Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 7 (2005), Issue 5 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Editorial

Original Papers

Viewpoints

Editorial

Going, Going, Still There: Using the WebCite Service to Permanently Archive Cited Web Pages

Gunther Eysenbach[1], MD, MPH; Mathieu Trudel[1], MASc

Centre for Global eHealth Innovation, Toronto, Canada

Corresponding Author:
Gunther Eysenbach, MD, MPH
Centre for Global eHealth Innovation
University of Toronto and University Health Network
190 Elizabeth Street
Toronto ON M5G 2C4
Canada
Phone: +1 416 340 4800 ext 6427
Fax: +1 416 340 3595
Email: geysenba@uhnres.utoronto.ca

Abstract

Scholars are increasingly citing electronic “web references” which are not preserved in libraries or full text archives. WebCite is a new standard for citing web references. To “webcite” a document involves archiving the cited Web page through www.webcitation.org and citing the WebCite permalink instead of (or in addition to) the unstable live Web page. This journal has amended its “instructions for authors” accordingly, asking authors to archive cited Web pages before submitting a manuscript. Almost 200 other journals are already using the system. We discuss the rationale for WebCite, its technology, and how scholars, editors, and publishers can benefit from the service. Citing scholars initiate an archiving process of all cited Web references, ideally before they submit a manuscript. Authors of online documents and websites which are expected to be cited by others can ensure that their work is permanently available by creating an archived copy using WebCite and providing the citation information including the WebCite link on their Web document(s). Editors should ask their authors to cache all cited Web addresses (Uniform Resource Locators, or URLs) “prospectively” before submitting their manuscripts to their journal. Editors and publishers should also instruct their copyeditors to cache cited Web material if the author has not done so already. Finally, WebCite can process publisher submitted “citing articles” (submitted for example as eXtensible Markup Language [XML] documents) to automatically archive all cited Web pages shortly before or on publication. Finally, WebCite can act as a focussed crawler, caching retrospectively references of already published articles. Copyright issues are addressed by honouring respective Internet standards (robot exclusion files, no-cache and no-archive tags). Long-term preservation is ensured by agreements with libraries and digital preservation organizations. The resulting WebCite Index may also have applications for research assessment exercises, being able to measure the impact of Web services and published Web documents through access and Web citation metrics.

(J Med Internet Res 2005;7(5):e60) doi:10.2196/jmir.7.5.e60

KEYWORDS

Archives; editorial policies; information management standards; digital libraries; periodicals standards; publishing standards; information storage and retrieval; Internet

Going, Going, Gone: Lost Internet References as a Growing Problem

Authors of scholarly publications increasingly cite (non-journal) Web pages and other Web-accessible documents in their articles. These cited materials may include for example descriptions of organizations on websites, “grey” research reports which have been published as Web page or Portable Document Format (PDF) files on the Web, online questionnaires, or even data files accessible for example through national statistics websites. As a general rule, published research should be transparent, replicable and falsifiable, and readers should have access to the cited materials, ideally seeing exactly the version authors saw when they cited the material. Yet, cited Web documents are at risk to be changed or even to disappear overnight, being unavailable for future generations of scholars. The unstable nature of Web references is increasingly recognized as a problem within the scientific community, and has been the subject of recent research and science policy discussions [1-8]. It also has been referred to as an issue “calling for an immediate response” by publishers and authors [7]. While services such as the Internet Archive or Google offer archiving (caching) of Internet documents, this is done randomly, does not focus on academic references, and cannot be initiated by authors, editors, or publishers wanting to cache a specific Web reference.

In journals like JMIR, where authors refer to Web services and online information perhaps more often than in other journals, the problem of “link rot” (“broken” links) in the references is particularly pertinent.

The Solution: Archiving Cited References With WebCite

To prevent “link rot” in scholarly references, JMIR is now among the first journals to adopt a new technology called WebCite (http://www.webcitation.org), which is designed to permanently archive and retrieve cited Internet references. This tool can be used by authors, readers, editors and publishers. It is free of charge for individual scholars (authors and readers), with participating publishers ultimately carrying the operating costs through a membership fee, similar to the CrossRef model, which is a not-for-profit consortium of publishers working on crosslinking between “traditional” journal references which carry a Digital Object Identifier (DOI). The WebCite consortium complements the CrossRef system as it caters to “non-traditional” cited material which does not carry a DOI, and which is therefore typically not permanently preserved in libraries or systems like the LOCKSS (Lots of Copies Keep Stuff Save) project at Stanford University [9].

The following briefly outlines how different groups of stakeholders such as scholars, editors and publishers can use WebCite. JMIR has amended its “Instructions for Authors” accordingly, asking its authors to archive (“cache”) cited URLs preferably before submitting a manuscript [10]. This ensures that peer-reviewers and readers have permanent access to the same version of the cited URL as the author. Thus, the following section entitled “Using WebCite as a ‘Citing Author’” is most pertinent for JMIR authors. If JMIR authors fail to cache cited URLs, archiving will be done later in the article production process, as described below under “Using WebCite as an Editor or Publisher”. However, in these cases the captured version of the cited page may differ from the version the author intended to cite if it has changed between the original access date and the article's processing date, thus authors are urged to prospectively cache cited Web material as early as possible, for example when they create a record in their bibliographic reference management software such as Reference Manager.

Using WebCite as a Citing Author

On the first level, the caching process can be initiated by the author of a manuscript wishing to cite a Web page (authors should note that it is usually *not* necessary to cache electronic *journal* articles if they have a DOI, as it can be assumed that these are permanently preserved in libraries. However, free articles from e-journals which appear not to be available in libraries, those without an ISSN and/or a DOI should be archived in case they vanish).

To initiate the process, the author goes to webcitation.org and submits the cited URL for archiving before citing it. This process is called to “WebCite®” a Web page or website. The WebCite tool takes a snapshot of the cited Web page and returns a “permalink” (permanent link), which the author should cite in the references section instead of (or in addition to) the unstable live link.

Alternatively, authors may also use a WebCite bookmarklet. A bookmarklet is a small JavaScript program that can be stored as a URL within a bookmark in most popular web browsers, or within hyperlinks on a Web page. The WebCite bookmarklet can be downloaded from the WebCite server and saved to the bookmarks (“favourites”) folder of any Web browser, so that the author can take a snapshot by selecting the bookmarklet whenever he encounters a Web page he might later want to cite.

Other third-party vendors may develop further tools such as browser plug-ins or add-ons to reference management software.

Authors may also cache multiple URLs by initiating a “combing” of a manuscript for URLs (currently this only works for HTML manuscripts). A request to comb the outbound links from a given HTML manuscript leads the WebCite server to present a checklist of outbound links from a manuscript to the user, who can then choose to archive the content of any of the outgoing links. This method is intended to be used during the prepublishing phase of manuscript preparation, in order to capture the content of cited Web pages which the author may have not archived with WebCite during their primary Internet search and writing up of the article. This method is deficient in that the captured version of the cited page may differ from the version the author intended to cite if it has changed between the original access date and the article's processing date. However, in cases where the original author did not include WebCite backed links for their references, this is nonetheless a better approach than simply not caching references at all.

Using WebCite as a Cited (Web)Author

Authors of online documents and websites which are expected to be cited by others can ensure that their work is permanently available by creating an archived copy using WebCite and providing the citation information including the WebCite link on the Web document(s). They may also put the WebCite bookmarklet as a link on the page(s) they expect to be cited. In the future, cited authors will also be able to retrieve WebCite statistics as an impact measure from webcitation.org.

Using WebCite as a Reader

Once the page(s) in question have been cached by WebCite, they can be accessed by users and publishers through the webcitation.org server, usually – if implemented by the publisher - just by clicking on a WebCite link next to the reference in question (see references [9] and [10] of this article for examples).

