionnaires in a cross-sectional study with convenience sampling. We performed confirmatory factor analyses to examine the postulated factor structures in both groups separately and multiple group confirmatory factor analyses to test whether the postulated factor structures of the questionnaires were equivalent across groups. With Cronbach alpha, we investigated the internal consistency of the postulated factors in the different questionnaires. Mann-Whitney U tests with the effect size “Probability of Superiority (PS)” were used to compare absolute values in the questionnaires between the two groups.

Results: In the paper-based sample, there were 311 subjects (73.3% female); in the online sample, we reached 255 subjects (69% female). The paper-based sample had a mean age of 39.2 years (SD 18.6); the Web-based sample had a mean age of 30.4 years (SD 10.5) and had a higher educational background. The original four factor structure of the CAARS-S could be replicated in both samples, but factor loadings were different. The Web-based sample had significantly higher total scores on three scales. The five-factor structure of the German short form of the WURS-k could be replicated only in the Web-based sample. The Web-based sample had substantially higher total scores, and nearly 40% of the Web-based sample scored above the clinically relevant cut-off value. The three-factor structure of the ADHS-SB could be replicated in both samples, but factor loadings were different. Women in the Web-based sample had substantially higher total scores, and 30% of the Web-based sample scored above the clinically relevant cut-off value. Internal consistencies in all questionnaires were acceptable to high in both groups.

Conclusions: Data from the Web-based administration of ADHD questionnaires for adults should not be used for the extraction of population norms. Separate norms should be established for ADHD online questionnaires. General psychometric properties of ADHD questionnaires (factor structure, internal consistency) were largely unaffected by sampling bias. Extended validity studies of existing ADHD questionnaires should be performed by including subjects with a diagnosis of ADHD and by randomizing them to Web- or paper-based administration.

(J Med Internet Res 2013;15(3):e47) doi:10.2196/jmir.2225

KEYWORDS

computers; Attention-Deficit/Hyperactivity Disorder; questionnaires; Internet; psychometrics

Introduction

Satisfactory psychometric properties in offline questionnaires do not guarantee the same outcome in Web-based versions. Any construct that is measured online should be compared to a paper-based assessment so that the appropriateness of online questionnaire data can be tested [1]. After analyzing common preconceptions about Internet questionnaires, Gosling et al [2] conclude that the quality of online data is comparable to traditional paper-and-pencil methods. The authors argue that Internet samples are not representative of the general population, but that traditional methods also do not achieve this. Web-based questionnaires are even considered to be more feasible in order to collect data in large population-based epidemiological studies [3].

Several studies did not find substantial differences between Web-based and paper-based modes of administration [4-7]. They were able to show similar psychometric properties or identical factor structures [8-10]. Those latter studies did not use clinical questionnaires to assess psychopathology, but no significant differences were found when depression questionnaires were administered in paper and online versions among the same individuals [11]. Comparable results regarding psychometric properties and absolute differences in Internet versus paper-and-pencil administration of several panic and agoraphobia questionnaires were found by Carlbring et al [12]. Both studies included pre-selected or self-recruited patient groups applying for treatment. A Web Screening Questionnaire for mental disorders yielded a high number of false positives though [13], while others reported satisfactory diagnostic accuracy in the Web-based detection of depressive disorders [14].

Buchanan [15] is skeptical of using Web-based questionnaires for normative comparisons, especially in clinical psychology. Several studies showed that score distributions between paper and online administration differed, with higher scores in Internet samples [16]. He therefore argued not to compare online questionnaire data to established norms.

Attention-Deficit/Hyperactivity Disorder (ADHD), with its core symptoms of inattention, hyperactivity, and impulsivity, is listed under disorders usually first diagnosed in childhood or adolescence in DSM-IV and ICD-10. It was shown that ADHD often persists into adulthood with prevalence rates between 4 to 5% [17-19]. We found only two studies in which measurement of ADHD via Web-based versions was examined. Steenhuis et al [20] applied a within-subject design to administer the ADHD section of the Diagnostic Interview Schedule for Children (DISC-IV) to parents. Intraclass correlation coefficients ranged between .87 and .94. A qualitative study examined acceptability of the Web-based version of the ADHD rating scale T-SKAMP in 19 teachers [21]. A large majority of teachers preferred the Web-based version over a paper version. They perceived it to be easier, shorter, simpler, and more informative, time saving, and flexible. Communication between teachers and physicians might be improved with this tool. No further ADHD diagnostic instruments, such as the SNAP and SWAN Scale for children [22] or common adult ADHD assessment instruments (see [23] for a review) were implemented as Web-based versions.

The Conners Adult ADHD Rating Scales (CAARS) [24] had satisfactory psychometric properties in their German translation [25-27] and were found to have the same factor structure as the American original, enabling them to be used for cross-cultural research. The aim of our study was to simultaneously collect paper-based and Web-based CAARS questionnaire data together with two other established ADHD questionnaires available in German and to compare the two data sources by different statistical measures regarding their equivalence in raw scores, measures of reliability, and factorial structures. To do so, we intended to collect normative data online and via paper questionnaires from subjects without a diagnosis of ADHD in order to examine whether online normative data can be merged with data from paper questionnaires.

Methods

Recruitment

We conducted a cross-sectional study on German adults with no serious chronic disease, who were over 18 years of age and without a lifetime diagnosis of ADHD. Participants in the paper-based sample were recruited by convenience sampling (university students, people from apprentice institutions, local neighborhoods, waiting areas such as airports, hairdressers, primary care physicians, and colleagues). Subjects were provided with a short study description and asked to complete the CAARS self-report (CAARS-S) as well as the German version of the Wender Utah Rating Scale (WURS-k), the German ADHD Self Rating Scale (ADHS-SB), and questions on age, gender, and education level. We disseminated approximately 500 printed questionnaires.

The Web-based questionnaire was also a cross-sectional convenience sample. We advertised for the online study on the websites of the Departments of General Practice/Family Medicine and Clinical Psychology at Philipps University Marburg and on a special Facebook page created exclusively for our study. Additionally, flyers with the online address of the Web-based questionnaire were distributed in the same recruitment areas as the paper-based questionnaires. For informed consent in the online study, the homepage prompted subjects to open a file with the study information and to check a box agreeing to participate in the study. Without checking this box, further pages of the questionnaire were not accessible. Since the survey was voluntary, all subjects had the ability to discontinue completing the questionnaire at any time. Subjects could see their progress in completing the questionnaire via a

small progress bar on the upper right side of the screen. On average, subjects needed 15:34 minutes to complete the survey; the majority of participants completed the survey during the afternoon (hour 15, or 3 p.m.). At the end of the questionnaire, subjects had the opportunity to receive feedback on their responses. This indicated whether their scores were within the normal range or higher. In cases of the latter, no diagnosis was offered, but it was suggested to seek professional assessment. Data protection was insured in that only the principal investigator (HC) had access to the unipark page [28] that generated and stored the data. Additionally, no personal information was requested of the subjects.

For development and testing, the paper versions of the questionnaires were entered into unipark. The research team and then students were asked to test this online version. The link was activated after testing for functionality and usability.

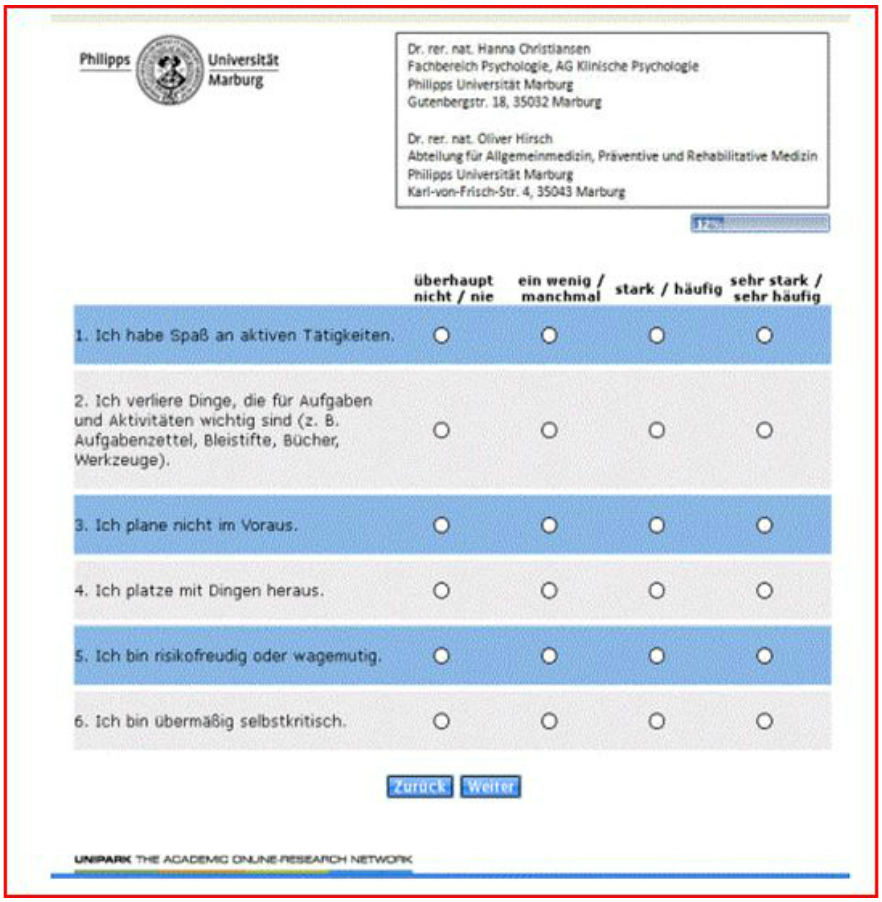
The survey was online from July 12, 2010, to August 30, 2011. On average, the page was accessed 26.16 times per week (view rate), though only 6.69 (25 %) subjects per week completed the survey (completion rate). Cookies were used to assign a unique user identifier to each client computer and were set on the first page. A session was valid for a total of 120 minutes.

The items were presented in the same order as the paper-and-pencil questionnaires, but only an average of six items were displayed per page (see Figure 1). When subjects did not complete all items on a page, they were asked to fill in the missing items in order to activate the "next" button. Therefore, no missing data could result in the online sample. Subjects were always able to review and change their answers with a "back" button. If subjects decided not to complete an answer, they could stop the survey. The majority of participants (57.44 %) discontinued filling out the survey on the first page.

For analyses, only questionnaires where subjects indicated they had not received a lifetime diagnosis of ADHD were analyzed. Apart from replacing missing items in paper versions with the expectation-maximization or the multiple imputation algorithms, no statistical corrections were performed.

Our study conforms to the Declaration of Helsinki and was approved by the local ethics committee of the Faculty of Medicine at the Philipps University in Marburg, Germany.

Figure 1. Example of survey items per page.



Measurements

Connors Adult ADHD Rating Scales

The German version of the CAARS-S assesses ADHD symptoms in adults aged 18 years or older. Symptoms are rated on a Likert-type scale (0 = *not at all/never* to 3 = *very much/very frequently*). The long version consists of 66 items, but only 42 items were included in the original factor analysis by Conners et al [24] due to statistical restrictions made by the authors. Four factors emerged from their analyses: inattention/memory problems, hyperactivity/restlessness, impulsivity/emotional ability, and problems with self-concept. Confirmatory factor analyses of the German version in healthy adults and ADHD patients supported this factor analytic solution [25,27]. The four

subscales were significantly influenced by age, gender, and the number of years in education. Symptom severity decreased with increasing age, males scored higher than females on hyperactivity and sensation-seeking behavior, and females scored higher than males on problems with self-concept. Overall symptom ratings were higher for individuals who had received less education. Test-retest reliability ranged between .85 and .92; sensitivity and specificity were high for all four subscales. The CAARS-S represents a reliable and cross-culturally valid measure of current ADHD symptoms in adults [26].

Wender Utah Rating Scale

The German version of the Wender Utah Rating Scale (WURS-k) [29,30] retrospectively assesses ADHD-relevant childhood behaviors and symptoms in adults. It consists of 25 items that distinguished patients with ADHD from a nonpatient comparison group. Subjects are instructed to rate 25 items that complete sentence stems such as ‘‘As a child I was or had...’’. Ratings are to be completed on a 5-point Likert scale (0 =*not at all or very slightly* to 4=*very much*). Test-retest reliability and Cronbach alpha were around .90. Factor analyses generated a 5-factor solution with the factors inattention/hyperactivity, impulsivity, anxiety/depression, oppositional behavior, and social adaptation by using 21 items. A total score >29 points hints at the possibility of ADHD during childhood.

ADHD Self Rating Scale

The German ADHD Self Rating Scale (ADHS-SB) consists of the 18 DSM-IV items that are broken down into the factors “inattention” (9 items), “hyperactivity”, and “impulsivity” (9 items together) [31]. The items are scored on a 4-point Likert scale (0=*not at all* to 3=*very pronounced/almost always the case*). Test-retest reliability coefficients were between .78 and .89. Correlations with subscales of the NEO Five Factor Inventory were in the expected directions. A total score >17 points hints at the possibility of adult ADHD.

Statistical Analysis

We performed confirmatory factor analyses to examine the postulated factor structures in both groups separately and multiple group confirmatory factor analyses using AMOS 19 to test whether the postulated factor structures of the questionnaires were equivalent across the groups. The factors were allowed to correlate because this is theoretically plausible in all three questionnaires. We used unweighted least squares as this estimation method makes no distributional assumptions [32].

Using multiple group analysis, we examined several levels of invariance between the groups. Configural invariance as the lowest level of invariance exists when the structure of the factor loading matrices is identical in all groups. Metric invariance occurs when factor loadings are identical in all groups. Scale invariance means that the measurement intercepts are the same across groups. Invariance of measurement errors exists if the error variables of measurement models, factor covariances, and factor variances are identical across groups.

We calculated several model fit indices to evaluate the results of our analyses. The root mean square residual (RMR) measures the mean absolute value of the covariance residuals [33]. Values less than .05 indicate a good model fit [32], but other authors state that a value of less than .10 signals an acceptable model fit [34-35]. The standardized root mean square residual (SRMR) eliminates scaling effects of the RMR. Values ≤ .10 indicate a good model fit [35]. The Global Fit Index (GFI) can measure the proportion of variance and covariance that a given model is able to explain. A GFI equal or higher than .90 can be considered as reflecting a good model fit [36]. The adjusted global fit index (AGFI) takes the number of parameters used in computing the GFI into account. An AGFI equal or higher than .90 can be considered as showing a good model fit [35]. These fit indices were calculated for each of the aforementioned invariance levels. Differences of fit indices between these invariance levels should not be larger than .01 [35], otherwise the criteria for a higher invariance level are not reached.

With Cronbach alpha, we investigated the internal consistency of the postulated factors in the different questionnaires. Values >.70 are considered to be acceptable [37].

Huber’s M estimators were calculated when standard deviation values were close to their respective means, signaling high variance [38].

Mann-Whitney U tests were used to compare absolute values in the questionnaires between the two groups. The effect size “Probability of Superiority (PS)”, PS=U/(n1*n2), indicates the probability that a randomly selected subject of group n1 has a higher score than a randomly selected subject of group n2. A PS of .50 means that both groups are equal regarding a specific variable, and that there is no effect. Consequently, the larger the effect, the more PS deviates from .50 [39].

The alpha level for statistical significance was set at .05 (two-sided). Missing responses in the paper versions were replaced using the expectation-maximization or the multiple imputation algorithms [40-42].

Results

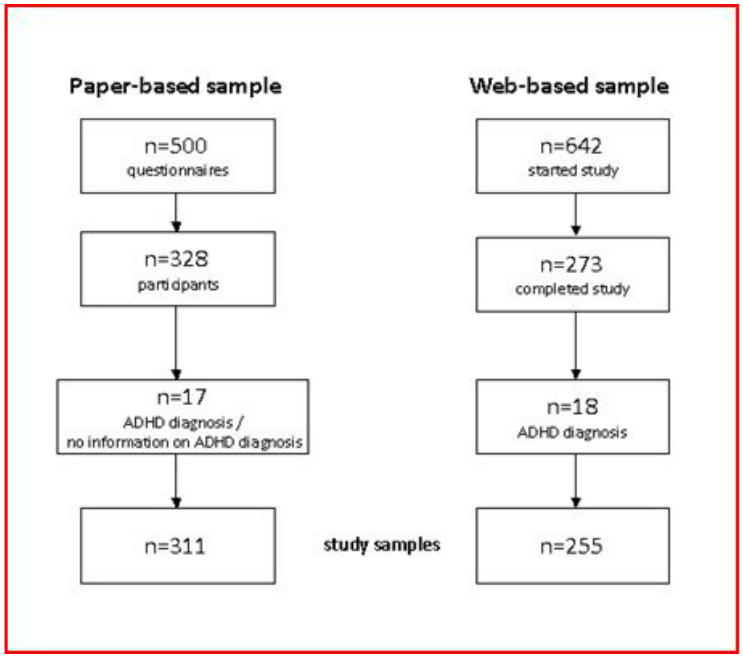
Samples

In the paper-based sample, we received responses from 328 participants of which 6 indicated they were diagnosed with ADHD, and 11 did not answer this question. Therefore, a total sample of 311 subjects resulted, meaning that 65.6% of our 500 printed questionnaires were returned. This cannot be regarded as a return rate as we did not record those subjects who were personally asked and refused to participate. In the Web-based sample, we received responses from 273 participants of which 18 indicated that they were diagnosed with ADHD so that a total sample of 255 subjects resulted. The flow of subjects in our study samples is depicted in Figure 2. Table 1 shows the demographic characteristics of the two samples.

Table 1. Demographic characteristics of the paper-based and Web-based samples.

		Paper-based (n=311)	Web-based (n=255)
Gender			
	Female	228 (73.3%)	176 (69.0%)
	Male	83 (26.7%)	79 (31.0%)
Age, mean years (SD)		39.2 (18.6)	30.4 (10.5)
Education			
	University	61 (19.7%)	73 (28.6%)
	Apprenticeship	86 (27.7%)	35 (13.7%)
	High school	66 (21.3%)	130 (51.0%)
	Middle school	61 (19.7%)	15 (5.9%)
	Basic school	36 (11.6%)	2 (0.8%)

Figure 2. Flow of subjects in the paper-based and Web-based samples.



The samples did not differ with respect to gender (χ^2 test: P=.26, Cramer V=.05). The Web-based sample was, on average, younger than the paper-based sample. This difference was statistically significant with a rather moderate effect size (Mann-Whitney U test: P<.001, PS = .41). In the Web-based sample, there were more participants with a university degree and more subjects attending high school, while in the paper-based sample, there were more participants attending middle or basic school or with a completed apprenticeship. This difference was statistically significant with a high effect size (χ^2 test: P<.001, Cramer V=.42).

Connors Adult ADHD Rating Scales

There was a maximum of 9% missing values on single variables in the paper sample; these were missing completely at random (Little's MCAR test, P=.27). They were replaced with the expectation maximization (EM) algorithm [40].

The four-factor model (df = 813) was supported in both groups. In the paper-based sample, the standardized RMR was .08, the RMR was .04, the GFI was .93, and the AGFI was .92. In the Web-based sample, the standardized RMR was .07, the RMR was .05, the GFI was .98, and the AGFI was .97. These fit indices signal a good model fit. Table 2 lists the correlations of CAARS items (loadings) with their postulated factors.

Except for Item 3 ("I don't plan ahead") of the factor "inattention/memory", Items 1 ("I like to be doing active things") and 5 ("I am a risk-taker or a daredevil") of the hyperactivity factor, and Item 43 ("I step on people's toes without meaning to") on the impulsivity factor, all other items have loadings > .40 in both samples.

The intercorrelations between the factors are consistently higher in the Web-based sample. The largest differences between the two groups were found in correlations involving "self-concept" (see Table 3).

Multiple group analysis revealed that the factor structures were the same in both samples, signaling configural invariance (SRMR=.04, RMR=.04, GFI=.99, AGFI=.99). However, factor

loadings were different (SRMR=.06, RMR=.10, GFI=.97, AGFI=.97) because all model fit indices increased > .01 when testing metric invariance. Consequently, other invariance assumptions were also not supported.

Cronbach alpha of the subscales ranged from .81 to .85 in the paper-based sample and from .89 to .91 in the Web-based sample.

Absolute subscale differences between the two groups were all significant with the Web-based sample scoring substantially higher with pronounced effect sizes (Table 4). For example, the probability that a randomly chosen subject from the paper-based sample has a higher inattention/memory score than a randomly chosen subject from the Web-based sample is .35.

As there is no normative data for Germany to date, we applied strict cut-off values based on American normative data T-value of 65, 94th percentile). The cut-off for "inattention/memory" was > 22 points; "hyperactivity" > 26 points; "impulsivity" > 22 points; and "self-concept" > 13 points. Regarding the total score of "inattention/memory", 27 subjects (10.6%) in the Web-based sample scored above this value while 4 (1.3%) did so in the paper-based sample. This difference was significant with a moderate effect size (χ^2 test: P<.001, Cramer V=.20). Regarding the total score of "hyperactivity", 5 subjects (2.0%) in the Web-based sample scored above this value while 2 (0.6%) did so in the paper-based sample. This difference was not significant (χ^2 test: P=.16, Cramer V=.06). Regarding the total score of "impulsivity", 20 subjects (7.8%) in the Web-based sample scored above this value while 6 (1.9%) did so in the paper-based sample. This difference was significant with a small effect size (χ^2 test: P=.001, Cramer V=.14). Regarding the total score of "self-concept", 30 (11.8%) in the Web-based sample scored above this value while 13 (4.2%) did so in the paper-based sample. This difference was significant with a small effect size (χ^2 test: P=.001, Cramer V=.14).

Wender Utah Rating Scale (Short Form)

There was a maximum of 2% missing values on single variables in the paper sample that were missing completely at random (Little's MCAR test, P=.57). These were replaced with the EM algorithm [40]. Due to technical difficulties, WURS-k data of 11 participants in the Web-based sample were not available, resulting in n=244.

The model in the paper-based sample was not admissible because the covariance matrix between the postulated five factors was not positive definite. This leads to the conclusion that the model is wrong [34], and it was thus rejected.

In the Web-based sample, the model (df = 179) was supported: SRMR = .07, RMR = .09, GFI = .98, AGFI = .97. Table 5 depicts the loadings on the postulated factors.

As shown in Table 5, except for Item 23 (problems with police) on the factor "social adaptation", all other items have high loadings > .4 on their postulated factors. The factors "inattention" and "impulsivity" correlated highest in the Web-based sample (r=.79), followed by "inattention" and "anxiety/depression" (r=.73), "impulsivity" and "oppositional behavior" (r=.72), and "inattention" and "oppositional behavior" (r=.72) (Table 6).

Due to the rejected model in the paper-based group, no multiple group analysis could be calculated.

Cronbach alpha of the subscales ranged from .68 to .82 in the paper-based sample and from .79 to .89 in the Web-based sample. No coefficients were calculated for the subscale "social adaptation" as it consists of only two items.

Absolute values of the total score were significantly higher (Mann-Whitney U test: P<.001; PS = .09) in the Web-based sample (mean 28.6, SD 14.0; Huber's M estimator 26.2) than in the paper-based sample (mean 11.0, SD 6.8; Huber's M estimator 9.6).

After applying the recommended cut-off value for the total score (> 29 points) [29,30], 38.1% in the Web-based sample scored above this value while merely 2.5% did so in the paper-based sample. This difference was significant with a high effect size (χ^2 test: P<.001, Cramer V=.46).

ADHD Self Rating Scale

There was a maximum of 1.3% missing values on single variables in the paper sample, except for Item 4 that asks for difficulties in the field of work. Student participants in the paper version did not complete this item, so 28.5% of missing at random data resulted. These were replaced with the multiple imputation algorithm by five imputations. The following calculations were done separately for the five imputations, and the respective results were averaged. Enders [40] recommends a larger number of imputations, but results showed only marginal differences between imputed datasets.

The four-factor model (df = 132) was supported in both groups. In the paper-based sample, the standardized RMR was .06, the RMR was .06, the GFI was .97, and the AGFI was .96. In the Web-based sample, the standardized RMR was .06, the RMR was .04, the GFI was .98, and the AGFI was .98. These fit indices signal a good model fit. Table 7 lists the loadings of ADHS-SB items on their postulated factors.

As shown in Table 7, except for Item 6 (avoidance of tasks with mental load) on the factor "inattention", and Item 14 (feel like driven by a motor) on the factor "hyperactivity", both in the paper-based sample, all other items have high loadings >.4 on their postulated factors.

The correlation between the factors inattention and hyperactivity is significantly higher in the paper-based sample, while the intercorrelations between the other factors are higher in the Web-based sample (Table 8).

Multiple group analysis revealed that the factor structures were the same in both samples, signaling configural invariance (SRMR = .06, RMR = .03, GFI = .98, AGFI = .98). However, factor loadings were different (SRMR = .15, RMR = .08, GFI = .83, AGFI = .79) because all model fit indices increased > .01 when testing metric invariance. Consequently, other invariance assumptions were also not supported.