Depending on the information a user has at hand, items cached by WebCite can be queried based on one of three methods: By explicit WebCite ID, by URL and citing article (DOI), or by URL and date.

Retrieval of a cached document by explicit WebCite ID (snapshot ID) is the preferred way to retrieve a specific snapshot. Every item added to the WebCite database (including Web pages, PDF files, and included images or stylesheets) is assigned a unique numeric ID. These IDs are unique, unambiguous and idempotent, and thus represent the ideal way of querying a given resource cached by WebCite. However, the use of this method requires knowing the ID for a given resource, and so it can not be used without premeditation. Upon completion of an archiving request, WebCite sends an email to the user who requested the archiving operation (or, in the case of an FTP uploaded file, to the prearranged technical contact for the DOI prefix of the citing article) containing the WebCite link with the unique ID.

Also possible is retrieval by URL and date. When queried in this manner, WebCite finds all cached versions of the given URL, and sorts them by proximity to the given date. Although this allows for a certain “fudge factor” with timestamps, it also means that these types of queries are inherently ambiguous, and are not guaranteed to be idempotent across queries. As such, these queries are intended to be used when the user has no information in hand other than the URL to query, and possibly the approximate date of the snapshot they would like to see.

The last option, by URL and citing article (identified through its DOI), is the preferred way many publishers may implement a WebCite link. This method allows for publishers which use WebCite as part of their (pre-)publishing workflow to easily create WebCite queries for their cached references with minimal coordination with WebCite before publishing. Publishers submit the citing articles (as XML) shortly before, on or after publication to WebCite, which automatically caches cited URLs. These queries are unambiguous, but are not necessarily guaranteed to be idempotent (the content of the URL may be recached by multiple submissions of a given page for combing).

Using WebCite as an Editor or Publisher

Journal editors and publishers can use WebCite at three different stages: At the (pre-)submission stage, the copyediting stage, and the publication stage. Ideally, an editor or publisher works with WebCite at all three stages.

On the first level, editors should ask their authors to cache all cited URLs “prospectively” before submitting their manuscripts to the journal, by adding a respective note to their “Instructions for authors” (see this journal [10] for an example).

On the second level, editors and publishers should instruct their copyeditors or “technical” editors (who are in charge of preparing the accepted document for publication) to cache cited Web material if the author has not done so already.

On the third level (a process that is currently tested with JMIR and BioMed Central as early adopters), publishers can submit the raw citing article to WebCite for processing. Ideally, this submission is done via file transfer protocol (FTP), and uses a well defined (preferably XML based) schema for article data. The exact dialect used for this purpose should be agreed on ahead of time by the publisher and WebCite. Currently, WebCite supports (X)HTML documents, NLM Journal Publishing DTD documents, and BioMed Central Article DTD documents. Adding new document types to this list is a straightforward process, and can be undertaken on a publisher by publisher basis by providing WebCite with a document DTD and sample document for testing.

While the first two levels are currently free of charge for publishers, the third level requires that the publisher becomes a member of the WebCite consortium.

A fourth option on how a publisher can use WebCite is retrospective archiving. WebCite also works as a focused crawler, and can – in collaboration with publishers - automatically comb citing articles “retrospectively” for cited URLs. The focused crawler can also be pointed to domains hosting academic journals, which have previously not asked authors to “WebCite” references before submission. Retrospective archiving has the obvious limitation that by the time references are being “WebCite archived” they may have disappeared already.

In a pilot test of the WebCite focussed crawler, WebCite analyzed 280752 references from 8381 articles published in all BioMed Central journals from August 1997 to April 5, 2005. 6627 (2.4%) of these references were “pure” URL citations (i.e. not a URL of a journal article etc.), of which 4919 were unique. 1571 cited an entire domain (i.e. a website as opposed to a specific webpage). 2938 cited a HTML page, 222 a PDF file, and 15 txt/doc files. Obeying a variety of robot-exclusion standards and “no-archive”/”no-cache” metatags or copyright restrictions, WebCite succeeded to archive 3198/4919 (65%) Web pages. 500 were not cached due to robot exclusions, but only 8 had a no-archive and 7 had no-cache restrictions. The remaining Web pages could not be cached because they were already inaccessible or had disappeared.

Due to the limitations of retrospective caching, prospective archiving of cited references by authors or publishers at the time the manuscript is written or published is the preferred way to solve the problem of unstable and dynamically changing Web citations. Since its official launch in October 2005, almost 200 journals are already using WebCite on a routine basis.

Copyright and Long-Term Preservation Issues

Two of the most frequently asked questions about WebCite concern copyright issues and long-term preservation issues.

First, how does WebCite deal with copyright? Caching and archiving Web pages is widely done (e.g. by Google, Internet Archive etc.), and is not considered a copyright infringement, as long as the copyright owner has the ability to remove the archived material and to opt out. In order to opt out, certain machine readable Internet standards are in use, such as robot exclusion standards, as well as no-cache and no-archive tags, which are all honoured by WebCite. Thus, Web authors of copyrighted material who do not want their work cached or permanently preserved can explicitly exclude it from being archived simply by including these standard tags. In addition to honouring the respective Internet standards, copyright owners of an archived Web page also may request manual removal. In the vast majority of jurisdictions, caching Web pages would also considered “fair use”, in particular because 1) usually only one Web page as part of a larger collection (a website) is quoted, 2) because the webcited document is usually “unpublished” in traditional venues, hence there is no economic impact, with the vast majority of cited authors actually benefiting from the citation, 3) because the webcitation is used in the context of research.

Despite these arguments it has to be acknowledged that copyright legislation and jurisdiction in this area are complex and in a state of flux, and part of a future iteration of WebCite may include a comprehensive licensing management system, allowing to pay royalties to authors of archived material, should they wish so.

Secondly, how can scholars and publishers who opt to use WebCite be sure that the webcitation permalinks themselves will never be broken, that webcitation.org will never disappear? The answer is threefold: First, through the largest possible degree of "openness": All WebCite code is Open Source, and all documentation is licensed under Creative Commons licenses. Secondly, through collaborations with libraries and consortia interested in preservation of digital material, who may act as a curator, custodian or trustee for the WebCite project. These long-term preservation partners may have agreed to hold backups of the service and to legally assume the domain name, all intellectual property such as trademarks, and the service itself, should for any reason the original WebCite service go out of business. Thirdly, the WebCite consortium will eventually be owned by (or through a membership scheme run by) publishers, who all have a vested interest in keeping the service alive.

Beyond Archiving: The WebCite Index as a Retrieval and Impact Evaluation Tool

Widespread adoption of the WebCite technology among scholars, editors and publishers will not only solve the problem of inaccessible cited documents, but also open up further possibilities, such as the building – in analogy to the "Science Citation Index" (SCI) – a global "WebCite Index" which has been proposed as early as in 1998 [11]. Such an index can be used as a tool to evaluate electronic publications and websites which are published outside of the traditional peer-reviewed journal publishing route. Currently, websites and electronic documents, even if they are cited heavily, contribute little to a researchers' career or institutions' reputation, as they are inadequately captured in the Science Citation Index and in traditional research assessment exercises. Data stored in the WebCite Index can the basis to calculate Web impact measures (the number of times Web documents are cited or accessed provides quality indicators of their importance), activity measures (indicators of research and development activity in the subject areas) and linkage measures (indicators of intellectual linkages between authors/organizations and knowledge linkage between their subject areas).

Acknowledgements

MJ Suhonos and Jean-Sebastian Dumais coded earlier versions of the WebCite service.

Conflicts of Interest

While the WebCite service is currently operating not-for-profit, the corresponding author is inventor of the WebCite technology and owns the registered trademark WebCite® (USPTO Reg. 3,030,242).

References

1. Schilling LM, Kelly DP, Drake AL, Heilig LF, Hester EJ, Dellavalle RP. Digital information archiving policies in high-impact medical and scientific periodicals. JAMA 2004 Dec 8;292(22):2724-2726. [Medline: 15585731] [doi: 10.1001/jama.292.22.2724]
2. Veronin MA. Where are they now? A case study of health-related Web site attrition. J Med Internet Res 2002 Nov 22;4(2):E10 [FREE Full text] [Medline: 22442442] [doi: 10.2196/jmir.4.2.e10]
3. Hester EJ, Heilig LF, Drake AL, et al. Internet citations in oncology journals: a vanishing resource? J Natl Cancer Inst 2004;96:969-971. [Medline: 15199120]
4. Johnson KR, Hester EJ, Schilling LM, Dellavalle RP. Addressing internet reference loss. Lancet 2004 Feb 21;363(9409):660-661. [Medline: 14987901] [doi: 10.1016/S0140-6736(04)15608-0]
5. Kelly DP, Hester EJ, Johnson KR, et al. Avoiding URL reference degradation in scientific publications. PLoS Biol 2004;2:E99. [Medline: 15094803] [doi: 10.1371/journal.pbio.0020099]
6. Crichlow R, Winbush N, Davies S. Accessibility and accuracy of web page references in 5 major medical journals. JAMA 2004 Dec 8;292(22):2723-2724. [Medline: 15585729] [doi: 10.1001/jama.292.22.2723-c]
7. Dellavalle RP, Hester EJ, Heilig LF, Drake AL, Kuntzman JW, Graber M, et al. Information science. Going, going, gone: lost Internet references. Science 2003 Oct 31;302(5646):787-788. [Medline: 22954831] [doi: 10.1126/science.1088234]
8. Spinellis D. The decay and failures of web references. Commun ACM 2003;46:71-77. [doi: 10.1145/602421.602422]
9. LOCKSS – How it works. URL: http://lockss.stanford.edu/works/how_it_works.htm [accessed December 29, 2005] [WebCite Cache ID 95899]
10. JMIR Instructions for Authors. URL: http://www.jmir.org/?Instructions_for_Authors:Instructions_for_Authors_of_JMIR#webcite [accessed December 28, 2005] [WebCite Cache ID 95905]

11. Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. BMJ 1998 Nov 28;317(7171):1496-1500 [FREE Full text] [PMC: 9831581] [Medline: 99051215]

Abbreviations

DOI: Digital Object Identifier
DTD : Document Type Definition
HTML: Hypertext Markup Language
LOCKSS: Lots of Copies Keep Stuff Save
NLM: National Library of Medicine
PDF: Portable Document Format
URL: Uniform Resource Locator
XML: eXtensible Markup Language

Submitted 27.12.05; this is a non–peer-reviewed article;accepted 29.12.05; published 30.12.05.

Please cite as:
Eysenbach G, Trudel M
Going, Going, Still There: Using the WebCite Service to Permanently Archive Cited Web Pages
J Med Internet Res 2005;7(5):e60
URL: http://www.jmir.org/2005/5/e60/
doi:10.2196/jmir.7.5.e60
PMID:16403724

Original Paper

Computers and the Internet: Tools for Youth Empowerment

Ruta K Valaitis, BScN, MHSc, PhD

Corresponding Author:
Ruta K Valaitis, BScN, MHSc, PhDAssociate Professor
School of Nursing
McMaster University
HSC Room 3N28E
1200 Main Street West
Hamilton, ON L8N 3Z5
Canada
Phone: +1 905 525 9140 ext 22298
Fax: +1 905 521 8834
Email: Valaitis@mcmaster.ca

Abstract

Background: Youth are often disenfranchised in their communities and may feel they have little voice. Since computers are an important aspect of youth culture, they may offer solutions to increasing youth participation in communities.

Objective: This qualitative case study investigated the perceptions of 19 (predominantly female) inner-city school youth about their use of computers and the Internet in a school-based community development project.

Methods: Youth working with public health nurses in a school-based community development project communicated with local community members using computer-mediated communication, surveyed peers online, built websites, searched for information online, and prepared project materials using computers and the Internet. Participant observation, semistructured interviews, analysis of online messages, and online- and paper-based surveys were used to gather data about youth's and adults' perceptions and use of the technologies. Constant comparison method and between-method triangulation were used in the analysis to satisfy the existence of themes.

Results: Not all youth were interested in working with computers. Some electronic messages from adults were perceived to be critical, and writing to adults was intimidating for some youth. In addition, technical problems were experienced. Despite these barriers, most youth perceived that using computers and the Internet reduced their anxiety concerning communication with adults, increased their control when dealing with adults, raised their perception of their social status, increased participation within the community, supported reflective thought, increased efficiency, and improved their access to resources.

Conclusions: Overall, youth perceived computers and the Internet to be empowering tools, and they should be encouraged to use such technology to support them in community initiatives.

(J Med Internet Res 2005;7(5):e51) doi:10.2196/jmir.7.5.e51

KEYWORDS

Computers; Internet; adolescent; power; public health; schools; social change

Introduction

The use of computers and the Internet can aid communities by supporting communication and access to information, thereby building social capital and community capacity [1,2]. Using computers can assist the community planning process [3], community participation, and information sharing [4-6]. Computer-mediated communication can build community awareness, encourage local decision making and dialogue between groups, and support disadvantaged communities. Status barriers can be reduced [7], and online communication with disenfranchised groups, such as women, people of color, and those with disabilities, can be promoted [8].

Youth are among the disenfranchised groups. Adults typically view youth as the cause of community deterioration rather than as a community asset [9,10]. Youth often feel they have little voice in their communities [11,12]. Youth participation in their communities can positively influence programs so that they are more responsive to youth's needs [3] and can help support youth's sense of self-determination from a community and individual perspective [4,5], thereby promoting their health. Increasing youth community participation, however, has been problematic. Since computers are an important aspect of youth culture, they may offer solutions to increasing and supporting community participation.

Youth have used computer technology to support various community projects [6-19]. Websites are seen as potential vehicles to support community building among groups that confront prejudice and domination [20]. Hart [21] argues that electronic publishing has the potential to broaden the reach of children's voices and provide instant feedback. Cockburn [22] supports that children can use information and communication technologies to increase their participation in public life through better access to information, collective action, a more level political playing field, and the ability to include their views in decision making. Despite the claims of the benefits and barriers of computer technology, no research was found that addressed its use in community development work with youth.

The purpose of this qualitative case study was to explore youth's use and perceptions of computers and the Internet as tools to support them in a school-based community development project. Objectives were to examine how youth used these tools and to explore youth's perceptions of how the use of these tools affected participation within their community and communication with community members.

Methods

Study Design

This case study was conducted in a Southwestern Ontario inner-city school. Nurse facilitators (the school public health nurse, 2 senior nursing students, and the author) worked with 19 well-functioning grade seven and eight students for 12 weeks on a school-based community development project. Criteria for eligibility included demonstrated responsibility and the ability to manage academically by missing a 50-minute school period three times per week. The principal saw the community development project as an enrichment program for youth who could benefit from the challenge. To increase project sustainability, it was decided to include grade seven students who could carry on in the following school year. The principal selected a group of 35 students in grades seven and eight in consultation with teaching staff. Of the 35 who were selected, 23 consented to participate, and of these, 19 completed the project. Parental consent was also obtained. Participants were primarily female (79%, n = 15) and were evenly distributed between grade seven (n = 13) and eight (n = 12), including 6 participants who were visible minorities and 13 who were on the school honor roll. Many of the selected youth held other duties in the school, such as office and lunch hour supervision. A number of them came from single parent families. Initially, the youth generally felt that they had little voice in their community and were not trusted by adults [12].