Cronbach alpha of the subscales ranged from .60 to .83 in the paper-based sample and from .79 to .91 in the Web-based sample.

Absolute differences between the two groups were significant with the Web-based sample (mean 12.8, SD 9.1; Huber's M estimator 11.1) scoring substantially higher than the paper-based sample (mean 2.2, SD 3.0; Huber's M estimator 1.4) with a high effect size (Mann-Whitney U test: $P<.001$, PS = .08).

After applying the recommended cut-off value for the total score (> 17 points) [31], 30% in the Web-based sample scored above this value, while only 0.6% did so in the paper-based sample. This difference was significant with a large effect size (χ^2 test: $P<.001$, Cramer V=.42).

Table 2. Correlations of CAARS items (loadings) with their postulated factors (latent constructs) in the paper-based and Web-based samples.

		Paper-based	Web-based
Inattention/Memory			
	ITEM 03	.14	.26
	ITEM 07	.53	.74
	ITEM 11	.59	.73
	ITEM 16	.47	.67
	ITEM 18	.55	.74
	ITEM 32	.57	.61
	ITEM 36	.69	.75
	ITEM 40	.52	.73
	ITEM 44	.55	.78
	ITEM 49	.55	.73
	ITEM 51	.55	.63
	ITEM 66	.57	.74
Hyperactivity			
	ITEM 01	.31	.06
	ITEM 05	.43	.32
	ITEM 10	.46	.50
	ITEM 13	.70	.74
	ITEM 20	.65	.69
	ITEM 25	.57	.46
	ITEM 27	.57	.82
	ITEM 31	.65	.66
	ITEM 38	.54	.76
	ITEM 46	.67	.80
	ITEM 54	.61	.71
	ITEM 57	.73	.81
Impulsivity			
	ITEM 04	.56	.62
	ITEM 08	.49	.72
	ITEM 12	.56	.69
	ITEM 19	.66	.58
	ITEM 23	.58	.61
	ITEM 30	.64	.76
	ITEM 35	.47	.54
	ITEM 39	.60	.67
	ITEM 43	.37	.59
	ITEM 47	.59	.78
	ITEM 52	.53	.62
	ITEM 61	.61	.69
Self-concept			
	ITEM 06	.59	.58
	ITEM 15	.60	.75

		Paper-based	Web-based
	ITEM 26	.58	.69
	ITEM 37	.81	.86
	ITEM 56	.75	.79
	ITEM 63	.81	.84

Table 3. Intercorrelations between the CAARS factors (latent constructs) in the paper-based and Web-based samples.

Factors			Paper-based	Web-based
Hyperactivity	<-->	Impulsivity	.73	.81
Inattention/Memory	<-->	Hyperactivity	.54	.74
Inattention/Memory	<-->	Impulsivity	.65	.79
Inattention/Memory	<-->	Self-concept	.47	.74
Hyperactivity	<-->	Self-concept	.24	.57
Impulsivity	<-->	Self-concept	.45	.71

Table 4. Means, standard deviations, and Huber's M estimators of the CAARS subscales in the paper-based and Web-based samples with their respective *P* and effect size values.

	Paper-based	Web-based	Mann-Whitney U Test (*P*) & effect size (PS[b])
Inattention/Memory	8.6 (SD 4.8) Huber's M[a] 8.2	12.3 (SD 7.1) Huber's M 11.2	<.001; PS=.35
Hyperactivity	9.0 (SD 5.3) Huber's M 8.0	11.2 (SD 6.1) Huber's M 10.3	<.001; PS=.38
Impulsivity	9.4 (SD 5.2) Huber's M 8.8	12.3 (SD 6.6) Huber's M 11.5	<.001; PS=.37
Self-concept	5.6 (SD 3.6) Huber's M 5.1	7.5 (SD 4.3) Huber's M 7.1	<.001; PS=.37

[a]Huber's M estimator.

[b]PS = probability of superiority.

Table 5. Correlations of WURS-k items (loadings) with their postulated factors (latent constructs) in the Web-based sample.

		Web-based
Inattention		
	ITEM 01	.82
	ITEM 02	.76
	ITEM 03	.78
	ITEM 06	.77
	ITEM 10	.75
	ITEM 15	.59
	ITEM 17	.67
	ITEM 24	.51
Impulsivity		
	ITEM 05	.75
	ITEM 11	.83
	ITEM 13	.83
	ITEM 16	.89
Anxiety/Depression		
	ITEM 07	.74
	ITEM 09	.62
	ITEM 18	.74
	ITEM 19	.80
Oppositional behavior		
	ITEM 08	.88
	ITEM 21	.54
	ITEM 22	.79
Social adaptation		
	ITEM 20	.64
	ITEM 23	.33

Table 6. Intercorrelations between the WURS-k factors (latent constructs) in the Web-based sample.

Factors			Web-based
Inattention	<-->	Impulsivity	.79
Impulsivity	<-->	Anxiety/Depression	.69
Inattention	<-->	Anxiety/Depression	.73
Inattention	<-->	Social adaptation	.50
Impulsivity	<-->	Oppositional behavior	.72
Impulsivity	<-->	Social adaptation	.49
Anxiety/Depression	<-->	Oppositional behavior	.35
Anxiety/Depression	<-->	Social adaptation	.57
Oppositional behavior	<-->	Social adaptation	.58
Inattention	<-->	Oppositional behavior	.72

Table 7. Correlations of ADHS-SB items (loadings) with their postulated factors (latent constructs) in the paper-based and Web-based samples.

		Paper-based	Web-based
Inattention			
	ITEM 01	.47	.72
	ITEM 02	.48	.73
	ITEM 03	.56	.68
	ITEM 04	.40	.65
	ITEM 05	.45	.63
	ITEM 06	.22	.62
	ITEM 07	.48	.54
	ITEM 08	.60	.72
	ITEM 09	.45	.64
Hyperactivity			
	ITEM 10	.71	.74
	ITEM 11	.66	.69
	ITEM 12	.58	.82
	ITEM 13	.54	.65
	ITEM 14	.31	.59
Impulsivity			
	ITEM 15	.68	.72
	ITEM 16	.49	.72
	ITEM 17	.48	.73
	ITEM 18	.58	.58

Table 8. Intercorrelations between the ADHS-SB factors (latent constructs) in the paper-based and Web-based samples.

Factors			Paper-based	Web-based
Inattention	<-->	Hyperactivity	.92	.63
Hyperactivity	<-->	Impulsivity	.64	.80
Inattention	<-->	Impulsivity	.66	.72

Discussion

We compared Web-based and paper-based administrations of three ADHD questionnaires for adults. Subjects in the online sample were older and had a higher educational background. The original four-factor structure of the Conners Adult ADHD Rating Scales could be replicated in both samples, but factor loadings were different. Internal consistencies were high in both groups, but the Web-based sample had significantly higher total scores in three subscales with 7.8 to 11.8% above clinically relevant cut-off values, compared to 1.3 to 4.2% in the paper-based sample. The five-factor structure of the German short form of the Wender Utah Rating Scale could be replicated only in the Web-based sample. Internal consistencies were acceptable to high in both groups. The Web-based sample had substantially higher total scores and nearly 40% of the Web-based sample scored above the clinically relevant cut-off value. The three-factor structure of the ADHD Self Rating Scale could be replicated in both samples, but factor loadings were different. Internal consistencies were acceptable to high in both groups. The Web-based sample had substantially higher total scores, and 30% of the Web-based sample scored above the clinically relevant cut-off value. Therefore, psychometric properties were similar in both samples, but the Web-based sample had substantially higher scores on all three questionnaires.

The relatively high dropout rate in our Web-based sample is also reported in the literature. Additional informed consent procedures were shown to increase early dropout in Web-based studies [43]. Kongsved et al [44] administered paper and online questionnaires to women referred for mammography; their questionnaires were comparable in length to our study. In their study, the Internet version had a higher completeness of data but a lower response rate. A lower response rate was also found in surgeons responding to an online questionnaire [45]. We cannot compare dropout rates in our samples because we did not record those who were personally asked and refused to participate in the paper-based sample.

Demographic differences (younger age, higher education) might have influenced the results [46]. ADHD is a disorder with a higher prevalence for men. Our predominantly female samples in both versions are therefore not suitable for deriving normative data. Women tend to participate more in psychological studies. The gender distributions in our study are quite comparable to common sample characteristics regarding online study participation [2,47]. Subjects in the Internet sample might have experienced psychological distress regarding ADHD symptoms so that they saw their participation in the context of assessing themselves for the disorder. One also has to consider that the probability that subjects in a certain geographic region have a high prevalence of ADHD symptoms in paper-based questionnaires is much lower than the probability of reaching such individuals via the Internet without any geographic barriers. On the other hand, an increased self-disclosure might also be an important factor. Subjects might have considered the completion of online questionnaires to be more anonymous than giving away hand-written information, since participants in an online study reported lower social anxiety and lower social desirability than those in the paper-based group [48].

Our results contradict the conclusion of Gosling et al [2] that subjects in Internet samples are not unusually maladjusted. It clearly depends on the context of the study. Even when the intention is to collect normative data for clinical questionnaires, one has to consider that the scores of online samples can be inflated [49].

On the other hand, our results corroborate the assumption of Rhodes et al [47] that previously hidden subgroups can be reached by Internet research. This might also be true for other medical disorders [50]. Whether our subgroup of women with higher scores on ADHD questionnaires has clinical significance must be determined by future studies with more controlled recruitment strategies.

Several limitations have to be mentioned. We did not randomize subjects to online and paper versions, so differences between the two groups might have arisen by sampling biases and should be replicated under randomized conditions. Different recruitment strategies for the paper and online samples might have influenced the results. Although relatively high discontinuation rates are common in online research, they might have caused bias in the results. In future online studies, leaving out questions should also be possible to create conditions similar to paper administration.

Conclusions

Data from the Web-based administration of ADHD questionnaires for adults should not be used for the extraction of population norms. Separate norms should be established for ADHD online questionnaires. General psychometric properties of ADHD questionnaires (factor structure, internal consistency) were largely unaffected by sampling bias. Extended validity studies of existing ADHD questionnaires should be performed by including subjects with a diagnosis of ADHD and by randomizing them to Web- or paper-based administration.

Acknowledgments

This study received no financial grants.

Conflicts of Interest

None declared.

References

1. Strickland OL, Moloney MF, Dietrich AS, Myerburg S, Cotsonis GA, Johnson RV. Measurement issues related to data collection on the World Wide Web. ANS Adv Nurs Sci 2003;26(4):246-256. [Medline: 14674574]
2. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust Web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004 Mar;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]
3. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using Web-based questionnaires in large population-based epidemiological studies. Eur J Epidemiol 2006;21(2):103-111. [doi: 10.1007/s10654-005-6030-4] [Medline: 16518678]
4. 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]
5. Van De Looij-Jansen PM, De Wilde EJ. Comparison of Web-based versus paper-and-pencil self-administered questionnaire: effects on health indicators in Dutch adolescents. Health Serv Res 2008 Oct;43(5 Pt 1):1708-1721 [FREE Full text] [doi: 10.1111/j.1475-6773.2008.00860.x] [Medline: 18479404]
6. Denscombe M. Web-based questionnaires and the mode effect. Social Science Computer Review 2006;24(2):246-254. [doi: 10.1177/0894439305284522]
7. Hardré PL, Crowson HM, Xie K, Ly C. Testing differential effects of computer-based, Web-based and paper-based administration of questionnaire research instruments. Br J Educ Technol 2007 Jan 2007;38(1):5-22. [doi: 10.1111/j.1467-8535.2006.00591.x]
8. Vallejo MA, Jordán CM, Díaz MI, Comeche MI, Ortega J. Psychological assessment via the internet: a reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms

Check-List-90-Revised (SCL-90-R). J Med Internet Res 2007;9(1):e2 [FREE Full text] [doi: 10.2196/jmir.9.1.e2] [Medline: 17478411]
9. Hardré PL, Crowson HM, Xie K. Differential Effects of Web-Based and Paper-Based Administration of Questionnaire Research Instruments in Authentic Contexts-of-Use. Journal of Educational Computing Research 2010 Jan 2010;42(1):103-133. [doi: 10.2190/EC.42.1.e]
10. Riva G, Teruzzi T, Anolli L. The use of the internet in psychological research: comparison of online and offline questionnaires. Cyberpsychol Behav 2003 Feb;6(1):73-80. [doi: 10.1089/109493103321167983] [Medline: 12650565]
11. Holländare F, Andersson G, Engström I. A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. J Med Internet Res 2010;12(5):e49 [FREE Full text] [doi: 10.2196/jmir.1392] [Medline: 21169165]
12. Carlbring P, Brunt S, Bohman S, Austin D, Richards J, Öst L, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. Computers in Human Behavior 2007 May 2007;23(3):1421-1434. [doi: 10.1016/j.chb.2005.05.002]
13. Donker T, van Straten A, Marks I, Cuijpers P. A brief Web-based screening questionnaire for common mental disorders: development and validation. J Med Internet Res 2009;11(3):e19 [FREE Full text] [doi: 10.2196/jmir.1134] [Medline: 19632977]
14. Lin CC, Bai YM, Liu CY, Hsiao MC, Chen JY, Tsai SJ, et al. Web-based tools can be used reliably to detect patients with major depressive disorder and subsyndromal depressive symptoms. BMC Psychiatry 2007;7:12 [FREE Full text] [doi: 10.1186/1471-244X-7-12] [Medline: 17425774]
15. Buchanan T. Internet-based questionnaire assessment: appropriate use in clinical contexts. Cogn Behav Ther 2003;32(3):100-109. [doi: 10.1080/16506070310000957] [Medline: 16291542]
16. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. J Med Internet Res 2011;13(4):e109 [FREE Full text] [doi: 10.2196/jmir.1593] [Medline: 22155721]
17. Davidson MA. ADHD in adults: a review of the literature. J Atten Disord 2008 May;11(6):628-641. [doi: 10.1177/1087054707310878] [Medline: 18094324]
18. Ramsay JR, Rostain AL. Adult ADHD research: current status and future directions. J Atten Disord 2008 May;11(6):624-627. [doi: 10.1177/1087054708314590] [Medline: 18417728]
19. Weiss M, Murray C. Assessment and management of attention-deficit hyperactivity disorder in adults. CMAJ 2003 Mar 18;168(6):715-722 [FREE Full text] [Medline: 12642429]
20. Steenhuis MP, Serra M, Minderaa RB, Hartman CA. An Internet version of the Diagnostic Interview Schedule for Children (DISC-IV): correspondence of the ADHD section with the paper-and-pencil version. Psychol Assess 2009 Jun;21(2):231-234. [doi: 10.1037/a0015925] [Medline: 19485678]
21. Bhatara VS, Vogt HB, Patrick S, Doniparthi L, Ellis R. Acceptability of a Web-based attention-deficit/hyperactivity disorder scale (T-SKAMP) by teachers: a pilot study. J Am Board Fam Med 2006;19(2):195-200 [FREE Full text] [Medline: 16513909]
22. Swanson JM, Schuck S, Mann-Porter M, Carlson C, Hartman CA, Sergeant JA, et al. Categorical and Dimensional Definitions and Evaluations of Symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. The International Journal of Educational and Psychological Assessment 2012;10:51-70.
23. Rösler M, Retz W, Stieglitz RD. Psychopathological rating scales as efficacy parameters in adult ADHD treatment investigations - benchmarking instruments for international multicentre trials. Pharmacopsychiatry 2010 May;43(3):92-98. [doi: 10.1055/s-0029-1242819] [Medline: 20127615]
24. Conners CK, Erhardt D, Sparrow E. Conner´s Adult ADHD Rating Scales (CAARS). Technical manual. North Tonawanda, NY: Multi-Health Systems; 1999.
25. Christiansen H, Hirsch O, Philipsen A, Oades RD, Matthies S, Hebebrand J, et al. German Validation of the Conners Adult ADHD Rating Scale-Self-Report: Confirmation of Factor Structure in a Large Sample of Participants With ADHD. J Atten Disord 2012 Mar 21. [doi: 10.1177/1087054711435680] [Medline: 22441889]
26. Christiansen H, Kis B, Hirsch O, Matthies S, Hebebrand J, Uekermann J, et al. German validation of the Conners Adult ADHD Rating Scales (CAARS) II: reliability, validity, diagnostic sensitivity and specificity. Eur Psychiatry 2012 Jul;27(5):321-328. [doi: 10.1016/j.eurpsy.2010.12.010] [Medline: 21392946]
27. Christiansen H, Kis B, Hirsch O, Philipsen A, Henneck M, Panczuk A, et al. German validation of the Conners Adult ADHD Rating Scales-self-report (CAARS-S) I: factor structure and normative data. Eur Psychiatry 2011 Mar;26(2):100-107. [doi: 10.1016/j.eurpsy.2009.12.024] [Medline: 20619613]
28. Unipark. URL: http://www.unipark.info/1-0-online-befragungssoftware-fuer-studenten-und-universitaeten-unipark-home.htm [accessed 2013-02-19] [WebCite Cache ID 6EY9AVRHg]
29. Retz-Junginger P, Retz W, Blocher D, Weijers HG, Trott GE, Wender PH, et al. [Wender Utah rating scale. The short-version for the assessment of the attention-deficit hyperactivity disorder in adults]. Nervenarzt 2002 Sep;73(9):830-838. [doi: 10.1007/s00115-001-1215-x] [Medline: 12215873]

30. Retz-Junginger P, Retz W, Blocher D, Stieglitz RD, Georg T, Supprian T, et al. [Reliability and validity of the Wender-Utah-Rating-Scale short form. Retrospective assessment of symptoms for attention deficit/hyperactivity disorder]. Nervenarzt 2003 Nov;74(11):987-993. [doi: 10.1007/s00115-002-1447-4] [Medline: 14598035]
31. Rösler M, Retz W, Retz-Junginger P, Thome J, Supprian T, Nissen T, et al. [Tools for the diagnosis of attention-deficit/hyperactivity disorder in adults. Self-rating behaviour questionnaire and diagnostic checklist]. Nervenarzt 2004 Sep;75(9):888-895. [doi: 10.1007/s00115-003-1622-2] [Medline: 15378249]
32. Blunch N. Introduction to structural equation modeling using SPSS and AMOS. London: Sage; 2008.
33. Kline RB. Principles and practice of structural equation modeling. New York: Guilford; 2005.
34. Arbuckle JL. Amos 17.0 User´s Guide. Chicago: SPSS Inc; 2008.
35. Weiber R, Mühlhaus D. Strukturgleichungsmodellierung Structural Equation Modeling. Berlin: Springer; 2009.
36. Raykov T, Marcoulides GA. A first course in structural equation modeling. Mahwah: Lawrence Erlbaum Associates; 2006.
37. George D, Mallery P. SPSS for Windows Step by Step: A Simple Guide and Reference, 18.0 Update. Boston: Allyn & Bacon; 2010.
38. Huber PJ, Ronchetti EM. Robust Statistics. Hoboken, NJ: John Wiley & Sons; 2009.
39. Grissom RJ, Kim JJ. Effect sizes for research. Univariate and multivariate applications. New York: Routledge; 2012.
40. Enders CK. Applied missing data analysis. New York: Guilford; 2010.
41. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006 Oct;59(10):1087-1091. [doi: 10.1016/j.jclinepi.2006.01.014] [Medline: 16980149]
42. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393 [FREE Full text] [Medline: 19564179]
43. O'Neil KM, Penrod SD, Bornstein BH. Web-based research: methodological variables' effects on dropout and sample characteristics. Behav Res Methods Instrum Comput 2003 May;35(2):217-226. [Medline: 12834076]
44. 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]
45. 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]
46. van den Berg MH, Overbeek A, van der Pal HJ, Versluys AB, Bresters D, van Leeuwen FE, et al. 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 [FREE Full text] [doi: 10.2196/jmir.1707] [Medline: 21955527]
47. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]
48. Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999 Aug;31(3):433-438. [Medline: 10502866]
49. Houston TK, Cooper LA, Vu HT, Kahn J, Toser J, Ford DE. Screening the public for depression through the Internet. Psychiatr Serv 2001 Mar;52(3):362-367. [Medline: 11239106]
50. 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]

Abbreviations

ADHD: Attention-Deficit/Hyperactivity Disorder
ADHS-SB: ADHD Self Rating Scale
AGFI: Adjusted Global Fit Index
CAARS: Conners Adult ADHD Rating Scales
CAARS-S: CAARS self-report
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition
GFI: Global Fit Index
ICD-10: International Classification of Diseases—Tenth Edition
MCAR: Missing completely at random
PS: Probability of Superiority
RMR: Root Mean Square Residual
SRMR: Standardized Root Mean Square Residual
WURS-k: Wender Utah Rating Scale-Short Form

Edited by G Eysenbach; submitted 17.06.12; peer-reviewed by L Whitehead, T Donker; comments to author 26.09.12; revised version received 15.10.12; accepted 17.01.13; published 21.03.13.

Please cite as:
Hirsch O, Hauschild F, Schmidt MH, Baum E, Christiansen H
Comparison of Web-Based and Paper-Based Administration of ADHD Questionnaires for Adults
J Med Internet Res 2013;15(3):e47
URL: http://www.jmir.org/2013/3/e47/
doi:10.2196/jmir.2225
PMID:23518816

Original Paper

Measuring Physical Activity in a Cardiac Rehabilitation Population Using a Smartphone-Based Questionnaire

Leila Pfaeffli[1], MA; Ralph Maddison[1], PhD; Yannan Jiang[1], PhD; Lance Dalleck[2], PhD; Marie Löf[3], PhD

[1]The National Institute for Health Innovation, University of Auckland, Auckland, New Zealand

[2]Department of Sport and Exercise Science, University of Auckland, Auckland, New Zealand

[3]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Leila Pfaeffli, MA
The National Institute for Health Innovation
University of Auckland
Private Bag 92019, Auckland Mail Centre
Auckland, 1142
New Zealand
Phone: 64 9 373 7599 ext 84728
Fax: 64 9 373 1710
Email: l.pfaeffli@nihi.auckland.ac.nz

Abstract

Background: Questionnaires are commonly used to assess physical activity in large population-based studies because of their low cost and convenience. Many self-report physical activity questionnaires have been shown to be valid and reliable measures, but they are subject to measurement errors and misreporting, often due to lengthy recall periods. Mobile phones offer a novel approach to measure self-reported physical activity on a daily basis and offer real-time data collection with the potential to enhance recall.

Objective: The aims of this study were to determine the convergent validity of a mobile phone physical activity (MobilePAL) questionnaire against accelerometry in people with cardiovascular disease (CVD), and to compare how the MobilePAL questionnaire performed compared with the commonly used self-recall International Physical Activity Questionnaire (IPAQ).

Methods: Thirty adults aged 49 to 85 years with CVD were recruited from a local exercise-based cardiac rehabilitation clinic in Auckland, New Zealand. All participants completed a demographics questionnaire and underwent a 6-minute walk test at the first visit. Subsequently, participants were temporarily provided a smartphone (with the MobilePAL questionnaire preloaded that asked 2 questions daily) and an accelerometer, which was to be worn for 7 days. After 1 week, a follow-up visit was completed during which the smartphone and accelerometer were returned, and participants completed the IPAQ.

Results: Average daily physical activity level measured using the MobilePAL questionnaire showed moderate correlation (r=.45; P=.01) with daily activity counts per minute (Acc_CPM) and estimated metabolic equivalents (MET) (r=.45; P=.01) measured using the accelerometer. Both MobilePAL (beta=.42; P=.008) and age (beta=–.48, P=.002) were significantly associated with Acc_CPM (adjusted R^2=.40). When IPAQ-derived energy expenditure, measured in MET-minutes per week (IPAQ_met), was considered in the predicted model, both IPAQ_met (beta=.51; P=.001) and age (beta=–.36; P=.016) made unique contributions (adjusted R^2=.47, $F_{2,27}$=13.58; P<.001).There was also a significant association between the MobilePAL and IPAQ measures (r=.49, beta=.51; P=.007).

Conclusions: A mobile phone–delivered questionnaire is a relatively reliable and valid measure of physical activity in a CVD cohort. Reliability and validity measures in the present study are comparable to existing self-report measures. Given their ubiquitous use, mobile phones may be an effective method for physical activity surveillance data collection.

(J Med Internet Res 2013;15(3):e61) doi:10.2196/jmir.2419

KEYWORDS

cellular phone; self report; motor activity; bias; cardiovascular diseases

Introduction

The case for a new technology to measure physical activity is compelling. Participation in regular physical activity is associated with a plethora of positive physical and mental health outcomes [1-3], and the burden associated with physical inactivity is considerable [4,5]. Much of the data supporting the commonly known benefits of physical activity are based on self-reported measures of physical activity. Physical activity questionnaires are commonly used to assess physical activity in large population-based studies because of their low cost and convenience [6]. Although many of the self-report questionnaires used in these studies have shown to be valid and reliable measures, they do have limitations. The self-report approach is subject to measurement errors and misreporting, including deliberate social desirability bias and unintentional bias, such as recall or comprehension error, all of which reduce the precision of the estimate of levels of activity [7-10]. Caution must be taken to select an appropriate physical activity questionnaire according to the purpose of the research and the population under investigation [7].

A major source of bias with self-report questionnaires is the recall period [11]. Typically, self-report measures require participants to remember their physical activities during specific periods of time, such as 3 to 7 days; however, the more distal the recall period, the greater the recall error [12,13]. For example, studies have shown that when using the International Physical Activity Questionnaire (IPAQ), people tended to overreport their physical activity and had difficulty accurately recalling the intensity of the activity done over 1 week [13]. For these reasons, researchers have used diaries or activity logs to record self-reported physical activity on a daily basis to enhance recall [12]. Such approaches require participants to complete paper-based records, which are associated with considerable participant burden and call for sustained cooperation [6].

The ubiquitous use of mobile phones offers a novel approach to measuring physical activity and to reduce participant burden. Mobile phones offer the potential to capture self-report physical activity on a daily basis and offer real-time data collection. Because most people carry a mobile phone most of the time, they have the potential to enhance recall of physical activity by frequent prompting and limiting the time lag between the behavior and data collection. This may reduce information bias and increase compliance; same day or previous day recall has been shown to reduce recall error because error tends to increase with recall duration [12]. Moreover, increased access and availability to mobile phone telecommunications increase the potential of this tool for large-scale data collection in population-based studies [14].

A recent study validated a mobile phone-delivered physical activity questionnaire against doubly labeled water [15]. Twenty-two women reported their physical activity over a 14-day period by answering 2 questions sent daily to their mobile phones. A small mean difference (0.014) with narrow limits of agreement (2 SD 0.30) was found between the mobile phone questionnaire and the reference estimates. In a second study [16], the mobile phone physical activity questionnaire was compared against accelerometry. Both methods showed high within-subject variations; however, the day-to-day variations in energy expenditure within subjects assessed using the mobile phone agreed well with corresponding accelerometer values. The authors concluded that the mobile phone questionnaire was a promising tool for assessing levels of physical activity.

Despite these positive effects, the mobile phone questionnaire has only been examined in healthy Swedish women (aged 20-45 years). Further research is needed to assess the reliability and validity of this questionnaire in males as well as females, with a wider age range. Given the importance of physical activity participation for prevention of chronic diseases and for the secondary prevention of cardiovascular disease (CVD), examination of this approach in a clinical population was warranted. Improving physical activity levels is a key objective following a cardiac event. Not all patients attend structured exercise-based cardiac rehabilitation programs, and they may also exercise in their own time, so it is important that the method used captures habitual activity. The purpose of this study was to determine the convergent validity of a mobile phone physical activity questionnaire against accelerometry in people with CVD. A second aim was to compare how the mobile phone questionnaire performed compared with a commonly used self-recall physical activity questionnaire.

Methods

Study Procedures

A 7-day convergent validation study was conducted from January to May, 2012. Participants were recruited from a local exercise-based cardiac rehabilitation clinic in Auckland, New Zealand, and were included if they had documented history of CVD, were currently participating in cardiac rehabilitation, and could safely perform exercise.

The study involved 2 visits. At the first visit, consenting participants completed a demographics questionnaire and underwent a 6-minute walk test (6MWT). Subsequently, participants were temporarily provided with a smartphone and accelerometer. All participants were provided with verbal and written instructions about how to complete the mobile phone physical activity level questionnaire (MobilePAL). Participants responded to 2 physical activity questions initiated by the smartphone application each day for 7 days. The questions were sent to all participants at 7:00 pm each evening. Participants were shown how to properly wear the accelerometer and instructed to wear it for the same 7-day period.

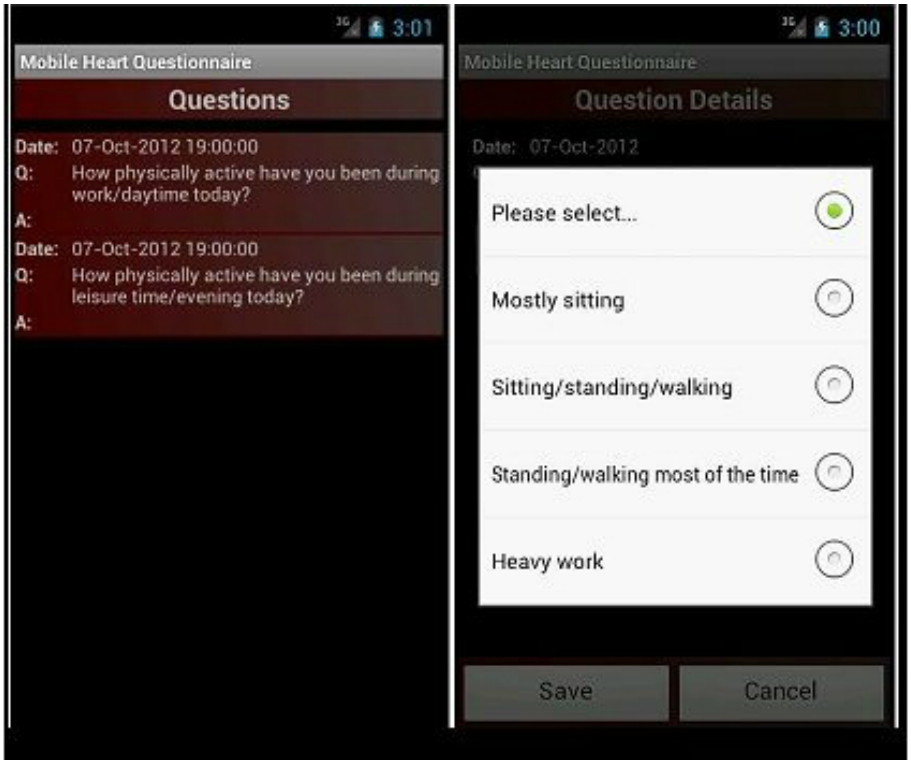
At the end of 1 week, participants completed the second visit, in which they returned the smartphone and accelerometer, and completed a paper copy of the IPAQ.

Mobile Phone Questionnaire Development

For this study, we adapted the original Java-based mobile phone questionnaire [15] for delivery via an Android application (see Figure 1). Smartphone use is increasing, with 44% of people in the United Kingdom [17] and 50% of Americans using smartphones in 2012 [18]. Affordability continues to improve

as the cost of smartphones and data plans decrease [19]. Administering the questionnaire by smartphone application offers several advantages over the original Java approach. Participants were not required to have a subscriber identity module (SIM) card or access the Internet to answer the questions, eliminating any costs to the user and reliance on cellular phone networks. Data were saved directly onto the phone in a comma-separated values (CSV) file format, which was then uploaded to a server where it could be safely stored. The application was pretested to ensure the questionnaire functioned and uploaded data correctly.

Figure 1. Screenshot of questionnaire and answer categories of the MobilePAL smartphone questionnaire.



Measures

Accelerometer

Some consensus exists that accelerometry-based activity monitors provide a useful comparison for evaluating self-report instruments. They can provide detailed information about activity patterns on a minute-by-minute basis and impose only minimal burden on participants [6]. According to Sirard and Pate's [20] measurement hierarchy, secondary measures such as accelerometers are an acceptable comparison for validating self-report methods. In the present study, accelerometry was used to provide convergent validity against the self-report measures.

Participants wore a dual axial Actigraph GT1M accelerometer (Model AM7164-2.2C Actigraph Ltd, Pensacola, FL, USA), a reliable and valid objective measure of physical activity [21]. The Actigraph is a small, closed device worn at the hip (belt clip or elastic band) that records body movements. Activity counts generated per minute (cpm) were used to determine time spent in light, moderate, and vigorous intensity activity. The following cut-off points were used to determine the intensity of physical activity: sedentary ≤100 cpm, light 101-2020 cpm, moderate 2021-5999 cpm, and vigorous ≥6000 cpm. A valid day consisted of wearing the accelerometer for ≥10 hours of valid time, which was defined as those minutes with <1 hour of consecutive zeros. Average activity counts per minute were calculated for each participant on each valid day, and then averaged over valid days to get the average daily activity count per minute (Acc_CPM). Average daily minutes spent in lifestyle, light, moderate, and vigorous physical activity was also calculated (Acc_PAmin). Activity count data were also used to estimate energy expenditure using the Freedson et al [22] metabolic equivalent (MET) regression equation. METs are multiples of resting metabolic rate during a specific activity, in which 1 MET is equivalent to rest. Daily MET values were calculated by averaging the MET values per minute (converted from raw activity counts per minute using the equation) over each valid day. Average daily METs (Acc_METs) were calculated by summing daily MET values and dividing by the number of days of valid data.

MobilePAL Questionnaire

The mobile application presented 2 questions each evening to the user about their physical activity that day, for a total of 7 days. Consistent with previous research [15], for each person and for each day, the answers to the 2 short questions (Table 1) were converted to physical activity level (PAL), which is the ratio between total energy expenditure and resting energy expenditure during 24 hours. The PAL was calculated by combining the PAL values obtained for work/daytime activities reported by Black et al [23] and an additional contribution to PAL from energy expended during leisure/evening activities (Table 1). The latter was calculated from published energy costs

expressed as MET values for walking and cycling [24]. PAL data were extracted from the smartphone and imported into Microsoft Excel. Average daily PAL was calculated by summing the daily PAL values and then dividing by 7 days (MobilePAL).

Table 1. The two questions administered daily by the smartphone application and their corresponding physical activity level (PAL) value score.

Question	Answer category	PAL score[a]
How physically active have you been during work/daytime today?	Mostly sitting	1.55
	Sitting/standing/walking	1.65
	Standing/walking most of the time	1.85
	Heavy work	2.2
How physically active have you been during leisure time/evening today?	Mostly sitting	+0
	Light/walking (30 min)	+0.06
	Moderate/cycling (>30 min)	+0.15
	Sport/cycling (>60 min)	+0.29

[a]Daily PAL was calculated by combining the PAL score from question 1 (work/daytime physical activity) and question 2 (leisure time/evening physical activity).

International Physical Activity Questionnaire (IPAQ)

The IPAQ is a reliable and validated 7-day recall measure, which provides a comprehensive evaluation of daily physical activities, and assesses time spent walking and doing light, moderate, and vigorous intensity activities across various domains [24]. Computation of the total scores required summation of the duration (in minutes) and frequency (days) for all the types of activities in all domains. Domain specific scores were calculated by summing the scores for walking, moderate, and vigorous intensity activities within the specific domain. Activity-specific scores were calculated by summing the scores for the specific type of activity across domains.

Two variables were derived from the IPAQ data: (1) average daily active minutes (IPAQ_PAmin), which was calculated by summing total time spent walking, in moderate, and in vigorous intensity physical activity, and then dividing by 7 days; (2) average daily physical activity level (IPAQ_met), which was calculated using total physical activity (MET-minutes per week) divided by 7 days. MET-minutes per week were derived as duration×frequency per week×MET intensity assigned to each category of activity [25].

Six-Minute Walk Test (6MWT)

The 6MWT is a test of physical capacity commonly used in the assessment of cardiac rehabilitation patients and was used to assess functional capacity of participants. Each participant completed the test once during their first study visit. The 6MWTs were administered by a research assistant using a standard protocol [26].

Analysis

Statistical analyses were performed using SAS version 9.2 (SAS Institute, Inc, Cary, NC, USA) and R version 2.15.0 (R Foundations for Statistical Computing, Vienna, Austria). All statistical tests were 2-tailed at a 5% significance level. Participants' characteristics and physical activity measurements were first summarized using descriptive statistics. Both Pearson product moment correlation (r) and Spearman rank order correlation were used to assess the strength of correlation between 2 measures with associated P values. Regression analyses were carried out to investigate the relationships between the MobilePAL questionnaire, the accelerometer, and self-report physical activity measured by the IPAQ. For accelerometer data, Acc_CPM was compared with MobilePAL because this captures all activities performed by the person when wearing the device, including incidental and lifestyle activities. As stated, we estimated average daily METs (Acc_METs), which reflects activity-related energy expenditure. In the present study, we did not collect body mass data; therefore, we were unable to estimate basal metabolic rate and could not truly estimate PAL. The Acc_METs was considered an appropriate proxy measure for comparison with MobilePAL. Energy expenditure obtained from the IPAQ (IPAQ_met) was chosen as a comparator because it includes a similar measurement unit as MobilePAL. Potential confounding effects of age and functional capacity (6MWT) were examined in all models. Repeated measures analysis was also conducted to evaluate the change in PAL over the 7-day period.

A Shapiro-Wilk test of normality was conducted with the relatively small sample size (N=30). When necessary, log transformation of the outcome variable was considered. Because the results with and without the log transformation were similar in all analyses, the original (nontransformed) data are presented.

Results

As shown in Table 2, most participants were New Zealand European (29/30, 97%), men (26/30, 87%), aged between 49 to 85 years (mean 65.6, SD 8.8). Thirty-six potential participants were approached to take part. Of these, 32 expressed interest and 30 completed the study (30/36, 83%). More than half were working full or part time (19/30, 63%) and 11/30 (37%) were retired. Thirteen participants (43%) had never smoked and 17/30 (57%) identified as previous smokers. Twenty-one participants (70%) consumed at least 1 alcoholic drink per week. Distance walked during the 6MWT ranged from 372 to 742 meters (mean 570.8, SD 96.3).

Descriptive summaries of all physical activity measurements obtained by the 3 different instruments are presented in Table 3. All participants provided at least 4 days of valid accelerometer data. Average daily active minutes measured by IPAQ (IPAQ_PAmin) were on average lower than the average daily valid minutes recorded by accelerometer (Acc_PAmin), which is not surprising because the IPAQ does not capture incidental movement.

Table 2. Participant characteristics (N=30).

Variables		Participants n (%)
Gender		
	Male	26 (87)
	Female	4 (13)
Ethnicity		
	New Zealand European	29 (97)
	Māori (indigenous)	1 (3)
Medical status[a]		
	High blood pressure	23 (77)
	High cholesterol	26 (87)
	Diabetes	5 (17)
	Atrial fibrillation	10 (33)
	Heart attack	14 (47)
	Angina	3 (10)
	Other forms of heart disease	12 (40)

[a]Some participants reported having more than 1 medical condition.

Table 3. Summary of average daily physical activity obtained by the 3 different instruments, the smartphone, the accelerometer, and the International Physical Activity Questionnaire (IPAQ).

Instrument and measurement		Mean (SD)	Range
Smartphone[a]			
	MobilePAL	1.77 (0.1)	1.6-2.1
Accelerometer[b]			
	Acc_CPM	313 (140)	108-702
	Acc_METs	1.69 (0.1)	1.5-2.0
	Acc_PAmin	302 (74)	188-528
IPAQ[c]			
	IPAQ_met	531 (468)	47-1840
	IPAQ_PAmin	149 (131)	14-504

[a]MobilePAL=daily physical activity level (PAL) measured by smartphone questionnaire.

[b]Acc_CPM=daily activity counts per minute measured by accelerometer; Acc_METs=average daily metabolic equivalent derived from accelerometer counts per minute; Acc_PAmin=daily minutes of lifestyle, light, moderate, or vigorous physical activity measured by accelerometer.

[c]IPAQ_met=MET-minutes per day measured by IPAQ; IPAQ_PAmin=daily minutes of walking, in moderate intensity, and in vigorous intensity physical activity measured by IPAQ.

The Pearson's correlation coefficients are presented in Table 4. Associations between the MobilePAL questionnaire and accelerometer daily activity counts (Acc_CPM), activity-derived METs (Acc_METs), and IPAQ-derived energy expenditure (IPAQ_met) were similar in magnitude. The conclusions were consistent using Spearman rank correlation and, therefore, are not reported here. Graphs illustrating the degree of spread in the data are shown in Multimedia Appendix 1. Functional capacity measured by the 6MWT was not associated with any variables except age (r=–.43; P=.02). Age was not associated with the self-report measures, but showed moderate correlations with Acc_CPM (r=–.53; P=.003) and Acc_PAmin (r=–.46; P=.01).

Table 4. Correlations between 3 different instruments, the smartphone, the accelerometer, and the International Physical Activity Questionnaire (IPAQ).

Instrument and measurement	Correlations, r[a]					
	MobilePAL	Acc_CPM	Acc_METs	Acc_PAmin	IPAQ_met	IPAQ_PAmin
Smartphone[b]						
MobilePAL	1.00					
Accelerometer[c]						
Acc_CPM	*.45*	1.00				
Acc_METs	*.45*	*1.00*	1.00			
Acc_PAmin	*.39*	*.81*	*.81*	1.00		
IPAQ[d]						
IPAQ_met	*.49*	*.62*	*.62*	*.40*	1.00	
IPAQ_PAmin	*.48*	*.61*	*.61*	*.41*	*.99*	1.00

[a]Statistically significant correlations (P<.05) are indicated in italics.

[b]MobilePAL=daily physical activity level (PAL) measured by smartphone questionnaire.

[c]Acc_CPM=daily activity counts per minute measured by accelerometer; Acc_METs= Average daily metabolic equivalent derived from accelerometer counts per minute; Acc_PAmin=Daily minutes of lifestyle, light, moderate, or vigorous physical activity measured by accelerometer.

[d]IPAQ_met=MET- minutes per day measured by IPAQ; IPAQ_PAmin=daily minutes of walking, in moderate intensity, and in vigorous intensity physical activity measured by IPAQ.

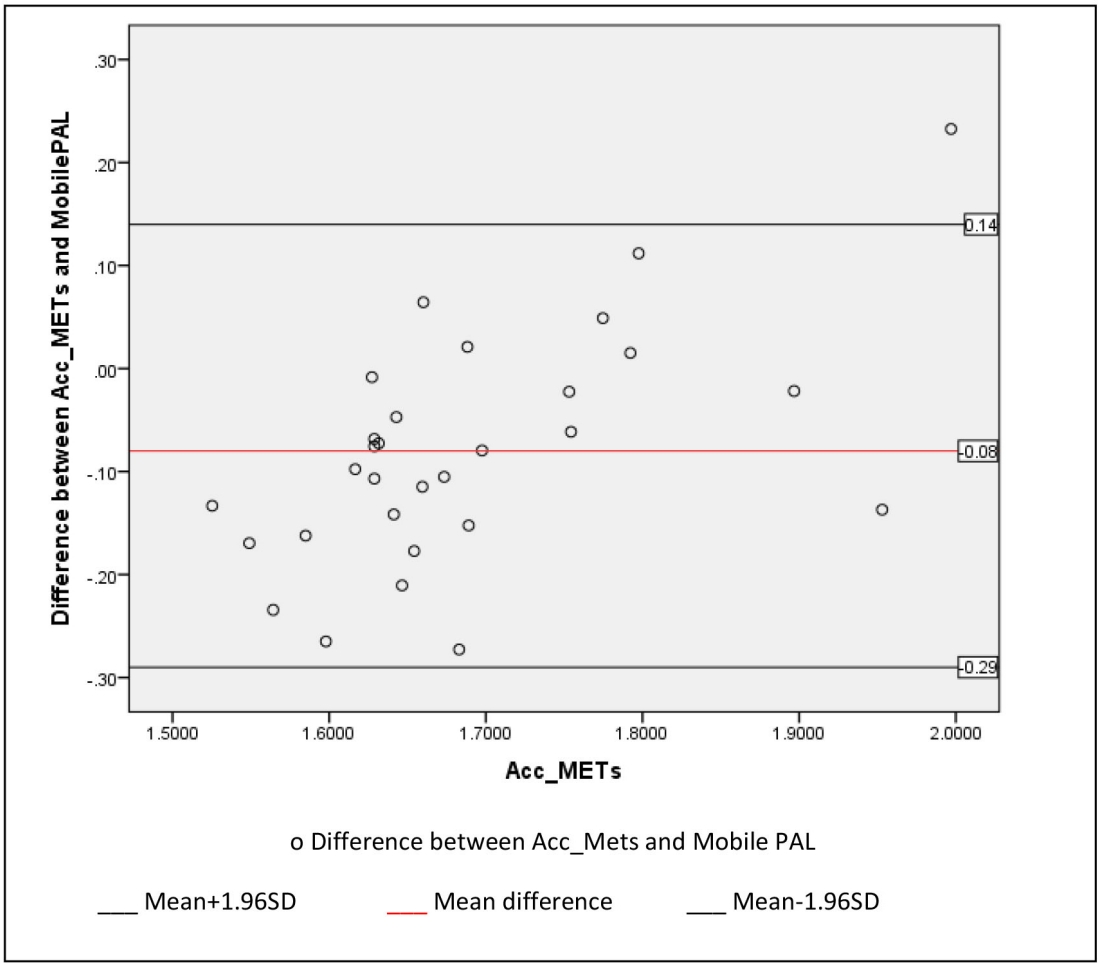
To determine the level of agreement between accelerometer-derived energy expenditure (ACC_METs) and Mobile PAL we conducted a Bland-Altman plot (see Figure 2). Overall, there was good agreement between the methods, with a mean bias of +0.08 METs.

Linear regression analyses were used to further investigate the relationships between the variables of interest. Because 6MWT was not associated with any of these variables, only age was adjusted in the regression analysis.

We first examined the association between MobilePAL and the reference standard of accelerometry (Acc_CPM). Both MobilePAL (beta=.42; P=.008) and age (beta=–.48; P=.002) were significant predictors of Acc_CPM (adjusted R^2=.40, $F_{2,27}$=10.58; P<.01). Because Acc_METs were derived from accelerometer cpm, a separate regression equation was not conducted. Similar findings were found when IPAQ-derived energy expenditure was considered in the model. Both IPAQ_met (beta=.51; P=.001) and age (beta=–.36; P=.016) made unique contributions to the predicted model (adjusted R^2=.47, $F_{2,27}$=13.58; P<.001). We also found that IPAQ_met (beta=.51; P=.007) was strongly associated with MobilePAL (adjusted R^2=0.19, $F_{2,27}$=4.32; P<.05), but not age (beta=.07; P=.682).

Repeated measures analysis on MobilePAL over 7 days revealed little within-day variability. No differences in MobilePAL between any 2 days were observed after Tukey-Kramer adjustment, except for days 6 and 7 (adjusted P=.04).

Figure 2. Bland-Altman plot comparison of energy expenditure obtained using accelerometry (Acc_METs) and the mobile phone questionnaire (MobilePAL).



Discussion

We found relatively good associations between the mobile phone-derived activity-related energy expenditure (MobilePAL) and the accelerometer-derived total daily activity counts (Acc_CPM), which was similar to that observed with a commonly used self-report questionnaire, the IPAQ, and slightly better than that demonstrated in the 12-country validation of the IPAQ [25]. The similar magnitude of correlations for IPAQ and MobilePAL with accelerometry data indicates the mobile phone questionnaire is as good as traditional paper-based questionnaires. There is potential that the mobile phone questionnaire may have enhanced recall on the IPAQ, as participants were answering questions each day about their physical activities. This constant recall may have led to stronger associations. Good agreement existed between mobile phone and accelerometer derived activity-related energy expenditure (MobilePAL and Acc_METs) with only slight overestimation of the mobile phone questionnaire compared to the reference standard of accelerometry. Taken together, these findings support the use of the mobile phone questionnaire to assess physical activity levels in people with CVD.

Our study sample included attendees of a cardiac rehabilitation program; however, many people with CVD are encouraged to exercise but do not attend cardiac rehabilitation programs [27-30]. Moreover, adherence at cardiac rehabilitation is low and many people exercise on their own time in their community. This mobile phone questionnaire offers a viable approach to monitoring activity levels of people with CVD, irrespective of attendance at cardiac rehabilitation. Such data are needed to truly understand the benefits of physical activity for secondary prevention of CVD.

In comparison to other cardiac rehabilitation populations, our sample had a similar mean age [31]. In the present study, age was a significant predictor in the regression analyses; the older participants in our sample were less active than the younger participants, as revealed by the inverse relationship between accelerometer counts per minute and age. Moreover, participants in our study had better physical function with a mean 6MWT of 570 meters, which is greater than that observed in other postcardiac rehabilitation populations (377-555 m) [31]. This may be due to a higher exercise dose as our sample participated in a supervised exercise program 3 times per week.

The repeated measures analyses showed little within-day variability for the mobile phone questionnaire. For 28 participants, day 6 fell on a Tuesday (n=23) or a Thursday (n=5). The study participants belonged to a cardiac rehabilitation exercise clinic that ran on Mondays, Wednesdays, and Fridays, so perhaps the lower PAL score on day 6 was a reflection of a rest day for some participants.

This study builds on previous validation research by Bexelius and colleagues [15] by testing the convergent validity of the mobile phone questionnaire in people with CVD. Previously, the questionnaire has demonstrated validity against doubly labeled water and accelerometry in Swedish women. Our findings extend the generalizability of this questionnaire; taken together, these findings suggest that this mobile phone questionnaire is a reliable and valid self-report measure of physical activity.