Over 12 weeks, facilitators and youth met during school hours, three days per week, for one and a half hours each day. Youth identified needs and assets of their school and neighboring community, prioritized problems, and planned and implemented actions to improve their environment. Community assessment activities included mapping where youth live, work, learn and play; conducting a neighborhood walkabout; and photographing images to illustrate community needs and assets. The youth participated in a visioning session, interviewed key community members, conducted face-to-face interviews, and surveyed peers online about their views. Small groups of 4 to 8 youth worked on specific tasks with at least one adult facilitator. Large group sessions were also held to define community boundaries, decide on group goals, and share updates. Occasionally, a group of 7 youth met independently to refine computer and Internet skills and work on computer-based activities to support the project.

Participants identified many health issues, such as violence, drugs and alcohol, pollution, and smoking. They eventually chose to work on "do-able" school improvements. They enhanced the school's general appearance by removing graffiti from school walls and beautified the school environment through a small greening project. They lobbied for improvements in washrooms and prepared two proposals: to establish a student-run school store and student council. Factors that enabled and constrained the community development process for youth are described elsewhere [23].

Youth accessed 6 computers in the library that were used to (1) construct websites containing their community assessment findings, (2) survey peers online, (3) create documents, (4) access information and resources through the Internet, and (5) communicate with each other, project facilitators, school staff, and community members using a computer-mediated communication system (FirstClass). This system, hosted by the local university, was used because it provided a private password-protected communication space for project participants. Participants could write private email messages or post messages to the group's bulletin board, which could be viewed by all project participants. Community members, who were trained and given access to the system, could post and read all existing notes. This provided a useful archive of all messages. An example of one youth's message enlisting help from an adult can be found in Textbox 1.

Textbox 1. Letter composed by a youth participant

> Hello. This is a message from the XXX School Community Development project. Hopefully, you [have received] this message [that was] sent to XXX Park. We have written this message to you in hopes that you will view our web site and give us ideas on how to improve our school with our "Greening and Cleaning" project to beautify our school. We could use any form of help from you. Maybe some people who work there or volunteer could visit XXX School and accompany us with some of the work. We are planning to start planting a small garden this spring. Maybe you could be able to help us plant and plan a long-term garden. Please e-mail us back, and if you can find any time please visit our web site at http://www.learnlink.mcmaster.ca/POWer.
>
> Sincerely,
>
> XXXX

Research Methodology

This qualitative case study used multiple methods of data collection, including participant observation, field notes, youth interviews (conducted individually, in pairs, or in a group of 3), facilitator interviews (conducted individually), online- and paper-based surveys, and quantitative analysis of computer-mediated communication notes. Interviews were transcribed, reviewed, and edited as needed. Table 1 outlines the timeline, method, and purpose of each data collection method. Ethics approval was obtained from the University of Toronto, Research Ethics Board.

Table 1. Data collection methods and purpose, in chronological order

Time	Data Collection Method	Purpose
Week 4	Open-ended, paper-based anonymous survey	To explore youth's general feelings about the use of computers in the project
Week 9	Open-ended, online survey in the computer-mediated communication system	To explore youth's perceptions about the community development project and computers
Weeks 1-13	Participant observation and raw field notes (Reflective notes were also added to the field notes.)	To obtain a detailed description of all project meetings, events, and communications with project participants (Reflective notes provided an audit trail.)
Weeks 14, 15	One-hour youth interviews (singly, in pairs, and, in one case, with a group of 3) (All interviews were taped, transcribed, checked, and edited for accuracy in transcription.)	To explore youth's perceptions about the use of computers in the project
Weeks 14, 15	One-hour adult facilitator interviews (All interviews were taped, transcribed, checked, and edited for accuracy in transcription.)	To enhance the author's interpretation, adult facilitators were interviewed about their impressions of youth's perceptions of computers in the project.
End of study	Quantitative analysis of notes	To investigate youth's involvement with online notes and to determine who wrote and opened notes

Richards and Richards suggest "working up from the data" and later reflecting and exploring it to form impressions and summaries [24] (p. 466). Qualitative coding of the interview transcripts, field notes, and open-ended surveys was conducted using this approach. The constant comparative method [25] was used to support the development of themes using qualitative analysis software (ATLAS.ti, ATLAS.ti Scientific Software Development GmbH, Germany). Two types of triangulation were used to build credibility of the qualitative findings [26]. One type used the constant comparison method, where themes were identified repeatedly, thus satisfying the existence of the theme. The second form was between-method triangulation. Interviews, participant observation, paper-based and online surveys as well as analysis of computer-mediated communication notes were used to gain an in-depth understanding of the phenomenon in question. Data can converge to a single proposition or demonstrate inconsistencies or contradictions. These differing outcomes in the analysis are valuable to consider to "construct meaningful propositions" or explanations about the data [27] (p. 15). The paper presents emergent themes, as well as inconsistencies uncovered in the analysis.

Results

The first section briefly reports on technology use by youth. Following this, themes are presented related to youth's perceptions. Quotes from male youth are marked as such.

Use of Technology

Youth worked with computer technology in varying degrees: 18 (94.7%) youth developed Web pages, and 10 (52.6%) youth conducted an online survey of their peers' opinions about community needs and assets. Of the 177 notes that were posted in the computer-mediated communication system, project youth posted 117 (66.1%) notes, whereas adults posted 60 (33.9%) notes. On average, youth created 6 notes (range 2-11, SD 2.4) and opened 20 notes (range 3-62, SD 15.7).

Youth's Perceptions of Technology

Four major themes describe youth's perceptions of computer and Internet technology used in the project: reduced social risk factors, increased community participation, increased opportunity for reflection, and increased resources (Textbox 2). Inconsistencies are also presented.

Textbox 2. Themes and subthemes of youth's perceptions of computer and Internet technology

1. **Reduced Social Risk Factors**
    - Reduced anxiety
    - Increased control
    - Increased social status
2. **Increased Participation in Community**
    - Sharing youth's views with the world
    - Getting others' opinions
    - Getting access to influential people
3. **Increased Opportunity for Reflection**
4. **Increased Resources**
    - Increasing efficiency
    - Increased access to information and materials
    - Improved record of progress

Reduced Social Risk Factors

Three subthemes explain the major theme of reduced social risk factors. Youth felt that using computers reduced their anxiety, increased their control, and increased their social status.

Reduced Anxiety

Youth wrote to adults online to get information and advice. For example, youth emailed the local police liaison officer for ideas about vandalism prevention in the school and contacted the local botanical gardens for advice on plants suitable for a school greening project. Many youth felt that writing to adults via the Internet significantly lessened their anxiety compared to face-to-face or phone discussions. Youth were asked how they felt about constructing online messages to adults. SL (male) thought that "you wouldn't stutter or have trouble saying what you want." This was explored further in interviews. PF explained, "Usually people get choked up over the [phone] line. If you write it, it is easier to say things." Two girls described their experience:

> *[On the computer] it wasn't like when you're in person. You're kind of nervous with adults.* [LP]
>
> *You don't want to say the wrong thing.* [CB]
>
> *They get mad at you or something.* [LP]

Computer-mediated communication was seen as neutral ground between youth and adults. YB (male) said, "If you talk face-to-face, it scares you.... On the Internet, it's okay. You don't know them, they don't know you. You [get] along." Two youth who wrote to adult facilitators further explained the safe online environment:

> *On the computer you can say whatever you want. But when you are face-to-face you are afraid to tell [adults] what you want to do.... They might laugh at you. Laugh in your face. Tell you to get lost....* [CC]
>
> *If you go up to any person and just start telling them you would probably feel shy....* [CP]
>
> *The computer was easier. You can say whatever you want.* [CC]

Online communication provided a safe way to initiate discussions with adults. CP explained how it was easier to write to the project facilitators at first. "We have to know the person before. We didn't know you guys and we didn't start talking to you guys at first." As NT (male) explained, "[I'd rather] write [the police officer] an email first...because I never met the guy." TS composed notes to reduce her anxiety before telephoning an agency. She preferred email because "I don't have to talk."