Recently, another mobile phone-based physical activity questionnaire has been developed. Sternfeld and colleagues [32] evaluated an activity diary as an application program to be administered using a mobile phone. Participants were asked to record their physical activities on their phone 3 times a day. Participants could choose activities across 15 different domains, and responses were associated with MET values derived from the Compendium of Physical Activities [24]. Compared to accelerometry, intraclass correlation coefficients ranged from 0.55 for light physical activity to 0.63 for vigorous activity, whereas correlations were of moderate magnitude and slightly higher than observed in the present study. As with our study, there was good user acceptability. Collectively, these studies highlight the utility of mobile phones for self-reporting physical activity.

Despite these findings, it is important that researchers do not necessarily develop a completely new series of mobile phone applications for self-reporting physical activity, but rather build on previous research by refining existing platforms. This would avoid the unnecessary proliferation of questionnaires observed in the field of physical activity research. To illustrate, there are currently more than 100 self-report measures of physical activity in use, with varying degrees of reliability and validity. A recent review by Helmerhorst and colleagues [7] identified 34 physical activity questionnaires developed between 1997 and 2011, and found that these were no more reliable or valid compared to existing measures. We acknowledge that the field of physical activity measurement and the resulting techniques and analytic approaches has progressed considerably as a function of the development of these questionnaires. However, we suggest that as researchers, it is in our best interest not to continue this trend of proliferation with mobile phone questionnaires. A more fruitful approach might be to build on existing measurement expertise and work to refine or develop a universal mobile phone questionnaire for population-level use, as has been done with the IPAQ or Global Physical Activity Questionnaire.

An important finding from this study is that our sample of middle-aged to older adults were able to use the smartphone application. The digital divide, whereby some groups (including older-aged people) may use technology less than others, has been considered a barrier for researchers using smartphone applications [19]. However, there is now abundant evidence that mobile phones offer unprecedented opportunities to improve reach into traditionally underserved population groups [33,34]. Indeed, the telecommunications industry has documented a trend toward a digital divide in reverse, whereby low income and ethnic minority groups use the technology more than others [14]. It is estimated that 80% to 90% of the UK population will have a smartphone within 10 years [19]. Given this increasing use and availability of smartphone technology, combined with real-time data collection that is easy to use, makes this a promising way to obtain physical activity data and retain participant compliance. Participants are required to answer only 2 questions, thereby reducing the burden typically associated with other commonly used instruments. It takes less than 30 seconds to answer the MobilePAL questionnaire each day, whereas the IPAQ took our participants between 10 to 15 minutes to complete. Previous research has highlighted the need to make mobile phone applications easy to use (ie, 1 click from main page) [35]. This was an important consideration in our application development; hence, participants required a single click from the main page to access the application, 1 click to select answers to the questions, and 1 click to save responses. Most participants had never used a smartphone prior to this study, and anecdotal responses indicated that they generally found the application easy to use. Collectively, these features make mobile phone questionnaires, such as the one presented here, suitable for population-level surveillance.

This study is not without limitations. First our findings are based on a small (N=30) and relatively homogeneous sample of New Zealand European men, which impacts on the variability and generalizability of the data. However, the mobile phone questionnaire has been validated in other populations using doubly labeled water and indirect calorimetry [15]. Welk [6] describes validity as an ongoing and community process because validation cannot be determined by one study alone. Future research is needed to validate this questionnaire in other subgroups and countries.

In conclusion, a mobile phone–delivered questionnaire is a relatively reliable and valid measure of physical activity, and is as good as existing self-report measures. Given their ubiquitous use, mobile phones may be an effective method of physical activity surveillance data collection.

Acknowledgments

The authors would like to thank all the participants and the University of Auckland's Cardiac Rehabilitation Clinic, in particular Ms Brighid McCaffrey. We would also like to acknowledge the National Institute of Health Innovation's IT team, especially Johan Strydom for his work on the development of the application. Dr Maddison is supported by a Health Research Council, Sir Charles Hercus Fellowship. This study was supported by funds from a Health Research Council Project Grant.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Degree of spread between variables.

[PDF File (Adobe PDF File), 187KB - jmir_v15i3e61_app1.pdf]

References

1. Blair SN, Kohl HW, Gordon NF, Paffenbarger RS. How much physical activity is good for health? Annu Rev Public Health 1992;13:99-126. [doi: 10.1146/annurev.pu.13.050192.000531] [Medline: 1599603]
2. Blair SN, Morris JN. Healthy hearts-and the universal benefits of being physically active: physical activity and health. Ann Epidemiol 2009 Apr;19(4):253-256. [doi: 10.1016/j.annepidem.2009.01.019] [Medline: 19344864]
3. O'Donovan G, Blazevich AJ, Boreham C, Cooper AR, Crank H, Ekelund U, et al. The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. J Sports Sci 2010 Apr;28(6):573-591. [doi: 10.1080/02640411003671212] [Medline: 20401789]
4. Bull FC, Armstrong TP, Dixon T. Physical inactivity. In: Essati M, Lopez AD, Rodgers A, Murray CJL, editors. Comparative Quantification of Health Risks Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Geneva: World Health Organization; 2004:1-881.
5. 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]
6. Welk GJ. Physical Activity Assessment for Health-Related Research. Champaign, IL: Human Kinetics; 2002.
7. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act 2012;9:103 [FREE Full text] [doi: 10.1186/1479-5868-9-103] [Medline: 22938557]
8. Lamonte MJ, Ainsworth BE. Quantifying energy expenditure and physical activity in the context of dose response. Med Sci Sports Exerc 2001 Jun;33(6 Suppl):S370-8; discussion S419. [Medline: 11427762]
9. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 2003 Jun;37(3):197-206 [FREE Full text] [Medline: 12782543]
10. Westerterp KR. Assessment of physical activity: a critical appraisal. Eur J Appl Physiol 2009 Apr;105(6):823-828. [doi: 10.1007/s00421-009-1000-2] [Medline: 19205725]
11. Ainsworth BE, Caspersen CJ, Matthews CE, Mâsse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. J Phys Act Health 2012 Jan;9 Suppl 1:S76-S84 [FREE Full text] [Medline: 22287451]
12. Dishman RK, Washburn RA, Schoeller DA. Measurement of physical activity. Quest 2001 Aug;53(3):295-309. [doi: 10.1080/00336297.2001.10491746]
13. Rzewnicki R, Vanden Auweele Y, De Bourdeaudhuij I. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. Public Health Nutr 2003 May;6(3):299-305. [doi: 10.1079/PHN2002427] [Medline: 12740079]
14. World Health Organization. mHealth: New Horizons for Health through Mobile Technologies: Second Global Survey on eHealth (Global Observatory for Ehealth). Geneva: World Health Organization; 2011.
15. 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]
16. 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 [FREE Full text] [doi: 10.2196/jmir.1686] [Medline: 21946080]
17. Ofcom. 2012. Adults media use and attitudes report 2012 URL: http://stakeholders.ofcom.org.uk/market-data-research/media-literacy/archive/medlitpub/medlitpubrss/adults-media-use-attitudes/ [accessed 2012-10-28] [WebCite Cache ID 6BlCgO42C]
18. Nielsen Mobile Insights. America's New Mobile Majority: A Look at Smartphone Owners in the US. 2012. URL: http://blog.nielsen.com/nielsenwire/online_mobile/smartphones-account-for-half-of-all-mobile-phones-dominate-new-phone-purchases-in-the-us/ [accessed 2012-10-28] [WebCite Cache ID 6BlDMCG9P]
19. Boulos MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]

20. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med 2001;31(6):439-454. [Medline: 11394563]
21. Welk GJ, Schaben JA, Morrow JR. Reliability of accelerometry-based activity monitors: a generalizability study. Med Sci Sports Exerc 2004 Sep;36(9):1637-1645. [Medline: 15354049]
22. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [Medline: 9588623]
23. 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]
24. 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]
25. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
26. Enright PL. The six-minute walk test. Respir Care 2003 Aug;48(8):783-785 [FREE Full text] [Medline: 12890299]
27. Bethell HJ, Evans JA, Turner SC, Lewin RJ. The rise and fall of cardiac rehabilitation in the United Kingdom since 1998. J Public Health (Oxf) 2007 Mar;29(1):57-61 [FREE Full text] [doi: 10.1093/pubmed/fdl091] [Medline: 17189295]
28. Bethell H, Lewin R, Dalal H. Cardiac rehabilitation in the United Kingdom. Heart 2009 Feb;95(4):271-275. [doi: 10.1136/hrt.2007.134338] [Medline: 18208830]
29. Suaya JA, Shepard DS, Normand SL, Ades PA, Prottas J, Stason WB. Use of cardiac rehabilitation by Medicare beneficiaries after myocardial infarction or coronary bypass surgery. Circulation 2007 Oct 9;116(15):1653-1662 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.701466] [Medline: 17893274]
30. Banerjee AT, Gupta M, Singh N. Patient characteristics, compliance, and exercise outcomes of South Asians enrolled in cardiac rehabilitation. J Cardiopulm Rehabil Prev 2007;27(4):212-218. [doi: 10.1097/01.HCR.0000281765.52158.be] [Medline: 17667016]
31. Bellet RN, Adams L, Morris NR. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness-a systematic review. Physiotherapy 2012 Dec;98(4):277-286. [doi: 10.1016/j.physio.2011.11.003] [Medline: 23122432]
32. Sternfeld B, Jiang SF, Picchi T, Chasan-Taber L, Ainsworth B, Quesenberry CP. Evaluation of a cell phone-based physical activity diary. Med Sci Sports Exerc 2012 Mar;44(3):487-495. [doi: 10.1249/MSS.0b013e3182325f45] [Medline: 21857369]
33. Ory MG, Lee Smith M, Mier N, Wernicke MM. The science of sustaining health behavior change: the health maintenance consortium. Am J Health Behav 2010;34(6):647-659. [Medline: 20604691]
34. Zurovac D, Talisuna AO, Snow RW. Mobile phone text messaging: tool for malaria control in Africa. PLoS Med 2012 Feb;9(2):e1001176 [FREE Full text] [doi: 10.1371/journal.pmed.1001176] [Medline: 22363212]
35. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]

Abbreviations

6MWT: 6-minute walk test
Acc_CPM: daily activity counts per minute measured by accelerometer
Acc_METs: average daily metabolic equivalent derived from accelerometer counts per minute
Acc_PAmin: daily minutes of lifestyle, light, moderate, or vigorous physical activity measured by accelerometer
CSV: comma-separated values
CVD: cardiovascular disease
IPAQ: International Physical Activity Questionnaire
IPAQ_met: MET-minutes per day measured by IPAQ
IPAQ_PAmin: daily minutes of walking, in moderate intensity, and in vigorous intensity physical activity measured by IPAQ
MET: metabolic equivalent
MobilePAL: daily physical activity level (PAL) measured by smartphone questionnaire
PAL: physical activity level
SIM: subscriber identity module

Edited by G Eysenbach; submitted 29.10.12; peer-reviewed by K Westerterp, H Namba, D Basset, Jr, S Bonn; comments to author 14.11.12; revised version received 19.12.12; accepted 22.01.13; published 22.03.13.

Please cite as:
Pfaeffli L, Maddison R, Jiang Y, Dalleck L, Löf M
Measuring Physical Activity in a Cardiac Rehabilitation Population Using a Smartphone-Based Questionnaire
J Med Internet Res 2013;15(3):e61
URL: http://www.jmir.org/2013/3/e61/
doi:10.2196/jmir.2419
PMID:23524251

Original Paper

Development and Testing of a Multidimensional iPhone Pain Assessment Application for Adolescents with Cancer

Jennifer N Stinson[1,2], RN, PhD; Lindsay A Jibb[1,2], MSc, RN; Cynthia Nguyen[1], BSc; Paul C Nathan[1,2], MD, MSc; Anne Marie Maloney[1,2], NP-Paediatrics, MSN; L Lee Dupuis[1,2], RPh, MScPhm; J Ted Gerstle[1,2], MD; Benjamin Alman[1,2], MD; Sevan Hopyan[1,2], MD, PhD; Caron Strahlendorf[3], MBBCh; Carol Portwine[4], MD, PhD; Donna L Johnston[5,6], MD; Mike Orr[7], BASc, MBA

[1]The Hospital for Sick Children, Toronto, ON, Canada
[2]University of Toronto, Toronto, ON, Canada
[3]BC Children's Hospital, Vancouver, BC, Canada
[4]McMaster University, Hamilton, ON, Canada
[5]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[6]University of Ottawa, Ottawa, ON, Canada
[7]Cundari Group Limited, Toronto, ON, Canada

Corresponding Author:
Jennifer N Stinson, RN, PhD
The Hospital for Sick Children
555 University Ave
Toronto, ON,
Canada
Phone: 1 416 813 8501 ext 4514
Fax: 1 416 813 8501
Email: jennifer.stinson@sickkids.ca

Abstract

Background: Pain is one of the most common and distressing symptoms reported by adolescents with cancer. Despite advancements in pain assessment and management research, pain due to cancer and/or its treatments continues to be poorly managed. Our research group has developed a native iPhone application (app) called Pain Squad to tackle the problem of poorly managed pain in the adolescent with cancer group. The app functions as an electronic pain diary and is unique in its ability to collect data on pain intensity, duration, location, and the impact pain has on an adolescent's life (ie, relationships, school work, sleep, mood). It also evaluates medications and other physical and psychological pain management strategies used. Users are prompted twice daily at configurable times to complete 20 questions characterizing their pain and the app transmits results to a database for aggregate reporting through a Web interface. Each diary entry represents a pain case filed by an adolescent with cancer and a reward system (ie, moving up through law-enforcement team ranks, built-in videotaped acknowledgements from fictitious officers) encourages consistent use of the diary.

Objective: Our objective was to design, develop, and test the usability, feasibility, compliance, and satisfaction of a game-based smartphone pain assessment tool for adolescents with cancer.

Methods: We used both low- and high-fidelity qualitative usability testing with qualitative semi-structured, audio-taped interviews and iterative cycles to design and refine the iPhone based Pain Squad app. Qualitative thematic analysis of interviews using constant comparative methodology captured emergent themes related to app usability. Content validity was assessed using question importance-rating surveys completed by participants. Compliance and satisfaction data were collected following a 2-week feasibility trial where users were alarmed to record their pain twice daily on the app.

Results: Thematic analysis of usability interviews showed the app to be appealing overall to adolescents. Analyses of both low- and high-fidelity testing resulted in minor revisions to the app to refine the theme and improve its usability. Adolescents resoundingly endorsed the game-based nature of the app and its virtual reward system. The importance of app pain diary questions was established by content validity analysis. Compliance with the app, assessed during feasibility testing, was high (mean 81%, SD 22%) and adolescents from this phase of the study found the app likeable, easy to use, and not bothersome to complete.

Conclusions: A multifaceted usability approach demonstrated how the Pain Squad app could be made more appealing to children and adolescents with cancer. The game-based nature and built-in reward system of the app was appealing to adolescents and may have resulted in the high compliance rates and satisfaction ratings observed during clinical feasibility testing.

(J Med Internet Res 2013;15(3):e51) doi:10.2196/jmir.2350

KEYWORDS

neoplasms; pain; child; adolescent; youth; cellular phone; game

Introduction

Pain is a common and significant burden to children and adolescents with cancer. This pain may result from cancer, the various cancer treatment methods, or associated procedures, and is known to lower quality of life compared to healthy children [1-6]. Adolescents with cancer (aged 10-18 years) reported that pain was the most distressing symptom of the cancer experience [7,8]. Poorly managed pain has been shown to be additionally problematic as it results in anxiety and distress related to subsequent treatment and negative long-term psychological effects [9,10].

Proper assessment is the essential first step to effectively manage pain experienced by children and adolescents with cancer. Following pain assessment, pain management strategies can be developed, refined, and evaluated to provide the best possible pain relief for adolescents. Still, several barriers prevent appropriate pain assessment for adolescents. There is no well-validated multidimensional tool evaluating the sensory (intensity, quality, and location of pain), affective (emotional effects), and evaluative (pain's interference with daily activities) dimensions of pain in adolescents with cancer [11]. In addition, existing tools have methodological problems and do not allow for longitudinal assessments in everyday settings [12,13].

Self-report in the form of paper-based approaches has been the most commonly used pain assessment modality [14]. However, patient recall of past experiences involves active reconstruction of events, which can lead to inaccuracies and biases in the reporting of the event [14]. Pain reports can also be influenced by a patient's affect, as well as the most salient and current state of pain. Finally, patients resort to hoarding (back filling) or completing paper-based pain diaries in longitudinal assessment diaries just prior to returning them to the research center, which reduces the accuracy of collected data [15].

Real-time data capture using electronically-based pain assessments represent a superior method for capturing patient pain data. Wireless mobile devices such as smartphones can: (1) conveniently collect pain reports in natural settings (ie, they are pocket-sized, have large memory capacity, and user-friendly interfaces), (2) be individualized through flexible programming, (3) create time- and date-stamped pain reports to circumvent back-filling, and (4) conveniently upload data to secure electronic databases for review by scientists and clinicians [15]. Recent studies in pediatric [16,17] and adult [18,19] populations have used Internet and mobile device modalities to effectively track and manage health conditions.

Adolescents have adopted mobile technology into their everyday lives in a much quicker and more engaging way than other generations. A recent report indicated that 75% of adolescents aged 12-17 years now own mobile phones [20]. Mobile technology is known to serve five major functions for adolescents - entertainment, information, communication, organization, and support. Of these functions, entertainment and information represent the most frequent uses of technology by adolescents [21]. Regarding the Apple smartphone devices specifically, a 2012 survey of 7700 teenagers found that 40% of teenagers owned iPhones, this was an increase from 23% of teens owning iPhones the year prior [22]. Capitalizing on the popularity of smartphone-based entertainment with adolescents, game-based mobile health applications (apps) may represent an important means to increase compliance with remote monitoring and treatment of health issues.

In the present study, we aimed to design and develop a smartphone-based pain diary app that would be engaging to adolescents with cancer through the gamification of pain assessment recordings. The gamification process involved adolescents playing the role of law-enforcement officers on a special investigative force that hunts down pain. Gamification also included a compliance-based rewards system (ie, proceeding up law-enforcement ranks and receiving videotaped acknowledgements for logging assessments by fictitious officers in the field). We sought to use a user-centered design approach where we actively engaged adolescents in all aspects of the research process from the app's inception (design and development) through usability and feasibility testing to refine the prototype.

Methods

Pain Assessment Questionnaire Development

Pain assessment questions were developed using the e-Ouch electronic juvenile idiopathic arthritis pain diary [23-25] as a template. Expert opinion (ie, captured during a meeting with 10 pediatric oncologists and 10 pediatric pain experts) was used to modify the arthritis diary questions to be cancer pain-specific. The following modifications were made: (1) the body-map was altered to include areas known to be commonly painful in cancer, (2) the addition of a list of possible pain-associated symptoms (eg, nausea), (3) lists of medications and other therapies were changed to include those commonly used to treat cancer pain, and (4) the addition of a list of possible pain sources (eg, treatments, procedures). The final pain assessment questionnaire consisted of 20 questions on the multidimensional (ie, sensory, affective, and evaluative) nature of pain, as well as questions related to pain management strategies used and their effectiveness.

Pain Squad App Design and Development

Pain Squad Design Principles

The design principles included: (1) a multidimensional pain diary to help adolescents track their pain and treatments that help and do not help to reduce pain, (2) a function to alert the research team if pain was moderate to severe for 2 consecutive entries, and (3) the integration of rewards and incentives into the system to sustain engagement with the tool for 2 weeks.

Pain Squad, a special police unit, was designed as an incentive for adolescents to use the app (Figure 1). Adolescents were recruited by a special investigative force whose mission was to "hunt down pain and put it behind bars". Twice a day, adolescents completed a pain-reporting mission by answering questions about pain intensity and location using a simple touchscreen interface. As mentioned, 20 questions about an adolescent's pain were asked. Questions took the form of: (1) touchable visual analogue slider scales to rate dimensions of pain (pain intensity, pain unpleasantness, pain interference, control over pain) from 0 to 10, (2) a selectable body-map to identify pain locations, (3) multiple choice questions about pain characteristics (duration of pain, causes of pain, pain management strategies used and their effectiveness.), (4) lists of selectable words describing pain and associated symptoms (not shown), and (5) a free-text question to collect additional information adolescents may wish to record (Figure 2). Filling out 3 reports in a row earned a promotion to the next rank in the squad. The promotions were delivered in the form of short videos featuring the stars of Canadian-filmed TV shows Flashpoint and Rookie Blue (Figure 3). The "badge/medal" and video earned by the adolescent through promotion could be re-watched if desired. Pain Squad is not a game in the strictest sense and does not have rules or objectives beyond participation. Instead, the software used features of games such as rewards and achievements. The design concept was refined during low-fidelity usability testing (outlined below).

Figure 1. Pain Squad app title screen.

Figure 2. Screenshots of Pain Squad app assessment functionality showing the visual analogue slider scale (top left), the selectable body-map (top right), a multiple-choice question (bottom left), and a free-text question (bottom right).

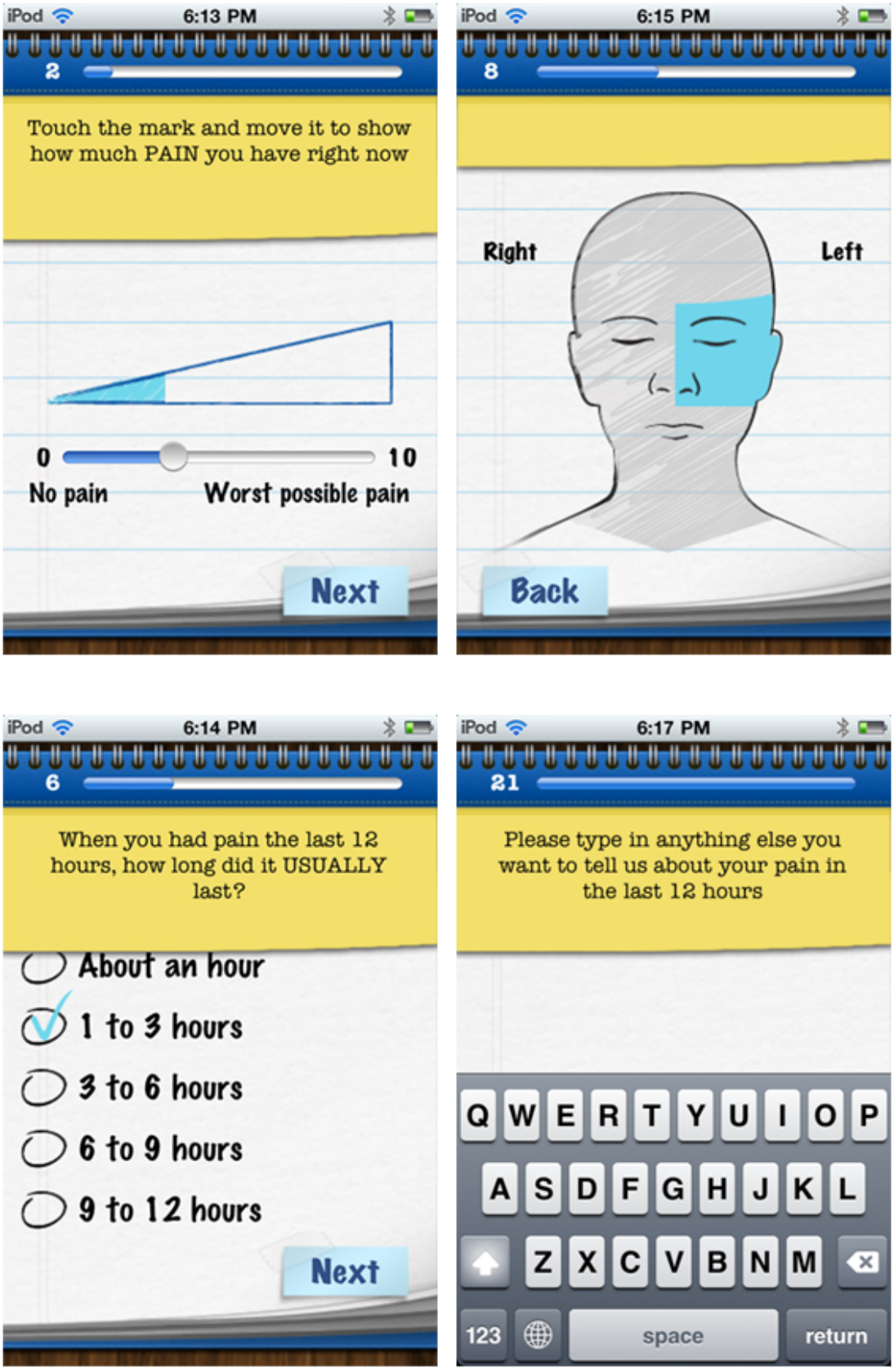


Figure 3. Screenshots of Pain Squad reward system showing compliance-based promotion scheme (left), and videotaped acknowledgement of the adolescent's compliance with completing pain assessments (right).

App Development

Once the overarching concept for the app and the app's design elements were developed in consultation with adolescents (ie, during low-fidelity usability testing), a prototype Pain Squad app was created for high-fidelity usability testing (outlined below). Pain Squad was programmed to be accessible to adolescents using an iPhone 4S. The app was built using client-server architecture. The client component was built in native Apple iOS code and the server component used iBATIS.NET to broker data exchange and an MS SQL Server 2008 database to manage incoming pain assessment data. The system also contained a password-protected Web-interface allowing researchers to create studies, add new study participants, and access results.

Pain Squad used iPhone audible alerts to notify adolescents to complete pain assessments. The system generated 3 alerts for each pain assessment to be completed. The timing of these alerts was tailored to the needs of adolescents by allowing adolescents to choose alarm times from a number of preset options. Data entered by adolescents were stored on the iPhone in an SQLLite database and communicated to the server whenever the application was online. Pain assessment data were not accessible via the iPhone and were transmitted to the server using an encrypted (HTTPS) protocol over a Secure Socket Layer (SSL) connection. The server was hosted at the tertiary care center behind a firewall in a network secure environment. A user name and password were required to access data.

Participant Characteristics

Adolescents involved in the usability and feasibility testing were recruited from one large university-affiliated tertiary hematology/oncology care center in Toronto, Canada over a 1-year period in 2011-2012. Children and adolescents were eligible to participate if they: (1) were able to speak and read English, (2) were 9 to 18 years old, (3) were diagnosed with cancer, (4) were being seen on an in- or out-patient basis by the oncology team, and (5) had self-reported pain at least once in the last week. Adolescents were excluded if they had severe cognitive impairments or major co-morbid medical or psychiatric illnesses that would preclude pain assessment by self-report. Consent was obtained prior to involvement in any phase of the study and all adolescents completed a brief questionnaire on their demographic characteristics. The research assistants gathered additional demographic and disease-related data from medical charts.

Phase 1: Usability Testing

Phase 1a : Low-Fidelity User-Centered Design

Following the development of the cancer pain assessment questionnaire, app interface designs were trialed with adolescents. A qualitative usability testing approach was used with 3 iterative cycles of semi-structured audiotaped interviews. Fifteen adolescents (5 in each of 3 cycles) were shown paper screenshots of the pain assessment application and were asked what they liked and disliked about the design, premise of the game, and question format. The list of questions was modified during the course of the interview process in light of emerging themes and field notes related to perceived ease of use.

Technical problems with the app were recorded by the research assistant. Adolescents were further asked to provide suggestions for improvement. Design elements were modified and new paper screenshots were generated and tested until no further changes were suggested.

Phase 1b: High-Fidelity User-Centered Design

Following the development of a fully functional iPhone-based Pain Squad prototype, usability testing (2 iterative cycles) with semi-structured audiotaped interviews was again conducted with 18 adolescents with cancer (21 adolescents were approached to participate and 3 declined). In this phase, a research assistant first provided adolescents with a brief (approximately 5 minutes) demonstration of the Pain Squad app on the iPhone using a standardized pain vignette. Adolescents were then asked to complete the app, recording their own pain while thinking aloud about likes, dislikes, and difficulties with the app. At the end of each session, a research assistant asked a series of standardized open-ended questions related to ease of use, and what adolescents liked or disliked about the app. The research assistant recorded the answers to questions that explored emerging themes and field notes on ease of app use. After the first iterative cycle, changes were made based on themes identified from adolescent opinion. A second iterative cycle was conducted with 8 adolescents, which generated no further recommendations for changes to the app.

Adolescents participating in the high-fidelity testing also rated the content validity of the pain diary questions. To do so, adolescents used a 4-point Likert scale to rate each question from “not important at all” to “very important”.

Data Analysis for Usability Testing

In both low- and high-fidelity testing, demographic data were analyzed using SAS version 9.1.3 [26]. Audiotaped usability interviews were transcribed verbatim. All transcripts from the usability testing phases were verified against the tapes and imported into NVivo 8.0 [27] for coding. Field notes taken during the interviews were also transcribed and included in the analytic process. Using grounded theory latent coding [28], data were coded according to the study objective and categorized to reflect emerging themes. Changes to the prototype were made based on feedback from each iterative cycle of testing. In addition, thematic analysis of high-fidelity usability interviews was performed according to adolescent age group (9-12 years and 13-18 years) to identify any age-specific differences in opinion. This analysis was performed in the same manner described above.

Phase 2: Clinical Feasibility Testing

Overview

Following usability testing, a clinical feasibility test was conducted with adolescents with cancer to determine compliance and satisfaction with the app. For this phase, adolescents were trained on app use with standardized pain vignettes. Fourteen adolescents were then loaned an iPhone 4S and asked to complete Pain Squad pain assessments twice daily (in the morning and the evening) for 14 days. Telephone assistance was available to adolescents in case of technical problems. The research team reviewed a summary of each adolescent’s pain reports daily so that patient safety issues could be identified and resolved.

On day 15, adolescents were prompted by the phone using an audible alert to complete the Pain Squad Evaluation Questionnaire on the iPhone, which ascertained likes and dislikes with the Pain Squad app. The Pain Squad Evaluation Questionnaire involved multiple 4-point Likert scales. As an example, selectable options for the question, “How much did you like using the pain diary?” were: “very much liked it”, “liked it okay”, “did not like it or disliked it”, and “did not like it at all”. Data generated from these surveys were used to establish satisfaction and ease-of-use of Pain Squad, the degree to which the app interfered with daily activities as well as information on how long adolescents would be willing to use the diary. The Evaluation Questionnaire also included a free-text question where adolescents were encouraged to enter any other information about the diary they felt important to discuss.

Clinical Feasibility Testing Data Analysis

Compliance was defined as 100% when 28/28 entries were completed within the 2-week period. Compliance and satisfaction were analyzed using SAS version 9.1.3 [26]. The level of significance was set at $P<.05$. Comparisons between compliance data generated during feasibility testing were made using *t* tests.

Results

Phase 1: Usability Testing

Participant Characteristics

Demographic and disease characteristics of adolescents included in the low- and high-fidelity user-centered testing are shown in Table 1. The age of participants in each phase of testing was similar (approximately 13 years). Both male and female participants were included in all phases of the study and participants also had a variety of cancer diagnoses. Time since diagnosis was on average less than 2 years for participants in each study phase.

Table 1. Demographics and disease characteristics of adolescents included in low- and high-fidelity usability testing, content validity testing, and feasibility testing.

	Phase 1a: Low-fidelity usability testing (n=15)		Phase 1b: High-fidelity usability and content validity testing (n=18)		Phase 2: Feasibility (n=14)	
Characteristics	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)
Age (years) [a]						
	13.9 (1.9)		13.4 (2.9)		13.2 (2.3)	
Gender						
Female		7 (47)		9 (50)		9 (64)
Male		8 (53)		9 (50)		5 (36)
Primary diagnosis						
ALL		4 (27)		6 (33)		7 (50)
AML		3 (20)		0 (0)		2 (14)
Ewing's Sarcoma		1 (7)		3 (17)		1 (7)
Non-Hodgkin's Lymphoma		2 (13)		1 (6)		2 (14)
Osteosarcoma		2 (13)		4 (22)		1 (7)
Rhabdomyosarcoma		1 (7)		2 (11)		0 (0)
Other		2 (13)		2 (11)		1 (7)
Patient type [a]						
Inpatient		8 (53)		15 (83)		6 (43)
Outpatient		7 (47)		3 (17)		8 (57)
Duration of illness (years) [a]						
	0.8 (0.8)		1.4 (4.2)		0.8 (0.7)	

[a]at the time of study recruitment

Phase 1a: Low-Fidelity User-Centered Design Phase

Overview

Transcript analysis from low-fidelity testing revealed 4 distinct themes, which guided the development of the game-based app: (1) a need to change the game theme, (2) the importance of the reward-system, (3) an appreciation for the aesthetics of the app, and (4) changes required to improve ease of use.

Change to the Game Theme

During low-fidelity usability testing, and before a capital investment in programming was made, we endeavored to find a theme for the app that was well-liked by adolescents. Adolescents with cancer had difficulty relating to the original game theme and expressed the need to change the game to improve app desirability. The original theme of the game centered around a private detective agency. According to the game, this agency was known as "Gum Shoe" and investigated pain cases. The gumshoe concept fared poorly in usability trials and was considered difficult to relate to by adolescents. In fact, the lack of understanding around the term gumshoe was the most strongly endorsed message in the first iteration of low-fidelity testing. Indeed, none of the adolescents interviewed knew what gumshoe meant. This was articulated by an adolescent who discussed unfamiliarity with the concept.

> *I don't know what gumshoe is. I would guess that gumshoe means…just gum on a shoe?* [Female, 13 years]

To address this issue, the game theme was changed to "Pain Squad". Pain Squad took a more modern approach by using a "pain detective" concept and saw adolescents playing the role of law-enforcement officers (Figure 4). This new concept was trialed with adolescents in the second iteration of low-fidelity testing and received overwhelmingly positive endorsements. No further changes to the design theme were required.

Figure 4. App prototype home screens: the Gum Shoe home screen depicting a detective case file (left), and the Pain Squad home screen depicting the desk of a law enforcement officer (right). Tapping "Start Survey" or "Start Case" begins the pain assessment. Tapping "Accomplishments" or "Rewards" allows review of rank and access to videotaped acknowledgements.

Importance of the Reward System

Embedded in the Pain Squad app design was a reward system to encourage pain assessment competition. This reward system was highly regarded by adolescents as illustrated below.

> *I think that [the reward system] is a really good idea! That makes it way more enjoyable. I really like the moving up…idea and the rewards.* [Female, 16 years]

Appreciation for the Game Esthetics

The color scheme, fonts, and graphics used in the Pain Squad app were considered attractive to adolescents, helping to create interest in the app. During low-fidelity testing, adolescents made several positive comments regarding the appearance of the app screenshots:

> *Green, blue, and yellow are really good colors for this app, for a detective. I like the layout, it's cool. And I like the font.* [Female, 13 years]

> *I would just scroll through and tap on the screen. I like how there are the different choices and the colors. That's really cool!* [Female, 12 years]

Changes Required to Improve Ease of Use

Based on the screenshots shown, adolescents suggested several changes to improve usability of the game-based app. These changes included adding verbal and numerical anchors to pain assessment scales to guide pain ratings (eg, "no pain" and "worst possible pain", 0 and 10 respectively) and providing clarification on body parts in the pain location body map. The quote below demonstrates one adolescent's recommendation that the torso and the abdomen should be combined.

> *These two parts aren't clear. I don't even know what part of the body that is! It's better to have them together. It's easier to understand.* [Female, 16 years]

These recommendations were addressed before proceeding to the high-fidelity testing as seen in Figures 5 and 6.

Figure 5. Screenshots of Pain Squad app visual analog scales showing before (left) and after (right) verbal and numerical rating anchors were added.

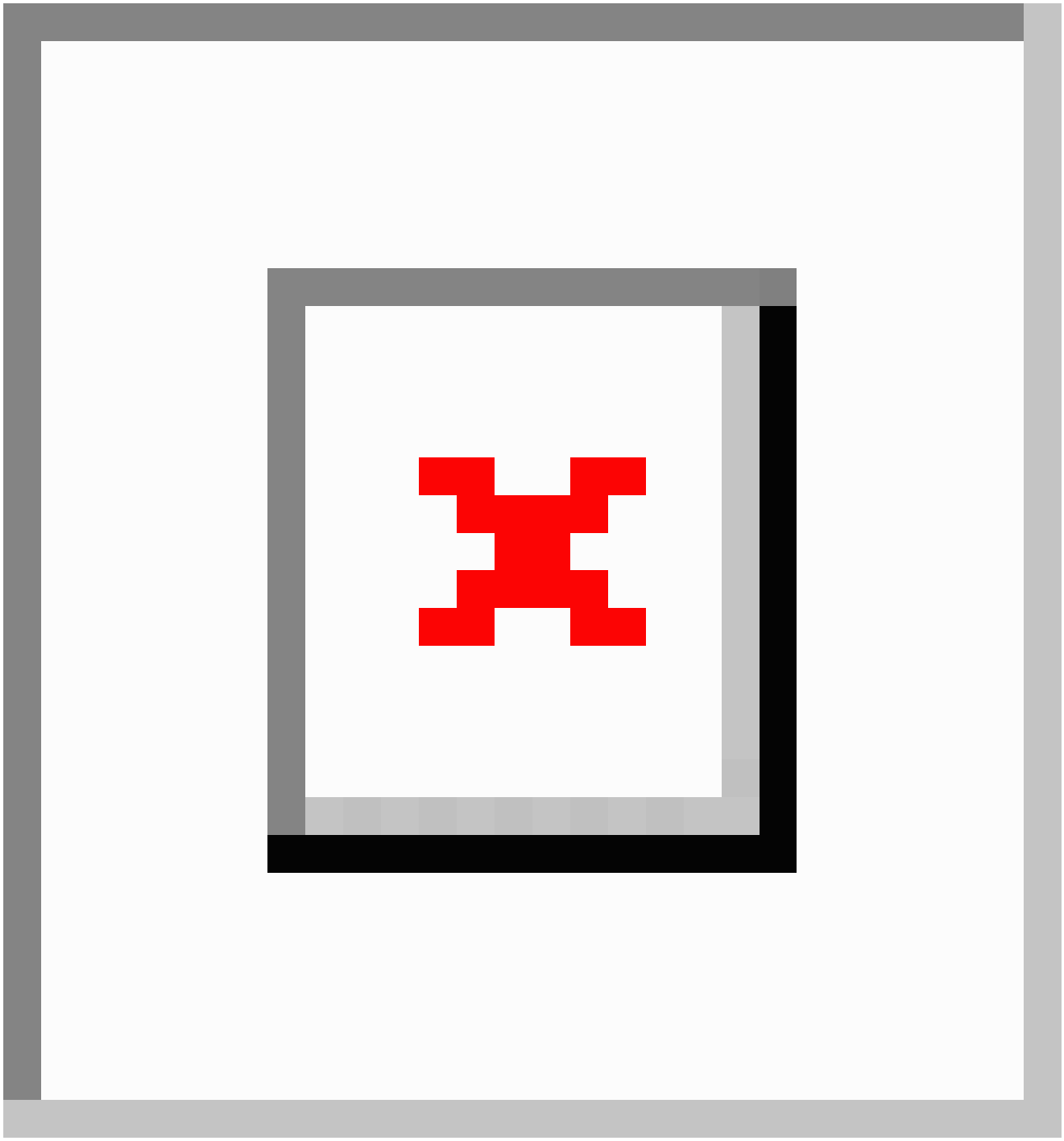
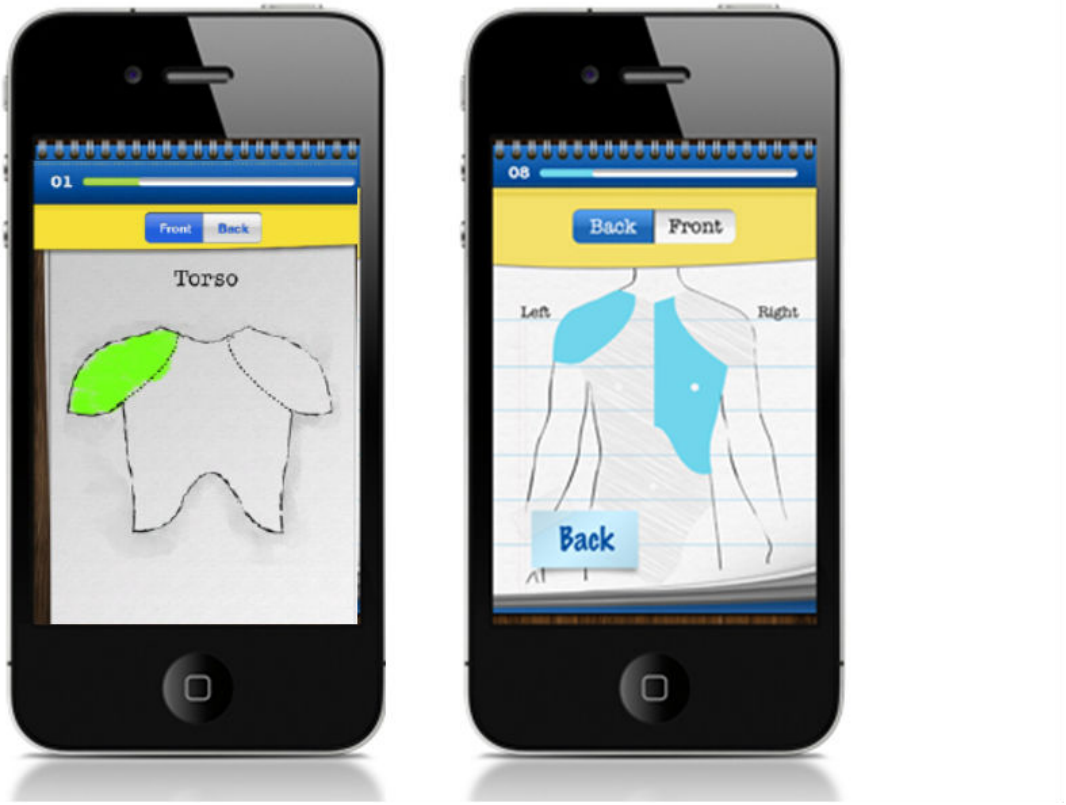


Figure 6. Screenshots of Pain Squad app "body map" showing before (left) and after (right) body parts were combined as recommended by adolescents.



Phase 1b: High-Fidelity User-Centered Design Phase and Content Validity Testing

Overview

Four themes highlighting the usability of the game-based app were identified through analysis of high-fidelity usability interviews. These themes were: (1) general endorsement of the app, (2) appeal of the game and rewards aspects of the app, (3) ease-of-use of app, and (4) recommendations for improving the user-interface of the app. We found no differences in endorsed themes between younger (9-12 year olds) and older (13-18 year olds) adolescents. These themes therefore represent the convergent perspectives of study participants.

General Endorsement of the App

Adolescents overwhelmingly endorsed the app as appealing and fun to use. Every adolescent interviewed said they would use the app daily for an extended period (ie, 2 weeks) if given the option. This satisfaction with the app is illustrated in the following quote from an adolescent with cancer:

> *Oh cool! It's so interesting…I liked that you could put down how you felt. Instead of talking to somebody you could just put down how you feel…It was really simple. I liked everything. I think that one, I'd add to my phone!* [Female, 11 years]

Appeal of the Game and Rewards Aspects of the App

The game-like features of the app and the rewards program were also seen as appealing to adolescents with cancer. Adolescents appreciated the ability to role-play as a detective and advance through the law enforcement team ranks.

> *Oh that's cool! I liked the way your like a spy trying to solve a case because it made it more interesting.* [Male, 10 years]
>
> *I like that you can see what rank you are and what you need to do [to advance].* [Female, 11 years]

Ease-of-Use of App

All adolescents interviewed also endorsed the ease-of-use of the app. All adolescents found the app to be "easy to understand" and "easy to navigate". Adolescents also discussed their own familiarly with using electronic devices such as smartphones and did not foresee themselves having any problems using Pain Squad.

Recommendations for Improving User Interface of the App

As in the low-fidelity testing, specific suggestions for improvements to the app were made by adolescents. Adolescents in the first iteration of high-fidelity usability testing found that the wording of some of the questions was confusing. For instance, one adolescent did not understand the question, "How much control do you feel you have over your pain?"

> *It sort of doesn't really make sense. Like how much control you had? How could you like really have like the control? So I didn't really like that one.* [Male, 9 years]

In the first iteration, 4 out of 10 adolescents also discussed the need to be able to select more specific areas on the body location map.

> *Umm the questions were pretty accurate, though I didn't like the body map because it wouldn't let you select exact body parts [in pain].* [Male, 11 years]

The final first-iteration recommendation by adolescents involved the placement of selectable buttons on the screen. Adolescents found it too easy to advance through questions by accidently tapping on the "Next" button before having time to decide on their answer.

These issues were addressed by further sub-dividing the body map such that more specific body areas could be selected (Figure 7), simplifying question wording and moving the "Next" button to prevent accidental tapping. No further recommendations for changes were made in the second iteration of testing.

Figure 7. Screenshots of Pain Squad app "body map" showing before (left) and after (right) specificity of selectable pain location was improved.

Content Validity Testing

Content validity testing affirmed the importance of pain assessment questions for adolescents with cancer. Results indicated that 88% of questions were rated as "important" or "very important" by the majority (>50%) of adolescents. Questions that did not achieve this rating related to the impact of pain on: (1) things adolescents did, (2) schoolwork, and (3) relationships with friends and family.

Phase 2: Clinical Feasibility Testing

Participant and Pain Characteristics

Demographic and disease characteristics are shown in Table 1. Data collected from the feasibility testing provided data on the pain experience of adolescents with cancer. These data showed that average pain reported by an adolescent with cancer at the time of completing the diary was 2.3/10 (SD 1.8). Pain was present for 14% (2/14) of adolescents on every pain assessment completed. Additionally, average worst pain in the 12 hours preceding a reported pain episode was 5.8/10 (SD 1.3). Pain reports received were reviewed by the research team. The team did not need to intervene with any adolescents participants for safety reasons.

Compliance

Rates of app compliance were high, with a mean of 81% (SD 22%) of pain assessments completed by adolescents. The relatively large variability (22%) in assessment completion was due in large part to low rates of completion by 2 adolescents. One adolescent completed only 14% (4/28) of the Pain Squad entries and cited that she forgot to keep the iPhone with her as the reason for non-compliance. Another adolescent was hospitalized for an emergency medical complication during the feasibility trial and did not bring the iPhone with her to the hospital. Upon hospital discharge she resumed Pain Squad assessments, completing a total of 57% (16/28) of entries. Average compliance with these 2 adolescents excluded from the analyses was 88% (SD 8%).

There were no differences in mean compliance with assessments between morning and evening reports (82% SD 23% vs 80% SD 23%, P=.77) or between week days and weekends (80% SD 24% vs 82% SD 23%, P=.84). In addition, compliance with Pain Squad remained high for the duration of the feasibility trial as no difference in assessment completion rates was seen between week 1 and week 2 (84% SD 20% vs 78% SD 29%; P=.55). Sub-group analyses of the feasibility trial cohort showed

no difference in compliance rates according to gender (females: 79% SD 28%, males: 84% SD 8%, *P*=.68) or initial treatment location (in-patient: 84.5% SD 15.6%, out-patient: 78% SD 27%, *P*=.59).

Satisfaction

Data from the post-study Pain Squad Evaluation Questionnaire showed that adolescents enjoyed using the game-based app, found it easy to use and found that it did not interfere with their daily activities. When asked about how much they liked using Pain Squad, 86% of adolescents (12/14) indicated that they "very much liked it" or "liked it okay" and only 14% (2/14) indicated that they "didn't like it" or "didn't like it at all". Regarding the overall appearance of the app, 64% of adolescents (9/14) "very much liked it" and 36% (5/14) "liked it okay". No adolescents surveyed indicated that they did not like the app design. Regarding ease-of-use, 79% of adolescents (11/14) found the app "very easy" or "easy" to fill out twice a day and 14% (2/14) found it "a little hard" to do. The majority of adolescents also indicated that completing Pain Squad did not affect activities and friendships, with only one participant (7% of respondents) stating that the app "interfered a lot" with these activities. The majority of adolescents (79%; 11/14) reported that 2 weeks was an appropriate amount of time to use the app, but longer periods of usage (ie, greater than 6 weeks) were endorsed by the remainder of respondents.

Discussion

We present the design and development of a game-based mHealth app to routinely assess pain in adolescents with cancer. We further present data on compliance with completing assessments collected during a 2-week clinical feasibility study. Low-fidelity, user-centered design led to the identification of needed changes to improve the game-based premise of the app and its ease of use (ie, changes to body map, inclusion of anchors to assessment scales). The high-fidelity, user-centered design approach led to the identification of additional needed revisions to the app, but also led to an appreciation of adolescents' positive reactions to the novel approach of gamification used in this app (video clips from officers in field and graduated reward system). This phase further affirmed the content validity of the pain assessment questions. The feasibility trial demonstrated high rates of compliance and overall satisfaction with the app.

To our knowledge, the present study was the first to report on a user-centered approach to the design and development of a smartphone-based multidimensional pain assessment app for adolescents with cancer. A report on the user-centered design of an mHealth cancer symptom assessment diary for adolescents has been previously published [29]. This study however, did not include adolescents in the design of the interface and did not take a multidimensional approach to the assessment of pain. Our study has used multiple iterative usability cycles (both low- and high-fidelity) to refine the Pain Squad app theme and design. Elucidating the opinions, insights, and recommendations of adolescents from the early stages in app design is critical to ensure a pain assessment tool that is well utilized and clinically helpful. We have also established the content validity of the pain assessment questionnaire through question-importance ratings by end users.