Increased Control

Some youth spoke about feeling more *in control* communicating with community members using computer-mediated communication. As described by SL (male), it "gives you more of a backbone." Two youth appreciated being able to prepare their dialogue carefully for adults.

> *[When talking] you sound kiddish. If you don't write anything like...* [SL (male)]
>
> *Big words.* [SG]
>
> *Yeah. I was ready to pull out a thesaurus and figure out what to say.* [SL (male)]
>
> *I think, umm, it was better on the computer than talking because...* [SG]
>
> *It's like, umm. Uhuhuh mmm. When you're talking.* [SL (male)]
>
> *Yeah exactly, they can't see you [online].* [SG]

Youth felt more prepared to talk to adults because they could read their notes first. SL (male) stated, "It kind of gives you an idea of how they talk and how they think." Writing on the computer also allowed youth to better prepare their notes. As MM noted, they could "spell check...and make sure the grammar is right." Having time to read other's notes and prepare their own gave youth a feeling of being more in control.

Increased Social Status

Communicating via the Internet raised youth's perception of their social status. Many youth felt that they were more professional and were perceived as having "smartness" (YB, male). CB preferred communicating with adults online "because then people think you're smarter." She later commented that writing to the principal online promoted being heard: "It was more professional." A facilitator concurred that youth's image was positively influenced through the use of technology: "I think they would have been seen differently by the community because they were using the technology.... [They] were seen as a bit smarter, brighter." Youth and adults felt that youth's credibility was also enhanced.

> *[On the computer] adults know we're not just jokers around that do whatever they want.* [CP]
>
> *Yeah this isn't just kidding.* [CC]

A facilitator reflected that "It certainly lent more credibility to the students' perception of what they were doing and external perceptions of what the students were doing within the school." MM corroborated this view: "Sometimes they might not listen to you, because they hear your voice or something. They will actually read what you write.... They still know you're just a kid doing stuff. But they are more reading it, because you are not actually there looking at them. They are actually going to read it."

Therefore using online communication raised youth's perception of their social status and professionalism, in turn, increasing their credibility in the community.

Inconsistencies Regarding Increased Social Status

On two occasions, youth interpreted online messages from adults as negative and rude, leading to feelings of inferiority and low social status.

> *You're telling us to write proper and good and stuff, and then they come back and pretty much [call] us stupid.* [SG]
>
> *That's what I was thinking.* [SL (male)]
>
> *Yeah, thanks for your support. We're retards!* [SG]

In one incident, a school police liaison officer replied to a youth's online question, "What would you change in the community?" The officer replied, "Your first question is huge!! I need you to be more specific. I would change a lot of things if I had the authority to do so...." SL (male) interpreted this communication negatively: "[It] sounded like he was getting rude on the thing. But it was just the way he was talking. Sounded like he was saying...get smarter." In another incident, US (male) wrote to the school staff member asking, "What kind of things do you think that this community needs?" The school staff member replied, "Is this a wish list?" CP spoke about her perception of this message:

> *He is giving you these good ideas. And then after he says, "Is this a wish list or something?" He should have put it down different. Not like that. When you think of a wish list you think like, "Oh, I want this, I want that." Like you're being greedy.... You're not trying to be greedy. You are just trying to help out.* [CP]

The school staff member reflected on the incident:

> *You have to choose your words carefully because words can be perceived differently. So in general, I'm quite careful with how I respond to email, and I read everything over before I send it, just to make sure the tone of it is the way I want it to be. But I think I was even a little more careful with the kids….* [school staff member]

Youth were careful writing to adults. Although communicating online increased their sense of control, US (male) noted that, "It's different writing to the principal because I have to use proper grammar and stuff.... I was careful." SL also felt "kind of nervous. [I held] back what I was going to [write]. Trying not to offend them." A few youth felt awkward writing to adults. "I was used to writing to [my peers]. It was hard to write to someone older than I am" (BG). DC said, "I wasn't exactly worried, but it was hard [because] I couldn't exactly think of what to put down." PF worried about the interpretation of her notes: "I don't know if they will like it, if they will understand it…if it shows the whole concept." Thus, writing to adults was somewhat problematic for youth.

Although youth perceived themselves as smarter in the online environment and felt less anxiety overall, the process of online communication was not stress free. To protect their anonymity, cyber names were used, such as "Purrfection" and "Hellokitty." This had the negative effect of reducing youth's credibility in the eyes of community members. One community member was contacted by phone when an email response was not forthcoming; this community member considered the youth's email a prank. To build credibility, youth began to describe the project in the body of their emails and referred to the project website once it was launched. This solved the problem.

Increased Community Participation

Youth felt that technology helped to increase their participation in the community. This is explained by three subthemes: sharing youth's views with the community, getting other's opinions, and getting access to influential people.

Sharing Youth's Views With the Community

Youth felt strongly that they could share their thoughts about their neighborhood through websites. Willing youth were taught to construct a simple website containing ideas about their neighborhood and photographs. The Web pages reported on community problems, such as pollution, bullying, violence, and smoking. Areas of pride, such as their school, the local football stadium, and recreation center, were also highlighted. Project progress was also reported. In the interview, LP said, "We made websites so we could show people what we have done so far in the project." NT (male) felt that the Internet helped youth share their ideas: "You can tell what we like and don't like about [our] communities." MM felt that "more likely [youth] are going to write out what they think on the computer. They are not just going to come up and tell you." Photographs augmented youth's stories.

> *Because, you have pictures and stuff so kids could see. If there's just sentences and stuff it's like...* [LP]
>
> *It's hard to imagine that in your head.* [CB]
>
> *You might be thinking about something else but they're not talking about that. It might just sound like that.* [LP]
>
> *Yes. I just think [the website] communicated it very well.* [CB]

Web pages helped youth extend their reach into the community. CB explained, "People understood you better I guess.... If we didn't have the computers we wouldn't be able to reach the community that was the around us." Unfortunately, youth's websites were not well publicized. They were disappointed that only two school staff members reported visiting the site. Despite this, given another opportunity, all youth stated that they would build Web pages again. "It took us a lot of time [to build websites]. It's worth it" (YB). Youth highly valued the potential global reach of their ideas, even if it had no impact on the outcomes of the project itself. One facilitator felt this itself was empowering for youth:

> *[Working with computers] made them feel part of the bigger world. Whether or not they contributed to the community…they were certainly communicating with people out there, people that they might not have ever interviewed or connected with.... I think these kids were empowered in the sense that they knew that their work was displayed for the whole world to see, and I think that was empowering alone....* [facilitator]

Getting Others' Opinions

Youth gained peer input through the Internet, which helped them identify and prioritize community needs. Students in grades six to eight were asked to rate community problems and strengths in an anonymous online survey. Youth visited the classrooms and invited their peers to participate. LP wrote:

> *[Computers] helped us...figure out what we wanted to focus on. If we didn't have computers to do [online] surveys, we would probably take twice as long to figure it out. And we could chat with people to find out what they thought of what we have done."* [LP]

Two youth debated benefits and limitations of online surveys.

> *[I liked the online survey] because we got to get to results. We got a chance to use the survey to see the [opinions] of other people.* [CP]
>
> *If we didn't have computers...we couldn't do as much as we did. Because we couldn't get other people's opinions unless they wrote us a letter or came face-to-face and talked to us.* [CC]
>
> *I think it would be better with [face-to-face] interviews. You can get more answers. If you just give them a choice out of 3…they might have more to say. So on the computer, they can't write all their thoughts.* [CP]

Despite the limitations of forced choice surveys, the majority of youth felt that computers helped them obtain other's opinions about their community. MM stated, "We learned to take everyone into consideration no matter what they have to say. We learned that everything anyone says is a good idea, because it is [their] idea." Clearly, youth valued their peer's opinions and successfully included them in the project using technology.