This study was also the first to report on the development and preliminary testing of a game-based mHealth pain assessment tool. We have gamified the concept of pain assessment by: (1) having adolescents with cancer role-playing as officers on the Pain Squad, and (2) rewarding compliance with assessment completion with videotaped acknowledgements and the ability to advance through the law-enforcement ranks. Given the popularity of smartphones with adolescents and the entertainment value of these devices, we reasoned that gamifying health information collection (ie, pain assessment data) using iPhones might improve compliance with reporting. The gamification of the app was highly endorsed by adolescents in both the low- and high-fidelity user-centered design phases. Adolescents helped to refine the game premise, positively rated app aesthetics (eg, colour scheme, font, graphics) and resoundingly endorsed the rewards component of the game.

The clinical feasibility examination showed high compliance rates with the diary irrespective of time of day (ie, morning vs evening) or week (weekday vs weekend). Compliance was also sustained for the duration of the trial, as there were no observed compliance differences between week 1 and week 2. Sub-group analyses also showed no difference in compliance rates between male and female participants or in- and out-patient participants. A systematic review of 62 publications on electronic pain assessment diaries in both children and adults showed compliance rates to be generally high [30]. The authors of the review used backward regression analysis to determine predictors of high compliance and found compliance to be positively associated with shorter diaries, older age of participants, having access to a user's manual, alarmed reminders, and financial compensation. Additionally important to consider is the clinical context of our study. All of the participants in the feasibility trial were actively being treated for cancer. Cancer treatment is known to be physically exhausting for adolescents [31,32] and could theoretically have adversely effected pain diary completion. Despite this potential adverse impact on reporting, the adolescents completed a large majority of pain assessments on a regular basis.

We propose several possible reasons for the high rates of compliance observed in the present study. First, audible tailored alarms were used to signal participants to complete pain assessments. Audible alarms have been used in previous mHealth assessment studies with adolescents and these studies have also reported high compliance rates [25,33-36]. These alarms act as a direct reminder to participants to log health information. Secondly, we propose that the gamification of the pain assessment app motivated adolescents to complete assessments. Gamification involves the use of videogame components in non-gaming systems to improve user satisfaction [37]. Several previous studies have used reward programs to encourage compliance with mHealth interventions by adolescents [33-35]. However, the reward programming in these studies involved external motivators (ie, money or gift cards given for completing reports). In the present study, we used no such incentive to encourage compliance but observed compliance rates similar to those previous studies. We

hypothesize that because of the adolescents' endorsement of the Pain Squad game premise and virtual rewards system, the app was internally motivating to complete. This is important because the financial feasibility of supplying patients with monetary rewards to log health information, especially over extended periods of time, is limited. Novel mechanisms, such as promotions to the next law-enforcement rank, could inspire improved compliance with health reporting and may ultimately be more practical for use by researchers, clinicians and the health system at large.

As the diary was used twice-daily for 2 weeks, we cannot comment directly on how compliance might be impacted if the app were used for a longer time period. Given that compliance did not significantly change from week 1 to week 2 and previously observed high rates of compliance over weeks to months [30], high compliance rates may be maintained for longer periods. Our data from the Evaluation Questionnaire also indicated that some adolescents were willing to use the app for longer periods (ie, potentially for a period longer than 6 weeks). However, a direct examination of the impact of gamification on long-term compliance with pain diaries remains to be conducted.

Our clinical feasibility test also showed that adolescents with cancer enjoyed using the app over the 2-week period, found it to be attractive and easy to use and felt that it did not interfere with their daily activities. These results are important as they relate to pain management in adolescents with cancer. Because of the complexity of pain management in this group [38-41], accurate and throughout pain assessment is required. Multidimensional pain assessment over extended periods of time (ie, weeks) and multiple times during the day is therefore needed. An electronic pain assessment diary that is well-liked, attractive, and not a burden to complete can assist in collecting important data from adolescents with cancer allowing for improved pain management.

Our study had several limitations that may have tempered interpretation of results and conclusions. First and importantly, we did not directly examine the causes of our observed high compliance rates and could therefore only speculate that they resulted from audible alarms and the gamification of pain assessment. Similarly, should audible alarms and gamification improve compliance, we do not know the degree to which each factor impacted on end user engagement. In addition, our user samples for all phases were drawn from the oncology program at one pediatric tertiary care center, thereby limiting the generalizability of our results. However, with samples as small as 5 end users per iterative usability cycle, the majority of usability problems can be identified [42]. We interviewed each user only once during usability testing and did not verify the results of our thematic analysis with adolescents. Finally, as the Pain Squad app was programmed in Apple iOS code, it was accessible only via Apple devices (ie, iPhone, iPad, iTouch). We are currently exploring options to improve the accessibility of Pain Squad by having it programmed for use on other platforms (eg, Blackberry, Android).

In conclusion, we have demonstrated a user-centered approach to the design and development (including the establishment of content validity) of a novel mHealth pain assessment tool for adolescents with cancer. We also report high rates of compliance and satisfaction with the app as demonstrated in a 2-week clinical feasibility test. We propose that the gamification of the app may have positively impacted compliance and satisfaction and suggest that the use of internal motivators, such as virtual reward systems, is an important area for future mHealth research. Multi-site testing of the validity (including responsiveness) and reliability of the Pain Squad electronic pain diary is currently underway. In addition, future research by our group will focus on using the Pain Squad app as a platform for a clinical decision support system to aid adolescents in making pain management choices. A valid and reliable electronic diary with pain management capabilities has the capacity to result in improved pain management, and ultimately improve quality of life for adolescents with cancer.

Acknowledgments

Funding for this study was provided by the C^{17} Children's Cancer & Blood Disorders Research Network.

Conflicts of Interest

None declared.

References

1. Collins JJ, Stevens MM, Berde CB. Pediatric cancer pain. In: Clinical pain management. London: Hodder Arnold; 2008.
2. Cheng KK, Lee V, Li CH, Yuen HL, Epstein JB. Oral mucositis in pediatric and adolescent patients undergoing chemotherapy: the impact of symptoms on quality of life. Support Care Cancer 2012 Oct;20(10):2335-2342. [doi: 10.1007/s00520-011-1343-1] [Medline: 22167295]
3. Jacob E, Hesselgrave J, Sambuco G, Hockenberry M. Variations in pain, sleep, and activity during hospitalization in children with cancer. J Pediatr Oncol Nurs 2007;24(4):208-219. [doi: 10.1177/1043454207299875] [Medline: 17588893]
4. Landolt MA, Vollrath M, Niggli FK, Gnehm HE, Sennhauser FH. Health-related quality of life in children with newly diagnosed cancer: a one year follow-up study. Health Qual Life Outcomes 2006;4:63 [FREE Full text] [doi: 10.1186/1477-7525-4-63] [Medline: 16987419]

5. Ljungman G, Gordh T, Sörensen S, Kreuger A. Pain variations during cancer treatment in children: a descriptive survey. Pediatr Hematol Oncol 2000;17(3):211-221. [doi: 10.1080/088800100276389] [Medline: 10779987]
6. Special Interest Group on Pain in Childhood, International Association for the Study of Pain. Cancer pain relief and palliative care in children. Geneva: World Health Organization; 1998.
7. Collins JJ, Byrnes ME, Dunkel IJ, Lapin J, Nadel T, Thaler HT, et al. The measurement of symptoms in children with cancer. J Pain Symptom Manage 2000 May;19(5):363-377. [Medline: 10869877]
8. Moody K, Meyer M, Mancuso CA, Charlson M, Robbins L. Exploring concerns of children with cancer. Support Care Cancer 2006 Sep;14(9):960-966. [doi: 10.1007/s00520-006-0024-y] [Medline: 16639553]
9. Chen E, Zeltzer LK, Craske MG, Katz ER. Children's memories for painful cancer treatment procedures: implications for distress. Child Dev 2000;71(4):933-947. [Medline: 11016557]
10. von Baeyer CL, Marche TA, Rocha EM, Salmon K. Children's memory for pain: overview and implications for practice. J Pain 2004 Jun;5(5):241-249. [doi: 10.1016/j.jpain.2004.05.001] [Medline: 15219255]
11. Anderson KO. Assessment tools for the evaluation of pain in the oncology patient. Curr Pain Headache Rep 2007 Aug;11(4):259-264. [Medline: 17686388]
12. Palermo TM. Assessment of chronic pain in children: current status and emerging topics. Pain Res Manag 2009;14(1):21-26 [FREE Full text] [Medline: 19262912]
13. Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. Pain 2006 Nov;125(1-2):143-157. [doi: 10.1016/j.pain.2006.05.006] [Medline: 16777328]
14. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials 2003 Apr;24(2):182-199. [Medline: 12689739]
15. Stone AA, Broderick JE. Real-time data collection for pain: appraisal and current status. Pain Med 2007 Oct;8 Suppl 3:S85-S93. [doi: 10.1111/j.1526-4637.2007.00372.x] [Medline: 17877531]
16. Jan RL, Wang JY, Huang MC, Tseng SM, Su HJ, Liu LF. An Internet-based interactive telemonitoring system for improving childhood asthma outcomes in Taiwan. Telemed J E Health 2007 Jun;13(3):257-268. [doi: 10.1089/tmj.2006.0053] [Medline: 17603828]
17. Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. Diabetes Technol Ther 2009 Feb;11(2):99-106. [doi: 10.1089/dia.2008.0022] [Medline: 19848576]
18. Mulvaney SA, Anders S, Smith AK, Pittel EJ, Johnson KB. A pilot test of a tailored mobile and Web-based diabetes messaging system for adolescents. J Telemed Telecare 2012 Mar;18(2):115-118. [doi: 10.1258/jtt.2011.111006] [Medline: 22383802]
19. Heinonen R, Luoto R, Lindfors P, Nygård CH. Usability and feasibility of mobile phone diaries in an experimental physical exercise study. Telemed J E Health 2012 Mar;18(2):115-119. [doi: 10.1089/tmj.2011.0087] [Medline: 22283356]
20. Lenhart A, Ling R, Campbell S, Purcell K. Teens and mobile phones. Washington, DC: Pew Internet & American Life Project; 2010. URL: http://pewinternet.org/~/media//Files/Reports/2010/PIP-Teens-and-Mobile-2010-with-topline.pdf [accessed 2013-02-25] [WebCite Cache ID 6EhcGFi8t]
21. Skinner H, Biscope S, Poland B, Goldberg E. How adolescents use technology for health information: implications for health professionals from focus group studies. J Med Internet Res 2003 Dec 18;5(4):e32 [FREE Full text] [doi: 10.2196/jmir.5.4.e32] [Medline: 14713660]
22. Munster G. Device ownership among teens. A Piper Jaffray and Company Report 2012.
23. Stinson JN, Petroz GC, Tait G, Feldman BM, Streiner D, McGrath PJ, et al. e-Ouch: usability testing of an electronic chronic pain diary for adolescents with arthritis. Clin J Pain 2006;22(3):295-305. [doi: 10.1097/01.ajp.0000173371.54579.31] [Medline: 16514331]
24. Stinson JN, Petroz GC, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, et al. Working out the kinks: testing the feasibility of an electronic pain diary for adolescents with arthritis. Pain Res Manag 2008;13(5):375-382 [FREE Full text] [Medline: 18958308]
25. Stinson JN, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, Dupuis A, et al. Construct validity of a multidimensional electronic pain diary for adolescents with arthritis. Pain 2008 Jun;136(3):281-292. [doi: 10.1016/j.pain.2007.07.002] [Medline: 17723279]
26. SAS Institute I. SAS Version 9.1.3. Cary NC: SAS Institute Inc; 2006.
27. QRS International. QSR NVivo8. In: QRS NVivo 8. Australia: QRS International Pty Limited; 2008.
28. Roberts CH. A matter of definition. In: Text analysis for the social sciences: methods for drawing statistical inferences from texts and transcripts. Mahwah NJ: Lawrence Erlbaum; 1997.
29. Gibson F, Aldiss S, Taylor RM, Maguire R, McCann L, Sage M, et al. Utilization of the Medical Research Council evaluation framework in the development of technology for symptom management: the ASyMS-YG Study. Cancer Nurs 2010;33(5):343-352. [doi: 10.1097/NCC.0b013e3181cb4bad] [Medline: 20495448]
30. Morren M, van Dulmen S, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: a systematic review. Eur J Pain 2009 Apr;13(4):354-365. [doi: 10.1016/j.ejpain.2008.05.010] [Medline: 18603458]

31. Erickson JM. Fatigue in adolescents with cancer: a review of the literature. Clin J Oncol Nurs 2004 Apr;8(2):139-145. [doi: 10.1188/04.CJON.139-145] [Medline: 15108414]
32. Gibson F, Mulhall AB, Richardson A, Edwards JL, Ream E, Sepion BJ. A phenomenologic study of fatigue in adolescents receiving treatment for cancer. Oncol Nurs Forum 2005 May;32(3):651-660. [doi: 10.1188/05.ONF.651-660] [Medline: 15897939]
33. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
34. Connelly M, Anthony KK, Sarniak R, Bromberg MH, Gil KM, Schanberg LE. Parent pain responses as predictors of daily activities and mood in children with juvenile idiopathic arthritis: the utility of electronic diaries. J Pain Symptom Manage 2010 Mar;39(3):579-590. [doi: 10.1016/j.jpainsymman.2009.07.013] [Medline: 20303032]
35. Connelly M, Bickel J. An electronic daily diary process study of stress and health behavior triggers of primary headaches in children. J Pediatr Psychol 2011 Sep;36(8):852-862 [FREE Full text] [doi: 10.1093/jpepsy/jsr017] [Medline: 21441572]
36. Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability. Pain 2004 Feb;107(3):213-219. [Medline: 14736583]
37. Deterding S, Sicart M, Nacke L, O'Hara KDixon D. Gamification: using game-design elements in non-gaming context. 2011 May Presented at: 2011 annual conference extended actracts on human factors in computing systems; 2011; Vancouver, British Columbia p. 7-12.
38. Ameringer S. Barriers to pain management among adolescents with cancer. Pain Manag Nurs 2010 Dec;11(4):224-233 [FREE Full text] [doi: 10.1016/j.pmn.2009.05.006] [Medline: 21095597]
39. Collins JJ, Devine TD, Dick GS, Johnson EA, Kilham HA, Pinkerton CR, et al. The measurement of symptoms in young children with cancer: the validation of the Memorial Symptom Assessment Scale in children aged 7-12. J Pain Symptom Manage 2002 Jan;23(1):10-16. [Medline: 11779663]
40. Mercadante S. Cancer pain management in children. Palliat Med 2004 Oct;18(7):654-662. [Medline: 15540676]
41. Woodgate RL, Degner LF, Yanofsky R. A different perspective to approaching cancer symptoms in children. J Pain Symptom Manage 2003 Sep;26(3):800-817. [Medline: 12967729]
42. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]

Abbreviations

App: application
SSL: Secure Socket Layer

Edited by G Eysenbach; submitted 13.09.12; peer-reviewed by J Mooney, C Chau; comments to author 09.11.12; revised version received 26.11.12; accepted 13.02.13; published 08.03.13.

Please cite as:
Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, Gerstle JT, Alman B, Hopyan S, Strahlendorf C, Portwine C, Johnston DL, Orr M
Development and Testing of a Multidimensional iPhone Pain Assessment Application for Adolescents with Cancer
J Med Internet Res 2013;15(3):e51
URL: http://www.jmir.org/2013/3/e51/
doi:10.2196/jmir.2350
PMID:23475457

Original Paper

A Smartphone-Based Intervention With Diaries and Therapist Feedback to Reduce Catastrophizing and Increase Functioning in Women With Chronic Widespread Pain. Part 2: 11-month Follow-up Results of a Randomized Trial

Ólöf Birna Kristjánsdóttir[1,2], RN, MSc; Egil A Fors[3,4], MD, PhD; Erlend Eide[2,5]; Arnstein Finset[1], MApsych, PhD; Tonje Lauritzen Stensrud[1], MApsych; Sandra van Dulmen[6,7,8], MApsych, PhD; Sigrid Hørven Wigers[9], MD, PhD; Hilde Eide[2,8], RN, MApsych, PhD

[1]Department of Behavioral Sciences in Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[2]Institute of Nursing, Faculty of Health, Oslo and Akershus University College of Applied Sciences, Oslo, Norway

[3]Department of Psychiatry; National Competence Centre for Complex Symptom Disorders, St Olav University Hospital, Trondheim, Norway

[4]Department of Public Health and General Practice, Faculty of Medicine, General Practice Research Unit, Norwegian University of Science and Technology, Trondheim, Norway

[5]Driv Digital, Tønsberg, Norway

[6]NIVEL (Netherlands Institute for Health Services Research), Utrecht, Netherlands

[7]Department of Primary and Community Care, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

[8]Faculty of Health Sciences, Buskerud University College, Drammen, Norway

[9]Jeløy Kurbad Rehabilitation Centre, Moss, Norway

Corresponding Author:
Ólöf Birna Kristjánsdóttir, RN, MSc
Department of Behavioral Sciences in Medicine
Faculty of Medicine
University of Oslo
Postbox 1111
Oslo, 0317
Norway
Phone: 47 22851020
Fax: 47 22851300
Email: kristjansdottir@gmail.com

Abstract

Background: Internet-based interventions are increasingly used to support self-management of individuals with chronic illnesses. Web-based interventions may also be effective in enhancing self-management for individuals with chronic pain, but little is known about long-term effects. Research on Web-based interventions to support self-management following participation in pain management programs is limited.

Objective: The aim is to examine the long-term effects of a 4-week smartphone-intervention with diaries and therapist-written feedback following an inpatient chronic pain rehabilitation program, previously found to be effective at short-term and 5-month follow-ups.

Methods: 140 women with chronic widespread pain, participating in a 4-week inpatient rehabilitation program, were randomized into two groups: with or without a smartphone intervention after the rehabilitation. The smartphone intervention consisted of one face-to-face individual session and 4 weeks of written communication via a smartphone, consisting of three diaries daily to elicit pain-related thoughts, feelings, and activities, as well as daily personalized written feedback based on cognitive behavioral principles from a therapist. Both groups were given access to an informational website to promote constructive self-management. Outcomes were measured with self-reported paper-and-pencil format questionnaires with catastrophizing as the primary outcome measure. Secondary outcomes included daily functioning and symptom levels, acceptance of pain, and emotional distress.

Results: By the 11-month follow-up, the favorable between-group differences previously reported post-intervention and at 5-month follow-up on catastrophizing, acceptance, functioning, and symptom level were no longer evident (P>.10). However, there was more improvement in catastrophizing scores during the follow-up period in the intervention group (M=-2.36, SD 8.41) compared to the control group (M=.40, SD 7.20), P=.045. Also, per protocol within-group analysis showed a small positive effect (Cohen's d=.33) on catastrophizing in the intervention group (P=.04) and no change in the control group from the smartphone intervention baseline to 11-month follow-up. A positive effect (Cohen's d=.73) on acceptance was found within the intervention group (P<.001) but not in the control group. Small to large negative effects were found within the control group on functioning and symptom levels, emotional distress, and fatigue (P=.05) from the intervention baseline to the 11-month follow-up.

Conclusion: The long-term results of this randomized trial are ambiguous. No significant between-group effect was found on the study variables at 11-month follow-up. However, the within-group analyses, comparing the baseline for the smartphone intervention to the 11-month data, indicated changes in the desired direction in catastrophizing and acceptance in the intervention group but not within the control group. This study provides modest evidence supporting the long-term effect of the intervention.

Trial Registration: Clinicaltrials.gov NCT01236209; http://www.clinicaltrials.gov/ct2/show/NCT01236209 (Archived by WebCite at http://www.webcitation.org/6FF7KUXo0)

(J Med Internet Res 2013;15(3):e72) doi:10.2196/jmir.2442

KEYWORDS

Internet-based personalized feedback; widespread chronic pain; fibromyalgia; pain management; eHealth; smartphone; Internet; cognitive therapy; catastrophization

Introduction

Chronic widespread pain (CWP) is a common cause of suffering in the adult population, with reported prevalence rates between 4% and 10% [1-5]. In addition to pain, fatigue, sleep disturbance, and emotional distress are common [1,5]. A subgroup has more severe symptoms and meets the diagnosis criteria of fibromyalgia [2,5]. Knowledge of the pathogenesis of CWP and fibromyalgia is still evolving; dynamic processes including biological, social, and psychological factors are known to be involved [6]. Multidimensional rehabilitation is the recommended treatment, including interventions based on cognitive behavioral therapy (CBT) where patients learn how thoughts, beliefs, and feelings can influence the pain experience and functioning [6,7]. The short-term effects are well established, but concerns about the long-term effects have been raised [7-11]. It has been indicated that for 30-60% of patients participating in pain management programs, the treatment gain is not maintained long-term (at 1- to 5-year follow-ups) [8,10]. The need for strategies to maintain treatment effects by supporting self-management following treatment has received little attention in the research field [8].

An increasing number of studies on Internet-based interventions, many based on CBT (iCBT), indicate their efficacy in supporting use of constructive self-management strategies in individuals living with chronic illness [12-15]. The results of research on iCBT for individuals with chronic pain are not entirely consistent, but in a recent systematic review, it was concluded that Internet-based interventions seem promising [16]. At least three recent randomized trials, not included in this review, support this conclusion. A randomized trial of an intervention consisting of a website with self-management information based on CBT and no therapist contact was found to reduce pain and increase physical functioning in individuals with fibromyalgia, compared with a control group receiving standard care alone, at a 6-month follow-up [17]. Another study, testing the effect of a no-therapist contact online intervention found positive effects on pain, catastrophizing, and disability for the intervention group [18]. In the third study, persons with persistent symptoms after multidisciplinary pain management rehabilitation received 8 weeks of guided iCBT. There was a medium between-group effect between the intervention group and the active control group, but a small within-group effect on catastrophizing after the intervention, and the improvements were maintained after 6 months [19].

To date, research on the effects of iCBT for persons with pain beyond 6-month follow-up is limited. Additionally, only a few studies on iCBT have investigated the effect of an intervention aiming to support self-management following participation in a traditional pain management program [19,20].

Most iCBT interventions for chronic pain are based on weekly modules with self-help material and involve weekly written communication with a therapist [16,19]. A few studies have investigated a different approach to iCBT with daily communication with a therapist over a few weeks, using a personal digital assistant (PDA) or smartphone [20-24]. The use of a smartphone instead of a desktop or laptop computer gives the participant the flexibility to register and receive information in different situations of daily life. In these studies, diaries with questions aiming to support awareness of disability-related thoughts (eg, catastrophizing) and feelings have been made available to patients on a Web-enabled mobile phone or a smartphone. Instead of weekly feedback from a therapist, the participants receive a daily written message personalized according to the recently registered information. Two randomized trials provide evidence for positive short-term effects (3-month/5-month follow-up) [20,24].

In our randomized controlled study, 135 women with CWP completing a 4-week inpatient rehabilitation program were included [20]. A large effect on catastrophizing was found between the groups for the completers after receiving personalized feedback via a smartphone for 4-weeks. At 5-month follow-up, the effects remained moderate for catastrophizing,

acceptance of pain, and functioning and symptom level [20]. The objective of the present paper is to report long-term results of the previously published trial on the smartphone intervention, ie, involving the same study with the same sample [20]. It was hypothesized that the intervention group would report less catastrophizing, better functioning, increased acceptance of pain, and success in values-based living than the control group at 11-month follow-up.

Methods

Study Design

The overall study design is shown in Figure 1. The design is a parallel-group randomized controlled trial. Further details of the study can be found in our earlier publication from this trial [20].

All participants attended a 4-week inpatient multidimensional rehabilitation program for chronic pain. The program included education in pain mechanisms and CBT-based pain management (approximately 20 hours), various forms of aerobic exercise, stretching, relaxation, individual myofascial pain treatment, and medication was administered as needed (see [25] for details of the program). In the fourth week of the program, participants were randomly assigned to one of the two study groups. A detailed description of the recruitment procedure is given in the previous report [20].

The intervention group received a smartphone intervention for 4 weeks after completing the inpatient rehabilitation. Both groups were given access to an informational website with self-help pain-management material. Self-reported assessments on paper were gathered at five time-points: before (T1) and after (T2) the inpatient program, 4 weeks after discharge (T3) when the intervention group had completed their smartphone intervention, and 5 (T4) and 11 months (T5) after the smartphone intervention period (ie, 12 months after discharge from the inpatient rehabilitation program). The first two questionnaires were received and completed at the rehabilitation center and the others were completed at home and returned by mail. In this paper, results of the first two assessments (T1 and T2) and the last (11-month follow-up, T5) are reported. The customary self-report administration mode at the rehabilitation center was a paper-and-pencil format and was therefore used in this study.

Participants

Participants were recruited consecutively from Jeløy Kurbad Rehabilitation Centre in Moss, Norway. Patients were referred to the center by their general practitioner, a medical specialist, or from a hospital. The inclusion criteria for the study were: female, 18 years or older, participating in the inpatient program for persons with chronic pain, having suffered from CWP for more than 6 months (with or without a diagnosis of fibromyalgia), not participating in another research project at the rehabilitation centre, being able to use the smartphone, and not being diagnosed with a profound psychiatric disorder.