Getting Access to Influential People

Some youth indicated that computers and the Internet provided access to influential people. DC wrote, "[I think computers have helped us] because we can ask important people if they have their own opinions on this community like [the police officer]...." BG shared that "[computers allowed] us to get important information from people outside of the school to help us with our project." PF explained, "Without the computers, we would have never gotten replies and such from the [Royal Botanical Gardens] or people like that." Through youth's Web pages, communication notes, and the online survey, youth viewed computers and the Internet as effective means to communicate with community members, including influential adults, other youth, and the broader community.

Increased Opportunity for Reflection

As a result of the asynchronous nature of computer-mediated communication, writing thoughts online gave youth time to reflect. A few youth felt that this helped clarify others' ideas. LP explained:

> *We could look and then we could come on another day and we could [write] to them about "What did you mean about that." [Then we could answer] "I think that's a good idea" or something.* [LP]

MM agreed:

> *It's going to be the same thing as you are going to say. But when you write it, you can think a lot more about what you are going to write. You don't have to worry about them being right there. You can go back and read it....* [MM]

Facilitators observed that youth took their time to communicate to adults online. A facilitator shared in her interview that "[computers] gave [youth] a chance to generate responses that were more thoughtful." CC noted that online polling provided more time for reflection compared to face-to-face interviews: "I think the computer is easier because people get to think more. If you're face to face…like if you're thinking a long time, people get might get bored of me." Technology provided more opportunity to think before acting.

Increased Resources

Participants felt that computers increased resources. This is explained by three subthemes: increased efficiency, increased access to informational and material supports, and a permanent record of progress.

Increased Efficiency

Computers and the Internet increased youth's efficiency in accessing resources. SL (male) felt that "you can get information faster and cheaper." CP explained, "The computer [online survey] gave you faster results." MM felt that technology provided a reliable and efficient means to communicate:

> *[By using computers] you know it's going to get to them. If you just call and leave a message, you have to worry about them calling you back. You can just check your email to see if they emailed you back. No waiting for them to call you back at a certain time.* [MM]

Increased efficiency was also noted in a facilitator's interview:

> *Email is quick and easy and accessible to everyone. At least that's the perception I think most of these students had.... [Youth] were able to write emails as soon as we found the sites that intrigued us or came up with ideas, and it was a very quick way, again, to start that dialogue…. They were able to do things right away.* [facilitator]

Youth, therefore, perceived a heightened sense of immediacy and access to quick feedback.

Information and Material Resources

Participants felt that the Internet enhanced the group's ability to search for and access community agencies, information, and material resources. Some youth searched for community contacts, such as the police and parks and recreation staff. They gathered information about forming a school council and starting a school greening project. MM felt that "[the Internet] helped with all our information that we needed to explore." Youth successfully obtained material resources. LP stated, "[through email we could] write to each other and to people [that] we needed supplies from." GC wrote, "I think the computers have helped to get donations from other people." PF wrote, "If we didn't have these computers, we would have never been able to email all those people that could help us." The project website helped publicize their project goals, which resulted in access to free resources. For example, the Children's Museum staff viewed the project website and then waived a room booking fee for a group function.

Permanent Record of Progress

Some youth felt that computer-mediated communication was valuable to document and store information. SL (male) identified that "the program holds information, [and] you can lose paper." Later, he commented online that "[computers] have helped us a lot by storing information…." MM indicated that "we wrote a lot of our ideas on the computer, and with that we can come back and read those ideas any time we want.... We will never have to worry about forgetting them." Thus, computer-mediated communication maintained a permanent record of project communications.

Inconsistencies Regarding Computers and the Internet as a Resource

A small number of participants generally disliked technology. They felt that using technology was time-consuming. BM felt "the logging in would take forever. And then you had all your mail you had to check. And then you had to do what you had to do, and it would take forever." GG spoke about contacting the local park: "[I would prefer] to phone because it will get there quicker." SM, who was very artistic, shared the following in the interview:

> *I don't like computers. If I would just [have] written it, it would have taken less time.... And it takes a long time to get into the program and we had to cut and paste it and all that stuff. I could have just written it and put it on pretty paper.* [SM]

Computers were not appealing to all youth. Many justifiably complained about slow connections and the Internet frequently being down.

Discussion

The major themes and subthemes that have been presented closely parallel a number of factors associated with powerlessness and empowerment in Wallerstein's Empowerment Education Model [28]. Therefore, these themes are interpreted using this model. Textbox 3 lists factors from the model that appear congruent with the major themes. The model supports the finding that youth perceived computers and the Internet to (1) reduce certain social risk factors thought to be associated with powerlessness and (2) to increase factors that are associated with empowerment. Despite technology limitations, computers and the Internet generally appear to be empowering tools for youth.

Textbox 3. Major concepts from Wallerstein's empowerment education model [28] that parallel themes

- **Powerlessness**
  - Social Risk Factors
  - High psychological demands
  - Low control
  - Low in hierarchy
- **Empowerment**
  - Increased Participation in Decision making
  - Critical Thinking/Conscientization
  - Resource Equity/Access

Overcoming Threats to Empowerment with Computers and the Internet

The themes—reducing levels of anxiety, gaining control, and increased social status—were related to online communication with adults. These themes parallel social risk factors in Wallerstein's model [28], namely high psychological demands, a feeling of low control, and being low in social hierarchy. Reduction of these factors supports empowerment. Computers and the Internet can, therefore, be viewed as supportive tools for youth to overcome threats to empowerment. Youth felt more confident, better prepared, and more knowledgeable about the adults with whom they communicated online.

Others found similar benefits from working with technology. Resnick et al [19] worked with youth in computer clubhouse projects where design experiences supported learning by giving youth a sense of control over the learning process. In this case study, youth felt in control communicating to adults online since they could manipulate the medium. They also felt increased status. Similarly, in an online nurse practitioner program, students identified that the novelty of taking a computer-based course and greater computer knowledge raised their status among their colleagues and families [29]. Reductions of social status differences have also been noted in computer-mediated communication research with adults in their work environments [7]. Further research into the impact of technology on communication processes between adults and youth is warranted.

Supporting Empowerment With Computers and the Internet

Wallerstein places emphasis on increasing participation in decision making to support empowerment [28]. Youth felt an increased sense of participation with their neighborhood and school community through online communication. They also gained others' perspectives using computer-mediated communication and online polling and, more importantly, shared their personal views on Web pages and in communication notes. Youth considered their website to be the most valuable strategy to express their ideas to the broader community. Resnick et al indicated that designing computer projects "facilitated personal connections to knowledge, because designers often develop a special sense of ownership (and caring) for the products (and ideas) that they design" [19] (p. 270). Perhaps this explains, in part, the value that youth placed on this activity. Wong et al [30] indicated that designing Web pages was a popular activity in Michigan elementary school computer clubs. They attributed this to three factors: providing youth with an authentic learning experience, increased participation with the broader community, and an increased sense of achievement. Community practitioners are thus encouraged to support youth in website creation.

This study supports websites as an effective narrative tool for youth. Rees [31] sees benefits from youth telling their story, in which a form of narrative therapy occurs and mutual education and sharing between young people and adults begins. Similarly, Schwab [32] reviewed community development initiatives and found that when youth tell their story, they grow. Media tools such as journals, masks, and drama that represent youth's lives were identified as valuable tools for self-expression and advocacy. Other creative methods include murals [33], zines (self-published magazines) [34], and videos [35]. Other researchers [36-39] describe using photography as a voice for community issues and as a method for participatory research and community assessment.

Most youth reported showing their personal Web pages to family and friends at home or at the public library. Despite the limitations for those without home Internet access, the potential for broad community reach through the websites and email was empowering. Wallerstein [28] indicated that stronger social networks support community empowerment. Networks can be enhanced through websites and online communication. Newer technologies such as Web conferencing and instant messaging have this potential.

Computers and the Internet were viewed by most youth as valuable resources. Wallerstein refers to a lack of resources, such as finances or access to systems, as a risk factor related to powerlessness, whereas access to resources supports community empowerment [28]. CB reflected that "I have learned that you can change your community if you have a lot of resources and a lot of people to help." Most youth felt that computers increased their efficiency; gave them more control; made information, people, and materials more accessible; and were useful for storing their ideas. Some youth commented that the time afforded to think and construct responses on the computer was beneficial. Bamberger [40] described work with youth in student labs at the Massachusetts Institute of Technology, which provided an environment for youth to "catch up with their own understandings—slowing down events and actions so as literally to grasp the 'goings on' of things and how they relate to ideas" (p. 239).

Limitations of Computers and the Internet

Youth noted limitations in using technology that had the potential to increase their feelings of powerlessness. Although online communication generally increased youth's perception of social status, at times, youth interpreted adults' online comments as "put downs." Community practitioners are cautioned when working with youth to word their email messages carefully. Care needs to be taken in selecting cyber names for youth. Since youth participants selected creative cyber names (eg, Sir Lancelot), their credibility was weakened. Further research into the impact of adults' different writing styles on youth's perceptions, and vice versa, is needed. Although synchronous communication was not used, one youth suggested "chatting" with neighboring schools. There is significant potential for synchronous communication to support community networks. A Canadian study by Skinner et al [41] found that the quality of Internet access for health information and resources was affected by privacy, gate-keeping, timeliness, and functionality. It would seem reasonable to consider these factors when applying computers and the Internet as tools to support future youth-driven community development initiatives.

A few youth indicated disinterest in computers and rarely chose to work with them other than to build Web pages; they were also indifferent or negative about the impact of computer technology on the project. Further research into what limits youth's engagement with computers, how computers can be used to serve more authentic purposes in community development, and how they can affect youth’s sense of empowerment is needed.

Conclusions

Overall, youth perceived computers and the Internet to be empowering tools in this community development project. Youth felt better able to participate in the community. The Internet provided a safe way to communicate with the neighborhood and school community by supporting youth’s ability to obtain others’ views and share their own. By communicating online, youth perceived themselves to have higher social status and increased credibility. The Internet’s potential to reach the broader community was empowering in itself. Technology was viewed as a useful resource that improved youth’s efficiency, supported critical reflection, and created a permanent record of their work. Despite the benefits, technical problems, computer access issues, the potential for online miscommunication, and the potential to raise youth’s feelings of inadequacy were drawbacks of using technology. This case study involved mostly female students who were high academic achievers. Thus, transferability of these results is limited to similar populations. It is unknown how results might differ with students with lower academic achievement or with a different gender mix.

Acknowledgments

I would like to acknowledge Drs. Lynn Davie and Ilze Kalnins for their support and assistance in the preparation of this manuscript.

I also acknowledge the Social Sciences and Humanities Research Council for its financial support through a doctoral fellowship, which helped to support this research.

Conflicts of Interest

None declared.

References

1. Goodman RM, Speers MA, Mcleroy K, Fawcett S, Kegler M, Parker E, et al. Identifying and defining the dimensions of community capacity to provide a basis for measurement. Health Educ Behav 1998 Jun;25(3):258-278. [Medline: 98277465]
2. Pigg K, Duffy Crank L. Building community social capital: the potential and promise of information and communication technologies. J Commun Informatics 2004;1(1):58-73.
3. Craig W. The Internet aids community participation in the planning process. Computers Environ Urban 1998;22(4):393-404. [doi: 10.1016/S0198-9715(98)00033-7]
4. Howley K. Equity, access, and participation in community networks: the case for human-computer interaction. Soc Sci Comput Rev 1998;16(4):402-410.
5. Shaw A, Shaw M. Social empowerment through community networks. In: High technology and low income communities: prospects for the positive use of advanced information technology. Cambridge, Mass: MIT Press; 1999:315-336.
6. Wellman B, Haase A, Witte J, Hampton K. Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. Am Behav Sci 2001;45(3):436-455. [doi: 10.1177/00027640121957286]
7. Tan B, Wei KL, Watson R, Walczuch R. Reducing status effects with computer-mediated communication: evidence from two distinct national cultures. J Manage Inform Syst 1998;15(1):119-141.
8. Van BE. Technology, democracy, and the creation of community. National Civic Review 2000;89(3):185-192. [doi: 10.1002/ncr.89301]
9. Davis A, Jones L. Whose neighbourhood? Whose quality of life? Developing a new agenda for children's health in urban settings. Health Educ J 1997;56:350-363.
10. Spencer C, Woolley H. Children and the city: a summary of recent environmental psychology research. Child Care Health Dev 2000 May;26(3):181-197. [Medline: 20376882] [doi: 10.1046/j.1365-2214.2000.00125.x]
11. Caputo T. Hearing the voices of youth: a review of research and consultation documents. Ottawa, Ontario: Health Canada; 1999:1-32.
12. Valaitis R. "They don't trust us; we're just kids:" views about community from predominantly female inner city youth. Health Care Women Int 2002;23(3):248-266. [Medline: 21998206] [doi: 10.1080/073993302317346325]
13. Meucci S, Schwab M. Children and the environment: young people's participation in social change. Social Justice 1997;24(3):1-10.
14. Hart C, Schwab M. Children's rights and the building of democracy: a dialogue on the International Movement for Children's Participation. Social Justice 1997;24(3):177-191.
15. Hart R, Daiute C, Iltus S, Kritt D, Rome M, Sabo K. Developmental theory and children's participation in community organizations. Social Justice 1997;24(3):33-63.
16. Milio N. Reaching out with information technology: educating children, engaging elders. Engines of empowerment: using information technology to create healthy communities and challenge public policy. Chicago, Ill: Health Administration Press; 1996:111-118.