Ethical Aspects

The study was approved by the Regional Ethics Committee in South-East Norway and by the Norwegian Social Science Services. All participants signed an informed consent form. The study is registered at ClinicalTrial.gov (NCT01236209).

Assessment Measures

The Pain Catastrophizing Scale (PCS [26]) was used as the primary outcome variable of the study. It is a 13-item questionnaire with questions on helplessness, magnification, and rumination. Patients rate items on a scale from 0 (not at all) to 4 (all the time). The total score range for the PCS is 0-52, with higher scores reflecting higher degrees of catastrophizing. In our sample, the internal consistency was high on all assessments (Cronbach alpha = .89-.94). Catastrophizing is among the psychological constructs that can play an important role in the development and maintenance of chronic pain [27,28]. Catastrophizing has consistently been found to be associated with distress and disability [28]. The Chronic Pain Acceptance Questionnaire (CPAQ) was used. It is scored on a 7-point Likert scale from 0 (never true) to 6 (always true) to give the total score (0-120). Higher scores reflect higher acceptance of pain and higher activities engagement. In our study, the Cronbach alpha coefficients were .81-.92. Emotional distress was measured with the questions from the 12-item General Health Questionnaire (GHQ) [29] with modified response alternatives. Bimodal scoring method was used (symptom present more than usual = 1, symptom present less than or as usual = 0). Total score range is 0 to 12; indicating the number of symptoms present more than usual during the last 2 weeks. In the current study, the Cronbach alpha coefficients were .72-.88. The Chronic Pain Values Inventory (CPVI) is a 12-item measure of importance and success in living according to one's own values in six domains (family, intimate relationships, friendship, work, health, and personal growth) [30]. Each item is rated on a scale from 0 to 5, with higher numbers indicating more importance or success. The mean success rating was used as a measure of values-based action (score range: 0-5). In the present study, the Cronbach alpha coefficients for the success scale were .75-.88. The original (1991) version of fibromyalgia impact questionnaire (FIQ) was used to measure the impact of fibromyalgia on functioning and symptom levels the last week. The score range is 0 to 100; a higher score indicates greater impairment [31]. The Cronbach alpha coefficients were .78-.87 (two questions related to work were excluded because of high missing rates). Short-form health survey (SF-8) was also used to measure functioning. Summary measure scales for mental health component and physical component were obtained by using SF-8 Scoring Software 4.5 [32]. The standardized scores have a mean of 50 and a SD of 10. Higher scores indicate better functioning. The Cronbach alpha for the mental component were .65-.74 and .79-.85 for the physical component.

The current levels (last couple of days) of pain, fatigue, sleep disturbance, and depression were assessed on visual analogue scales (VAS) from 0 (no pain/fatigue/sleep disturbance/depression) to 100 (worst imaginable). One question on subjective global improvement was included: "How do you

feel now compared to before you attended the inpatient program?"

Treatment Procedures

Smartphone Intervention: Diaries and Daily Situational Feedback (Intervention Group Only)

The main theoretical framework was based on the cognitive behavioral fear-avoidance model [27] and CBT, and comprised, more specifically, elements from the acceptance and commitment therapy (ACT) [33,34]. ACT has been found to reduce catastrophizing and disability in chronic pain patients [35-37]. The aim was to support continued use of the self-management strategies learned at the rehabilitation center (eg, exercise and stretching) and to promote improved daily functioning and values-based living. In the rehabilitation center, a traditional CBT approach was used, not ACT. Therefore, ACT elements such as mindfulness exercises were added as new components. The smartphone intervention had the following four components.

Face-to-Face Session

The intervention started with an approximately 1-hour individual session between a nurse and the participant. The session took place in the last week before discharge. The participants received information (name and qualifications) about their therapist for the intervention, which, in some cases was the nurse at the meeting. The nurse attending the face-to-face session summarized the meeting and passed this background information to the relevant therapist. For the duration of the study, the participant was lent a smartphone and could call a member of the research group (OBK, HE) for technical support.

Web-Based Diaries

The participant was asked to complete three diary entries per day using the smartphone. Examples of the smartphone's screen display are shown in Figures 2 and 3. The diaries included 16-24 questions about the current level and interference of pain, feelings, and thoughts related to avoidance, catastrophizing, and acceptance. They also included questions about planned and previous use of self-management activities learned at the rehabilitation center and daily values-based and practical activities. Lists of self-management activities (eg, mild exercise, stretching, resting, aerobic exercise, pleasurable activity) were provided as a reminder. The questions were chosen to support awareness and reflection of experience relevant to self-management. Participants answered most questions by choosing predefined alternatives or using scales. The diaries included a comment field giving participants the opportunity to write a short personal message to the therapist.

At the time scheduled for diary completion, a short message service (SMS) message with a link to a secure website, where the diary could be opened and questions answered and posted, was received by the participant. The participants completed the first diary entry during the face-to-face session and continued during the last week before discharge with the goal of getting used to the diaries before discharge (a run-in period). After discharge, the diaries were completed every day for 4 weeks.

Written Personalized Feedback

For 4 weeks after discharge, excluding weekends, participants received daily written feedback from a therapist. The feedback was empathic and personalized according to each participant's situation as reported in the diary. It included repetition of content reported in the diaries, positive reinforcement, reminders of self-management information given at the rehabilitation center, ACT exercises, and reflective questions. The aim was to encourage nonjudgmental awareness of cognitions, feelings, and emotions and to stimulate mindfulness and willingness to engage in meaningful activities despite pain or other discouraging intrusions, eg, to reduce the impact of catastrophizing on self-management behavior. The instructions for the exercises were written directly in the feedback or the participant was referred to exercises available on the smartphone and/or the website (see below). The feedback was also personalized according to the summary of personal information given at the face-to-face session (eg, family situation and health-related goals) and results on self-reported discrepancy between values and values-based living assessed with the CPVI at the end of the rehabilitation program. The feedback was usually available for the participant within 90 minutes after they had completed the second diary of the day. If this diary was not submitted, feedback based on information from the latest submitted diary was sent. There was no limitation on the length of the feedback, which ranged from a few sentences to a few paragraphs.

The feedback was written by any of 3 of the authors (OBK, TLS, and HE); all with a background in health care sciences (nursing and/or psychology).

Audio Files With Guided Mindfulness Exercises

A few audio files with short mindfulness exercises guided by the authors were available on the smartphones.

Informational Website With Self-Help Pain Management Material (All Participants; Control Group Received Only This Intervention)

All participants received access to a static website with information on self-management strategies for people with chronic pain. The website also included a few written ACT exercises and audio files with mindfulness exercises (as described above). See Multimedia Appendix 1 for a screenshot from the website. No specific instruction about frequency of use was given.

Statistics

To investigate differences in demographic variables and baseline characteristics, independent sample *t*-tests, nonparametric tests, and Chi-square tests were used. Paired *t*-tests were used to compare 11-month follow-up (T5) results to the baseline for the inpatient program (T1) and the smartphone intervention (T2). Independent *t*-tests or nonparametric tests (Mann-Whitney) were used to compare outcomes between groups at T5. The Cohen's *d* effect sizes (ES) were calculated using the difference between the groups' means divided by the mean of the standard deviation of both groups. If one or two items were missing on the GHQ, they were scored as present less than usual or as usual (= 0). If another instrument included one or two missing items,

the items were replaced with the mean of other items from the participant's instrument. If two response alternatives were marked, the healthier option was chosen. Total score was not computed if more than two items were missing, and the case was categorized as missing a total score for the instrument. The number of participants included in each analysis is provided. In the intention-to-treat analysis (ITT), the results of complete case analysis for the primary outcome is reported. In addition, two methods for replacing missing variables for the primary outcome at endpoint (T5) were applied: last observation carried forward (LOCF) and multiple imputations (MI). In the MI analysis, 50 imputations were made. The following clinically significant variables were included in the MI regression model: age, SF-8 physical component, and VAS for pain, sleep, fatigue, and depression at admission to the rehabilitation center. Six of the participants who withdrew from the smartphone intervention contributed questionnaires at the 11-month follow-up (T5). The ITT analyses included all participants (n=135) except those who met the exclusion criteria after randomization. In the analyses of secondary outcomes, only those who completed the interventions were included (n=112). IBM SPSS Statistics (versions 19 and 20) was used. A significance level of P=.05 was chosen, and a tendency toward difference was defined as P<.1. Effect sizes were categorized as small (<.5), medium (.5-.8), and large (>.8) in accordance with Cohen [38].

Figure 1. Study flow chart.

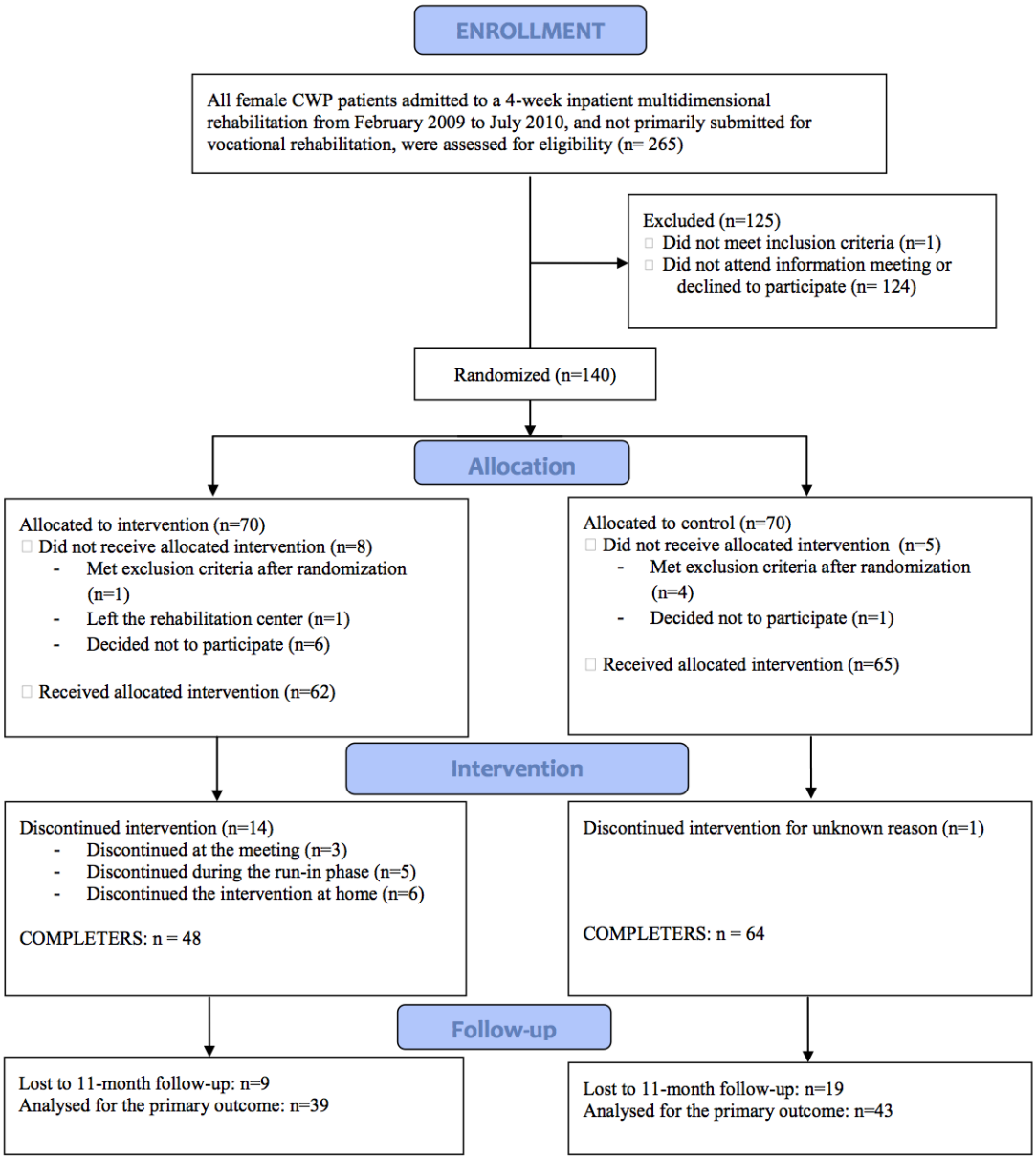


Figure 2. The smartphone's screen showing a diary in Norwegian.

Figure 3. The smartphone's screen showing feedback in Norwegian.

Results

Participants

265 women who were eligible for the study during the study period were invited to informational meetings about the project. Of these, 124 did not attend a meeting or declined to participate, and 1 did not meet the inclusion criteria. 140 were randomized to one of the two study arms (Figure 1). 5 subjects met the exclusion criteria after randomization (they were originally submitted for vocational rehabilitation and thus included in another research project), and 8 discontinued participation before receiving the allocated intervention. In the intervention group, 14 patients did not complete the intervention. Demographic data and baseline characteristics of the sample by groups are given in Table 1. All participants had CWP, and the majority was diagnosed with fibromyalgia. Despite randomization, the groups differed in mean pain level (P=.02) and physical functioning measured by SF-8 (P=.03) at admission to the rehabilitation center.

Primary Outcome: Catastrophizing

Descriptive results for catastrophizing are shown in Table 2 and follow-up differences in Table 3.

Between-Group Effects

At the 11-month follow-up (T5), there was no difference between the groups on the measure of catastrophizing (PCS); neither according to the ITT-analysis (LOCF P=.22 and MI P=.31) nor the per protocol analysis (complete case analysis, P=.18 and analysis with MI .23). When all intended-to-treat were analyzed (LOCF), catastrophizing seemed to improve more during the follow-up period (T2-T5) in the intervention group (M=-2.36, SD 8.41) than the control group (M=.40, SD 7.20), P=.045, using t-test of change scores.

Within-Group Effects

There were small positive within-group effects for the intervention group between T2 and T5 on catastrophizing by ITT (LOCF) and per protocol analyses, P=.02 and P=.04, respectively. In the analysis, where missing variables were imputed, there was a tendency towards a small positive within-group effect during this period (T2-T5). There were improvements regarding to catastrophizing in the T1 to T5 period for both the intervention and control groups, P<.001. Between 5-month follow-up (T4) and 11-month follow-up (T5), paired samples t-tests did not show changes in the intervention group (M=.06, SD 5.05, n=31), P=.94 for catastrophizing. However, in this period (T4-T5), there was a reduction in catastrophizing in the control group (M=-3.25, SD 7.09, n=34), P=.01.

Secondary Outcomes: Functioning and Symptom Levels

Table 4 shows descriptive statistics for the secondary outcomes at admission to the rehabilitation center, at discharge, and at the 11-month follow-up. In the per protocol analysis, no significant group differences were detected at discharge from the rehabilitation center on any of the outcome variables (all Ps>.05; GHQ, and depression (VAS), P=.08, see Table 4).

Between-Group Effects

No between-group differences were found at the 11-month follow-up; P-values ranged from .13 (SF-8, physical) and .17 (CPAQ) to .98 (sleep disturbance).

Within-Group Effects

Table 5 shows within-group changes for the secondary outcomes. When comparing the smartphone intervention's baseline data (T2) to the 11-month follow-up (T5), there was a moderate positive effect on acceptance (CPAQ) in the intervention group, but not in the control group. There was a small negative effect on functioning and symptom levels measured by the FIQ in the control group, but not in the intervention group. For the physical component of SF-8, there was a small negative effect for the intervention group (P=.046), but not the control group. For GHQ, there was a large negative effect for emotional distress but only in the control group. No significant changes were detected for the mental component of SF-8. There was a tendency towards improvement in values-based living in the intervention group but not in the control group. There was a moderate negative effect on fatigue (VAS), and a tendency towards a small negative effect on sleep (VAS) and pain (VAS) in the control group. No changes were detected on these symptoms in the intervention group.

When comparing baseline data for the inpatient program (T1) to the follow-up data (T5), improvement in acceptance, mental health measured by SF8, and values-based living was found in both groups (see Table 5). Reduction in disease impact (measured by FIQ) was found for the intervention group only (Cohen's d=.42, P=.03). There was a significant reduction in pain level in the intervention group between T1 and T5, mostly due to changes during the inpatient program (see Table 4). There was a trend towards improvement on fatigue and depression (VAS scales) in the intervention group only, between T1 and T5. When both groups were analyzed together, with all intended-to-treat included (complete case analysis), there was a reduction in pain level (M=5.68, SD 24.66, n=89) between admission to the inpatient program (T1) and the long-term follow-up (T5), P=.03.

When the 5-month follow-up results (T4) were compared to the 11-month follow-up (T5), no changes were found for acceptance, pain level, functioning, and symptom level (measured by FIQ), sleep disturbance, fatigue, and mental health (all P values > 0.10). During this period (T4-T5), the control group showed improvement in values-based living (M=.25, SD .70, n=33, P=.046), whereas the intervention group did not. The control group also showed improvement in depression (measured by a VAS, M=8.29, SD 20.13, n=35, P=.02), whereas no significant change was found in the intervention group during this (T4-T5) period. Reduction in physical functioning (measured by SF8) was found in the intervention group (M=3.45, SD 7.76, n=31, P=.02) and a trend towards improvement in the control group (M=2.30, SD 7.10, n=34, P=.07).

Of the completers, 47.4% (n=18) in the intervention group and 40.0% (n=18) in the control group reported feeling better now than before the inpatient program. 13.1% (n=5) in the

intervention group and 11.1% (n=5) in the control group reported feeling worse now compared to before the inpatient program. No change was reported by 39.5% (n=15) in the intervention group and by 48.9% (n=22) in the control group.

Withdrawal From Participation

Of the 135 participants (of 140 randomized) that met the inclusion criteria, 112 completed the study period (Figure 1). 21 withdrew from the intervention group (30.4%) and 2 withdrew from the control group (3.0%). There was a trend toward the completers being younger (M=43.33, SD 11.18) than the ones who withdrew (M=48.43, SD 10.06), *P*=.07. Additionally, there was a trend towards a higher level of depression (measured by a single VAS) in the group who withdrew (M=43.62, SD 28.57) compared to the completers (M=31.81, 28.92), at admission to the inpatient program, *P*=.06.

Response Rates to 11-month Questionnaires in Intervention Group and Control Group

The response rate for all included participants (n=135) was 66.7% at 11-month follow-up (T5) (n=45 in the intervention group and n=45 in the control group). When only completers (n=112) were included, the response rate at T5 was 81.3% (n=39) in the intervention group and 70.3% (n=45) in the control group. Among the completers, those who returned the questionnaire at T5 had better physical functioning (M=34.60, SD 7.53), at admission to the inpatient program measured by SF8, compared to those who did not return them (M=31.19, SD 6.93), *P*=.01. Those who returned the questionnaires at T5 reported also less disease impact (M=56.45, SD 16.82) on FIQ at admission to the inpatient program compared to those who did not return them (M=64.22, SD 14.59), *P*=.03. Those who returned the T5 assessments reported more success in living according to values (M=2.13, SD .81) than those who did not return them (M=1.74, SD .85), *P*=.04. There was also a tendency towards more severe depression (measured by a VAS scale) at both admission and discharge among those who did not return questionnaires at T5 compared to those who returned them, *P*=.08-.09.

Table 1. Characteristics at admission to the inpatient rehabilitation center.

Characteristic		Smartphone intervention (n=69)[a]	Control (n=66)[a]
Age, mean (SD), n		44.59 (11.13), 69	43.80 (11.20), 65
Marital status			
	Married or cohabiting	60.9% (n=42)	68.2% (n=45)
	Divorced	13.0% (n=9)	9.1% (n=6)
	Single	18.8% (n=13)	15.2% (n=10)
	Widow	5.8% (n=4)	3.0% (n=2)
	Unknown	1.4% (n=1)	4.5% (n=3)
Years of education			
	≤ 10 years (elementary)	18.8% (n=13)	12.1% (n=8)
	11-13 years (high school)	27.5% (n=19)	45.5% (n=30)
	>13 years (College/University)	43.5% (n=30)	34.8% (n=23)
	Unknown	10.1% (n=7)	7.6% (n=5)
Employment status			
	Working/studying	21.7% (n=15)	12.1% (n=8)
	Unemployed	4.3% (n=3)	1.5% (n=1)
	On sick leave	39.1% (n=27)	51.5% (n=34)
	On disability pension	17.4% (n=12)	19.7% (n=13)
	Part time working/studying and part time sick leave	11.6% (n=8)	7.6% (n=5)
	Other combination of the above	5.8% (n=4)	6.1% (n=4)
	Unknown	0%	1.5% (n=1)
Diagnosed with fibromyalgia (valid %)		80.9% (n=55)	84.4% (n=54)
Duration of symptoms (years), mean (SD), n		13.11 (8.78)	15.47 (12.09)
PCS[b], mean (SD), n		21.24 (10.33), 63	20.80 (9.45), 62
CPAQ[b], mean (SD), n		56.48 (15.02), 58	53.87 (13.81), 57
FIQ[b], mean (SD), n		58.75 (16.39), 69	58.58 (16.04), 66
SF-8, physical[b]; mean (SD), n		31.91 (7.57), 65	34.75 (7.35), 62
SF-8, mental[b], mean (SD), n		39.33 (10.49), 65	39.34 (9.61), 62
GHQ-12[b], mean (SD), n		3.32 (3.38), 62	3.02 (3.38), 61
CPVI[b], mean (SD), n		2.07 (0.95), 64	2.01 (0.73), 61
VAS [b] recordings of current level of (last couple of days):			
	Pain, mean (SD), n	67.08 (17.47), 69	57.85 (21.60), 66
	Fatigue, mean (SD), n	67.40 (23.73), 69	64.72 (21.02), 66
	Sleep disturbance, mean (SD), n	57.24 (26.22), 68	55.16 (23.38), 66
	Depression, mean (SD), n	34.73 (29.15), 68	32.93 (29.26), 65

[a] Patients meeting exclusion criteria after randomization were not included in this analysis.

[b] VAS, visual analogue scale (0-100[c]); PCS, Pain Catastrophizing Scale (score range 0-52[c]); CPAQ, Chronic Pain Acceptance Questionnaire (score range 0[c]-120); FIQ, Fibromyalgia Impact Questionnaire (0-100[c]); SF-8 (0[c]-100), Short Form; GHQ-12, questions from the General Health Questionnaire (score range 0-12[c]); CPVI, Chronic Pain Values Inventory (success score, range 0[c]-6).

[c] Values that indicate maximum symptom scores/least health.

Table 2. Means and standard deviations for the primary outcome at admission to the inpatient rehabilitation (T1), at discharge (T2), and 11 months (T5) after the intervention period.

Primary outcome measure, Pain Catastrophizing Scale	Group	T1[a] Mean (SD), n	T2[a, b] Mean (SD), n	T5[a] Mean (SD), n
ITT (complete case analysis)				
	Intervention	21.24 (10.33), 63	15.12 (9.61), 63	11.50 (8.68), 44
	Control	20.80 (9.45), 62	15.41 (9.62), 59	14.73 (9.95), 43
ITT(LOCF)				
	Intervention	21.24 (10.33), 63	16.06 (10.37), 68	13.72 (10.02), 69
	Control	20.80 (9.45), 62	15.33 (9.31), 65	15.57 (10.40), 66
ITT (MI)				
	Intervention			12.80, 69
	Control			14.74, 66
Per protocol (complete case analysis)				
	Intervention	20.56 (10.08), 43	14.61 (8.93), 45	11.92 (8.97), 39
	Control	20.78 (9.59), 60	15.46 (9.76), 57	14.73 (9.95), 43
Per protocol (MI)				
	Intervention			12.25, 48
	Control			14.66, 64
% (valid) with PCS score >24				
	Intervention	30.2%, 13	15.6%, 7	12.8%, 5
	Control	33.3%, 20	17.5%, 10	18.6%, 8

[a] T1, at admission to the inpatient program; T2, at discharge from the inpatient program; T5, 11-month follow-up

[b] No differences between groups at T2 (all *P* values >.05).

Table 3. Within-group differences and effect sizes (ES) for the primary outcome.

Primary outcome, PCS		Mean difference T2-T5[a] (n)	95% CI T2-T5	ES for T2-T5	*P*-value[b]	Mean difference T1-T5[a] (n)	95% CI T1-T5	ES for T1-T5	*P*-value[b]
ITT (complete case analysis)									
	Intervention	-2.60 (8.72), 42	-5.32, .11	.29	.06	-7.36 (7.87), 39	-9.91, -4.81	.80	< .001
	Control	-.21 (7.32), 40	-2.55, 2.13	.02	.86	-5.30 (7.30), 41	-7.60, -2.99	.56	< .001
ITT (LOCF)									
	Intervention	-2.36 (8.41), 68	-4.39, -.32	.23	.02	-7.57 (8.23), 63	-9.64, -5.50	.76	< .001
	Control	.40 (7.20), 65	-1.39, 2.18	-.04	.66	-4.65 (7.43), 62	-6.54, -2.76	.47	< .001
ITT (MI)									
	Intervention	-2.61, 63	-5.33, .10		.06	-8.22, 63	-10.86, -5.57		< .001
	Control	-.15, 59	-2.75, 2.46		.91	-5.79, 62	-8.12, -3.46		< .001
Per protocol (complete case analysis)									
	Intervention	-3.04 (8.74), 37	-5.95, -.12	0.33	.04	-7.23 (8.01), 34	-10.03, -4.44	.77	< .001
	Control	-.21 (7.32), 40	-2.55, 2.13	0.02	.86	-5.30 (7.30), 41	-7.60, -2.99	.56	< .001
Per protocol (MI)									
	Intervention	-2.45, 45	-5.29, .38		.09	-7.93, 43	-10.72, -5.14		< .001
	Control	-.27, 57	-2.84, 2.30		.84	-5.85,60	-8.15, -3.54		< .001

[a] T1, at admission to the inpatient program; T2, at discharge from the inpatient program; T5, 11-month follow-up.

[b] *P* values for paired samples *t*-tests.

Table 4. Means and standard deviations for the secondary outcomes at admission to the inpatient rehabilitation (T1), at discharge (T2), and 11 months (T5) after the intervention period, for the completers.

Secondary outcome measures	Group	T1[a] Mean (SD), n	T2[a,b] Mean (SD), n	T5[a] Mean (SD), n
CPAQ				
	Intervention	56.45 (15.22), 40	62.00 (13.62), 44	71.62 (14.11), 39
	Control	53.94 (13.92), 56	62.21 (10.15), 57	67.05 (15.18), 42
FIQ				
	Intervention	58.46 (17.26), 48	46.38 (16.92), 47	49.24 (21.34), 38
	Control	58.35 (16.18), 64	49.10 (17.32), 62	53.75 (17.73), 45
SF-8, physical				
	Intervention	32.12 (7.74), 45	36.68 (8.42), 40	34.38 (9.88), 39
	Control	34.98 (7.13), 60	35.86 (8.24), 49	37.35 (7.65), 44
SF-8, mental				
	Intervention	39.50 (10.67), 45	45.70 (8.06), 40	45.50 (10.70), 39
	Control	39.09 (9.61), 60	44.83 (9.69), 49	43.87 (9.09), 43
GHQ-12				
	Intervention	3.19 (3.21), 43	1.20 (2.02), 45	1.95 (2.64), 38
	Control	2.97 (3.43), 59	0.63 (1.01), 57	2.20 (2.82), 44
CPVI				
	Intervention	2.05 (0.95), 44	2.47 (0.91), 46	2.78 (1.00), 39
	Control	2.02 (0.74), 59	2.52 (0.68), 54	2.50 (0.77), 43
Pain, VAS				
	Intervention	66.59 (17.58), 48	53.07 (22.20), 47	56.28 (28.24), 38
	Control	57.32 (21.56), 64	52.99 (21.27), 61	55.85 (22.73), 45
Fatigue, VAS				
	Intervention	69.29 (23.98), 48	51.38 (27.75), 47	60.79 (28.56), 38
	Control	64.08 (21.01), 64	50.10 (24.28), 61	61.63 (23.63), 45
Sleep disturbance, VAS				
	Intervention	54.77 (26.99), 47	43.97 (25.77), 47	51.68 (30.45), 38
	Control	54.59 (23.31), 64	48.12 (24.57), 62	53.08 (25.95), 45
Depression, VAS				
	Intervention	30.68 (28.71), 47	19.84 (24.08), 47	22.82 (25.89), 38
	Control	32.65 (29.29), 63	27.36 (28.51), 61	28.93 (27.71),45

[a] T1, at admission to the inpatient program; T2, at discharge from the inpatient program; T5, 11-month follow-up.

[b] No differences between groups at T2 (all P values >.05; GHQ and depression (VAS), *P*=.08).

Table 5. Mean differences for the secondary outcomes within-groups, confidence intervals (CI) and effect sizes (ES) for the completers.

Secondary outcome measures		Mean difference T2-T5 (n)[a]	95% CI T2-T5[a]	ES for T2-T5[a]	*P*-value[b]	Mean difference T1-T5 (n)[a]	95% CI T1-T5[a]	ES[a] for T1-T5	*P*-value[b]
CPAQ									
	Intervention	10.20 (9.57), 35	6.91, 13.49	.73	<.001	14.12 (12.46), 33	9.70, 18.54	.95	< .001
	Control	2.15 (10.77), 39	-1.34, 5.64	.17	.22	8.45 (14.00), 38	3.85, 13.05	.62	.001
FIQ									
	Intervention	1.40 (21.15), 37	-5.65, 8.45	-.08	.69	-8.15 (21.95), 38	-15.36, -.93	.42	.03
	Control	7.05 (17.24), 44	1.81, 12.29	-.40	.01	-2.27 (16.23), 45	-7.15, 2.60	.13	.35
SF-8, physical									
	Intervention	-3.20 (8.84), 33	-6.33, -.06	-.36	.046	1.11 (8.99), 36	-1.93, 4.15	.12	.46
	Control	.54 (8.18), 34	-2.32, 3.39	.06	.71	1.01 (6.99), 42	-1.17, 3.19	.14	.36
SF-8, mental									
	Intervention	.39 (10.09), 33	-3.19, 3.97	.04	.82	6.45 (11.71), 36	2.49, 10.42	.58	.002
	Control	-1.61 (11.53), 33	-5.70, 2.48	-.17	.43	3.10 (6.99), 42	.93, 5.28	.36	.006
GHQ-12									
	Intervention	.69 (3.06), 35	-.36, 1.74	-.28	.19	-1.12 (4.07), 34	-2.54, .30	.36	.12
	Control	1.68 (2.94), 40	.74, 2.61	-.85	.001	-.61, 3.39, 41	-1.68, .46	.20	.26
CPVI									
	Intervention	.29 (.90), 37	-.01, .59	.30	.06	.56 (.72), 35	.32, .81	.59	< .001
	Control	-.11 (.88), 36	-.41, .19	-.16	.46	.40 (.61), 41	.21, .59	.54	< .001
Pain, VAS									
	Intervention	1.43 (32.58), 37	-9.44, 12.29	-.06	.79	-9.86 (25.54), 38	-18.25, -1.46	.43	.02
	Control	5.53 (20.73), 43	-.85, 11.91	-.25	.09	-1.82 (23.65), 45	-8.92, 5.28	.08	.61
Fatigue, VAS									
	Intervention	8.71 (33.56), 37	-2.48, 19.90	-.32	.12	-8.94 (31.93), 38	-19.43, 1.56	.36	.09
	Control	14.96 (24.50), 43	7.42, 22.49	-.61	<.001	-.93 (22.65), 45	-7.73, 5.87	.04	.78
Sleep disturbance, VAS									
	Intervention	5.55 (33.26), 37	-5.54, 16.64	-.20	.32	-2.29 (33.46), 37	-13.45, 8.86	.08	.68
	Control	7.43 (27.09), 44	-.81, 15.66	-.28	.08	.45 (30.11), 45	-8.59, 9.50	-.02	.92
Depression, VAS									
	Intervention	1.11 (28.71), 37	-8.46, 10.68	-.05	.82	-8.13 (27.78), 37	-17.40,1.13	.29	.08
	Control	7.27 (27.43), 43	-1.18, 15.71	-.26	.09	.99 (23.13), 45	-5.96, 7.94	-.04	.78

[a]T1, at admission to the inpatient program; T2, at discharge from the inpatient program; T5, 11-month follow-up.

[b]*P* values for paired samples *t*-tests.

Discussion

The results of the study are ambiguous. On one hand, there were no significant differences in mean values on any variables between the groups at 11-month follow-up. Thus, the favorable effects previously reported on catastrophizing, acceptance, functioning, and symptom levels at 5-month follow-up were not evident at long-term follow-up. However, there was significantly more improvement in catastrophizing scores during the follow-up period in the intervention group compared to the control group. Moreover, the within-group analyses, comparing the baseline for the smartphone intervention to the 11-month data, revealed changes in the desired direction in catastrophizing and acceptance in the intervention group but not within the control group. Also, increase in disease impact, emotional distress, and fatigue were seen in the control group but not within the intervention group. Additionally, effects on most

variables were maintained in the intervention group from the 5-month follow-up to the 11-month follow-up. Unexpectedly, between the two follow-ups, the control group reported some improvement in several variables (catastrophizing, values-based living, and depression) whereas the intervention group did not. We have no data to support an explanation for this improvement. One could speculate that it takes time for changes in thoughts, behavior, and priorities promoted by the multidimensional inpatient rehabilitation program to settle and cause positive effects. The controls did not get the smartphone intervention that could promote these changes early after discharge, and thus the changes may have been achieved at an earlier stage in the intervention group. We do not have exact login information for visits to the website. As mentioned in our previous paper, most participants in the control group (26 of the 38 who reported this information) visited it rarely (2 times or less). The impression of the administrator of the website (HE) was that it was seldom accessed. Based on the limited use of the website, it is not assumed to have caused any changes seen in the control group. The spontaneous improvement in the control group, large variations within variables, relatively few participants, and small effect sizes may explain the lack of significant differences between the groups. It is important to acknowledge that the effects of the inpatient program were sustained at the long-term follow-up in the control group for many of the outcome variables, ie, catastrophizing, acceptance, mental health, and values-based living. Improvement in those variables indicates that the participants cope better with their situation.

Some of the limitations regarding the generalizability of the study have been discussed in the previous report, eg, the difference in the completers groups versus those withdrawing [20]. Again, at this follow-up we have the limitations of a response rate below 70%. Those not returning the follow-up questionnaires reported generally more symptoms at admission to the inpatient program than those who returned them, thus having the possibility to influence the results. Multiple imputations have been recommended to improve the validity of results in trials with incomplete datasets [39]. In the ITT analysis, the level of catastrophizing in the control group at endpoint (T5) was almost the same for the complete case analysis (mean 14.73, n=43) and the MI analysis (mean 14.74, n=66). In the intervention group, the catastrophizing level was somewhat higher with MI (mean 12.80, n=69) compared to the complete case analysis (mean 11.50, n=44). This might partly be explained by higher baseline scores on two variables (pain and SF-8 physical component), which were included in the MI regression model. Importantly, in the per protocol analysis, the difference between the mean levels of catastrophizing with MI or without (complete case analysis) was small. This provides some support for the validity of our results of secondary outcomes, where results of complete case analysis is reported. However, in the within-group analysis, the difference between the intervention baseline (T2) and 11-month follow-up (T5) in the intervention group was significant applying complete case analysis (P=.04) but only borderline significant in the analysis with MI (P=.09), thus indicating that the results for complete case analysis should be interpreted with some caution.

The withdrawal rate of 30% indicates that this type of secondary intervention may not be found feasible by all. The withdrawal rate is similar to those reported in many iCBT, where an average dropout rate of 27% has been reported [40]. The patients who withdrew tended to score higher on depression and were older than the completers, which could have influenced their interest and capacity to participate. We do not have information on the reasons for withdrawal for all participants. However, many of those who withdrew before or during the run-in period reported that the combination of the smartphone intervention and participation in the inpatient program was stressful or expected to be stressful. Therefore, closer collaboration with the rehabilitation center and flexibility in start-up date of the smartphone intervention might contribute to reduction in withdrawal rates. At the 11-month follow-up, the subjective global improvement measure could have been improved by including a question to assess the participants' evaluation of the smartphone intervention.

Medications, education, CBT, and physical exercises are among the recommended treatment options for individuals with CWP and fibromyalgia [7,9]. The short-term effects are well established, but concerns about the long-term effects have been raised. In a recent longitudinal study including 1555 patients with fibromyalgia receiving standard care, with a mean follow-up period of 4 years, no clinically meaningful improvement in overall symptom severity was found for the sample. Only about one fourth of the sample showed meaningful improvement, including 10% with substantial improvement in symptom severity [41]. The goals of most nonpharmacological treatment are to provide knowledge and teach self-management skills aiming to reduce symptoms or support constructive coping. Adherence to recommended self-management strategies after treatment seems important for long-term effect. The research literature on Internet-delivered interventions to support self-management in individuals with chronic pain is rapidly evolving, with studies on therapist-guided and unguided intervention, as well as on applications for smartphones [16,42,43]. The present smartphone intervention was primarily meant to support use of constructive coping skills and implementation of recommended lifestyle changes the first weeks after discharge from inpatient rehabilitation, in order to prevent the fading of positive effects from the given rehabilitation. The smartphone intervention introduced elements from ACT, including mindfulness, which had not been presented in the inpatient program. We do not know if this influenced the results. Nevertheless, we acknowledge that the smartphone intervention could have been more strongly integrated in the rehabilitation program, eg, including the same health care professionals. The feasibility and long-term efficacy of the intervention might possibly be improved by providing the diaries on the individuals' own smartphones and by providing feedback on a more long-term basis. An 8-week, guided iCBT intervention following multidisciplinary treatment was found to reduce catastrophizing in a randomized trial with 72 persons with residual symptoms post treatment [19]. The intervention lasted twice as long as our intervention but included less therapist contact. The effects on catastrophizing were remained at 6-month follow-up, but more long-term effects are not reported [19]. A smartphone application based on ACT to support

values-based living was found feasible in an exploratory study including 11 healthy volunteers [43]. Ways to tailor the diary content and provide tailored feedback should be investigated. Booster periods with therapist-feedback might be beneficial or a longer period with less frequent therapist contact, eg, once a week/month.

To conclude, the results of this randomized trial are ambiguous. No significant between-group effect was found on the study variables at 11-month follow-up. However, more improvement in catastrophizing scores was seen in the intervention group than the control group in the period between discharge from the inpatient program and the follow-up. Moreover, the within-group analyses, comparing the baseline for the smartphone intervention to the 11-month data indicated changes in the desired direction in catastrophizing and acceptance in the intervention group but not within the control group. Also, increases in disease impact, emotional distress, and fatigue were seen in the control group but not within the intervention group. This kind of smartphone intervention may therefore be suited for providing self-management support following inpatient pain management program. Research on strategies to provide feasible self-management support on a long-term basis for individuals with CWP and ways to enhance cost-effectiveness is needed.

Acknowledgments

The study is primarily funded by the Research Council of Norway (Grant No. 182014) (OBK, HE, EE, and TLS). We would like to thank the women who were willing to participate in the trial and the staff at Jeløy Kurbad Rehabilitation Centre. We thank Are Pripp and Lien My Diep (Department for biostatistics and epidemiology, Oslo University Hospital) for statistical advice.

All authors participated in the development of the intervention. OBK and HE recruited participants and collected data, helped by SHW who was responsible for the patients' diagnostics and data collection at the inpatient center. OBK, TLS, and HE performed the role of the therapist. EE was responsible for the design and development of the technological system. HE was project leader. All authors contributed to and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot from the website.

[PNG File, 179KB - jmir_v15i3e72_app1.png]

Multimedia Appendix 2

CONSORT-EHEALTH Checklist V1.6.2 [44].

[PDF File (Adobe PDF File), 1MB - jmir_v15i3e72_app2.pdf]

References

1. Croft P, Rigby AS, Boswell R, Schollum J, Silman A. The prevalence of chronic widespread pain in the general population. J Rheumatol 1993 Apr;20(4):710-713. [Medline: 8496870]
2. Clauw DJ, Crofford LJ. Chronic widespread pain and fibromyalgia: what we know, and what we need to know. Best Pract Res Clin Rheumatol 2003 Aug;17(4):685-701. [Medline: 12849719]
3. Shipley M. Chronic widespread pain and fibromyalgia syndrome. Medicine 2010;38(4):202-204.
4. Lindell L, Bergman S, Petersson IF, Jacobsson LT, Herrström P. Prevalence of fibromyalgia and chronic widespread pain. Scand J Prim Health Care 2000 Sep;18(3):149-153. [Medline: 11097099]
5. Cöster L, Kendall S, Gerdle B, Henriksson C, Henriksson KG, Bengtsson A. Chronic widespread musculoskeletal pain - a comparison of those who meet criteria for fibromyalgia and those who do not. Eur J Pain 2008 Jul;12(5):600-610. [doi: 10.1016/j.ejpain.2007.10.001] [Medline: 18024204]
6. Flor H, Turk DC. Chronic Pain: An Integrated Biobehavioral Approach. Chronic pain: IASP Press; 2011.
7. Goldenberg DL, Burckhardt C, Crofford L. Management of fibromyalgia syndrome. JAMA 2004 Nov 17;292(19):2388-2395. [doi: 10.1001/jama.292.19.2388] [Medline: 15547167]
8. Morley S. Relapse prevention: still neglected after all these years. Pain 2008 Feb;134(3):239-240. [doi: 10.1016/j.pain.2007.12.004] [Medline: 18164134]
9. Häuser W, Bernardy K, Arnold B, Offenbächer M, Schiltenwolf M. Efficacy of multicomponent treatment in fibromyalgia syndrome: a meta-analysis of randomized controlled clinical trials. Arthritis Rheum 2009 Feb 15;61(2):216-224 [FREE Full text] [doi: 10.1002/art.24276] [Medline: 19177530]
10. Turk DC, Rudy TE. Neglected topics in the treatment of chronic pain patients--relapse, noncompliance, and adherence enhancement. Pain 1991 Jan;44(1):5-28. [Medline: 2038489]

11. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary rehabilitation for fibromyalgia and musculoskeletal pain in working age adults. Cochrane Database Syst Rev 2000(2):CD001984. [doi: 10.1002/14651858.CD001984] [Medline: 10796458]
12. Barak A, Grohol JM. Current and Future Trends in Internet-Supported Mental Health Interventions. Journal of Technology in Human Services 2011;29:155-196. [doi: 10.1080/15228835.2011.616939]
13. Barak A, Hen L, Boniel-Nissim M, Shapira N. A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. Journal of Technology in Human Services 2008 Jul 2008;26(2-4):109-160. [doi: 10.1080/15228830802094429]
14. McGeary DD, McGeary CA, Gatchel RJ, Allison S, Hersh A. Assessment of Research Quality of Telehealth Trials in Pain Management: A Meta-Analysis. Pain Pract 2012 Sep 27:1-10. [doi: 10.1111/j.1533-2500.2012.00601.x] [Medline: 23017210]
15. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
16. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]
17. Williams DA, Kuper D, Segar M, Mohan N, Sheth M, Clauw DJ. Internet-enhanced management of fibromyalgia: a randomized controlled trial. Pain 2010 Dec;151(3):694-702 [FREE Full text] [doi: 10.1016/j.pain.2010.08.034] [Medline: 20855168]
18. Ruehlman LS, Karoly P, Enders C. A randomized controlled evaluation of an online chronic pain self management program. Pain 2012 Feb;153(2):319-330 [FREE Full text] [doi: 10.1016/j.pain.2011.10.025] [Medline: 22133450]
19. Buhrman M, Fredriksson A, Edström G, Shafiei D, Tärnqvist C, Ljótsson B, et al. Guided Internet-delivered cognitive behavioural therapy for chronic pain patients who have residual symptoms after rehabilitation treatment: Randomized controlled trial. Eur J Pain 2012 Nov 9. [doi: 10.1002/j.1532-2149.2012.00244.x] [Medline: 23139021]
20. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. J Med Internet Res 2013;15(1):e5 [FREE Full text] [doi: 10.2196/jmir.2249] [Medline: 23291270]
21. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, van Dulmen S, Wigers SH, et al. Written online situational feedback via mobile phone to support self-management of chronic widespread pain: a usability study of a Web-based intervention. BMC Musculoskelet Disord 2011;12:51 [FREE Full text] [doi: 10.1186/1471-2474-12-51] [Medline: 21352516]
22. Sorbi MJ, Mak SB, Houtveen JH, Kleiboer AM, van Doornen LJ. Mobile Web-based monitoring and coaching: feasibility in chronic migraine. J Med Internet Res 2007;9(5):e38 [FREE Full text] [doi: 10.2196/jmir.9.5.e38] [Medline: 18166526]
23. Nes AA, van Dulmen S, Eide E, Finset A, Kristjánsdóttir OB, Steen IS, et al. The development and feasibility of a web-based intervention with diaries and situational feedback via smartphone to support self-management in patients with diabetes type 2. Diabetes Res Clin Pract 2012 Sep;97(3):385-393. [doi: 10.1016/j.diabres.2012.04.019] [Medline: 22578890]
24. Oerlemans S, van Cranenburgh O, Herremans PJ, Spreeuwenberg P, van Dulmen S. Intervening on cognitions and behavior in irritable bowel syndrome: A feasibility trial using PDAs. J Psychosom Res 2011 Mar;70(3):267-277. [doi: 10.1016/j.jpsychores.2010.09.018] [Medline: 21334498]
25. Wigers SH, Finset A. [Rehabilitation of chronic myofascial pain disorders]. Tidsskr Nor Laegeforen 2007 Mar 1;127(5):604-608 [FREE Full text] [Medline: 17332816]
26. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. Psychological Assessment 1995 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
27. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000 Apr;85(3):317-332. [Medline: 10781906]
28. Sullivan MJ, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, et al. Theoretical perspectives on the relation between catastrophizing and pain. Clin J Pain 2001 Mar;17(1):52-64. [Medline: 11289089]
29. Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med 1997 Jan;27(1):191-197. [Medline: 9122299]
30. McCracken LM, Yang SY. The role of values in a contextual cognitive-behavioral approach to chronic pain. Pain 2006 Jul;123(1-2):137-145. [doi: 10.1016/j.pain.2006.02.021] [Medline: 16564627]
31. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991 May;18(5):728-733. [Medline: 1865419]
32. Saris-Baglama RN, Dewey CJ, Chisholm GB, Plumb E, King J, Rasicot P, et al. Quality Metric Outcomes Scoring Software 4.5. User's Guide. Quality Metric Incorporated: Lincoln; 2011.
33. McCracken LM. Contextual cognitive-behavioral therapy for chronic pain. Seattle, Wash: IASP Press; 2005.
34. Dahl J, Wilson KG, Luciano C, Hayes SC. Acceptance and commitment therapy for chronic pain. Reno, NV: Context Press; 2005.

35. Vowles KE, McCracken LM, Eccleston C. Processes of change in treatment for chronic pain: the contributions of pain, acceptance, and catastrophizing. Eur J Pain 2007 Oct;11(7):779-787. [doi: 10.1016/j.ejpain.2006.12.007] [Medline: 17303452]
36. Bailey KM, Carleton RN, Vlaeyen JW, Asmundson GJ. Treatments addressing pain-related fear and anxiety in patients with chronic musculoskeletal pain: a preliminary review. Cogn Behav Ther 2010 Mar;39(1):46-63. [doi: 10.1080/16506070902980711] [Medline: 19697175]
37. Veehof MM, Oskam MJ, Schreurs KM, Bohlmeijer ET. Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. Pain 2011 Mar;152(3):533-542. [doi: 10.1016/j.pain.2010.11.002] [Medline: 21251756]
38. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New York: Psychology Press; 1988.
39. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
40. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of Web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
41. Walitt B, Fitzcharles MA, Hassett AL, Katz RS, Häuser W, Wolfe F. The longitudinal outcome of fibromyalgia: a study of 1555 patients. J Rheumatol 2011 Oct;38(10):2238-2246. [doi: 10.3899/jrheum.110026] [Medline: 21765102]
42. Rini C, Williams DA, Broderick JE, Keefe FJ. Meeting Them Where They Are: Using the Internet to Deliver Behavioral Medicine Interventions for Pain. Transl Behav Med 2012 Jan 3;2(1):82-92 [FREE Full text] [doi: 10.1007/s13142-011-0107-2] [Medline: 22924084]
43. Ly KH, Dahl J, Carlbring P, Andersson G. Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. SpringerPlus 2012;1:11.
44. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ACT: acceptance and commitment therapy
CBT: cognitive behavioral therapy
CPAQ: chronic pain acceptance questionnaire
CPVI: chronic pain values inventory
CWP: chronic widespread pain
ES: effect size
FIQ: fibromyalgia impact questionnaire
GHQ: general health wuestionnaire
iCBT: Internet-based cognitive behavioral therapy
ITT: intention-to-treat
LOCF: last observation carried forward
M: mean
MI: multiple imputations
PCS: pain catastrophizing scale
SF: short-form health survey
SMS: short message service
VAS: visual analog scale

Edited by G Eysenbach; submitted 15.11.12; peer-reviewed by M Buhrman; comments to author 10.12.12; revised version received 15.01.13; accepted 13.02.13; published 28.03.13.

Please cite as:
Kristjánsdóttir ÓB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, Wigers SH, Eide H
A Smartphone-Based Intervention With Diaries and Therapist Feedback to Reduce Catastrophizing and Increase Functioning in Women With Chronic Widespread Pain. Part 2: 11-month Follow-up Results of a Randomized Trial
J Med Internet Res 2013;15(3):e72
URL: http://www.jmir.org/2013/3/e72/
doi:10.2196/jmir.2442
PMID:23538392

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