17. Schon D, Sanyal B, Mitchell W. High technology and low income communities: prospects for the positive use of advanced information technology. Cambridge, Mass: MIT Press; 1999.
18. Tardieu B. Computer as community memory: how people in very poor neighborhoods made a computer their own. In: High technology and low income communities: prospects for the positive use of advanced information technology. Cambridge, Mass: MIT Press; 1999:287-314.
19. Resnick LB, Rusk N, Cooke S. The computer clubhouse: technological fluency in the inner city.. In: High technology and low income communities: prospects for the positive use of advanced information technology. Cambridge, Mass: MIT Press; 1999:263-314.
20. Uhler Cart C. Online computer networks: potential and challenges for community organizing and community building now and in the future. In: Minkler M, editor. Community organizing and community building for health. New Brunswick, NJ: Rutgers University Press; 1997:325-338.
21. Hart R. Children's participation: the theory and practice of involving young citizens in community development and environmental care. London: Earthscan Publications Ltd.; 1999.
22. Cockburn T. New information communication technologies and the development of a children's community of interest. Commun Dev J 2005 Jan. [doi: 10.1093/cdj/bsi016]
23. Valaitis R, O'mara L. Enabling youth participation in school-based computer-supported community development in Canada. Health Promot Int 2005 Sep;20(3):260-268. [Medline: 15764685] [doi: 10.1093/heapro/dah611]
24. Richards T, Richards L. Using computers in qualitative research. In: Handbook of qualitative research. Thousand Oaks, Calif: Sage Publications; 1999:445-462.
25. Lincoln Y, Guba E. Processing naturalistically obtained data. In: Naturalistic Inquiry. Beverly Hills, Calif: Sage Publications; 1995:332-356.
26. Cutcliffe JR, Mckenna HP. Establishing the credibility of qualitative research findings: the plot thickens. J Adv Nurs 1999 Aug;30(2):374-380. [Medline: 99388428] [doi: 10.1046/j.1365-2648.1999.01090.x]
27. Mathison S. Why triangulate? Educational Researcher 1988;17(2):13-17.
28. Wallerstein N. Powerlessness, empowerment, and health: implications for health promotion programs. Am J Health Promot 1992;6(3):197-205. [Medline: 94907907]
29. Cragg CE. Distance learning through computer conferences. Nurse Educ 1994;19(2):10-14. [Medline: 95157846]
30. Wong D, Packard BGM, Pugh K. The opposite of control: a Deweyan perspective on intrinsic motivation in "after 3 technology programs". Computers in Human Behavior 2000;16(3):313-338. [doi: 10.1016/S0747-5632(00)00009-1]
31. Rees S. Empowerment of youth. In: Empowerment in social work practice: a sourcebook. Toronto, Ontario: Brooks/Cole Publishing Company; 1998:130-145.
32. Schwab M. Sharing power: participatory public health research with California teens. Social Justice 1997;24(3):11-32.
33. Delgado M, Barton K. Murals in Latino communities: social indicators of community strengths. Social Work 1998;43(4):346-350.
34. Chu J. Navigating the media environment: how youth claim a place through zines. Social Justice 1997;24(3):71-85.
35. Saunders D. Invisible youth reappear! A review of two youth produced videos. Social Justice 1997;24(3):100-105.
36. Morrow V. 'Dirty looks' and 'trampy places' in young people's accounts of community and neighbourhood: Implications for health inequalities. Critical Public Health 2000;10(2):141-152.
37. Wang C, Burris MA. Empowerment through photo novella: portraits of participation. Health Educ Q 1994;21(2):171-186. [Medline: 94292382]
38. Wang C, Burris MA. Photovoice: concept, methodology, and use for participatory needs assessment. Health Educ Behav 1997 Jun;24(3):369-387. [Medline: 97302733]
39. Strack RW, Magill C, Mcdonagh K. Engaging youth through photovoice. Health Promot Pract 2004 Jan;5(1):49-58. [Medline: 14965435] [doi: 10.1177/1524839903258015]
40. Bamberger J. Action knowledge and symbolic knowledge: the computer as mediator. In: High technology and low-income communities: prospects for the positive use of advanced information technology. Cambridge, Mass: MIT Press; 1999:235-262.
41. 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] [Medline: 14713660] [doi: 10.2196/jmir.5.4.e32]

submitted 07.07.05; peer-reviewed by S Flicker; comments to author 20.07.05; revised version received 09.08.05; accepted 10.08.05; published 04.10.05.

Please cite as:
Valaitis RK
Computers and the Internet: Tools for Youth Empowerment
J Med Internet Res 2005;7(5):e51
URL: http://www.jmir.org/2005/5/e51/
doi:10.2196/jmir.7.5.e51
PMID:16403715

Viewpoint

A Historical Overview of Health Disparities and the Potential of eHealth Solutions

Michael C Gibbons, MD, MPH

Corresponding Author:
Michael C Gibbons, MD, MPH
Johns Hopkins Urban Health Institute
Johns Hopkins Medical Institutions
111 Market Place, Suite 850
Baltimore, MD 21202
Phone: +1 410 895 1104
Fax: +1 410 895 1111
Email: mgibbons@jhsph.edu

Abstract

Over the past decade, a rapidly expanding body of literature has demonstrated the existence of disparities in health and health care. While consensus has not emerged regarding the causes of disparities, they are generally thought to be related to sociocultural, behavioral, economic, environmental, biologic, or societal factors. To effectively address disparities, several authorities have suggested the need for greater information technology research and investments. eHealth researchers may be able to make significant contributions in this area through research and its applications. This paper begins with a historical overview of health disparities in the United States and Europe. It then discusses the role that the Internet, and access to the Internet, may play in the genesis of health disparities. Finally, this paper closes with a discussion of the potential benefits of eHealth applications and the possible contributions of the field to overcoming disparities in health and health care.

(J Med Internet Res 2005;7(5):e50) doi:10.2196/jmir.7.5.e50

KEYWORDS

access to information; socioeconomic factors; Internet; disparities; eHealth; minorities

International Origins

Over the past decade, a rapidly expanding body of scientific evidence has been put forth documenting differences in health status among US racial and ethnic groups. Evidence has also mounted suggesting that these differences may be related to both medical and nonmedical determinants. Internationally, however, neither the evidence nor the realization of a link between nonmedical sociobehavioral factors and health outcomes is new. The earliest reported observation of a hypothesized association between socioenvironmental risk factors and health outcomes occurred in Italy over three centuries ago when Bernardino Ramazzini detailed an unusually high frequency of breast cancer in Catholic nuns [1]. Not long thereafter, in 1775, British surgeon Sir Percival Pott reported a cluster of scrotal cancer cases among British chimney sweeps [1].

By the mid 19th century, large-scale epidemiologic evidence began to corroborate these early observations. In 1840, Edwin Chadwick, British civil servant and statistician, demonstrated mortality differentials between the social classes living in Liverpool, England. Chadwick asserted that these differences were likely due to poverty and lifestyle factors common to the poorer working classes [2]. German physician Rudolph Virchow went a step further when, in 1849, he asserted that, because diseases of the populace are traceable to defects in society, the focus of medicine should shift from changing the individual to that of changing the society [3]. Finally, in France, French physician Louis Villerme recommended improving school and working conditions as social interventions that would reduce class differences in mortality [3]. Thus, in Europe, by the beginning of the 20th century, the existence of class variations in morbidity and mortality were clearly evident in the scientific literature [2].

Throughout the 20th century, the study of social class differences in health status continued across Europe, especially in Britain were epidemiologists began using decennial census data to evaluate national mortality trends. The insights gained from these analyses enabled them to construct an occupational social class grading system that correlated inversely with infant mortality. It also was the basis of the claim made by the Registrar General of Britain that at least 40% of British infant mortality was entirely preventable if the social conditions of poor infants could be elevated to that of upper-class infants [2].

Two British researchers, Titmuss and Logan, evaluated regional class-based mortality trends and documented that the disparity in infant mortality rates between upper- and lower-class infants continued to increase from 1910 to 1950 [2]. This data, along with the Depression and World War II, encouraged the British government, in 1942, to respond by instituting the welfare state and promoting several policy initiatives designed to address the "five giants of Want, Disease, Ignorance, Squalor and Idleness" [2,4]. Despite this government investment, however, problems attributable to social inequalities and inadequate access to health care persisted. In fact, by the mid 1970s, some 30 years later, the evidence seemed to indicate that the problems were still increasing and that the health of British citizens was slipping behind that of other industrialized nations [4]. Thus, in 1977, the British government formed the Research Working Group on Inequalities in Health and selected Sir Douglas Black as its chair. The committee's report, issued three years later in 1980, became known as the Black Report, and it represents the first attempt by a national government to systematically study, understand, and explain health inequalities [4]. In summary, the health improvement recommendations of the report emphasized the need to improve the physical and the social environment in which the poor and lower classes lived [4].

Domestic Recognition

Across the Atlantic in the United States, scientific evidence from several lines of inquiry examining outcomes and patterns of health care delivered to defined populations began to converge and suggest the importance of the socioenvironment in determining health outcomes. Researchers using small area analysis and geographic information systems analytic techniques demonstrated that a significant amount of nonrandom practice variability existed between clinical practices in different geographic locales, despite treating clinically similar patients [5,6]. As public awareness grew, the US government became involved. In 1984, the US Department of Health and Human Services released a report on the health of the nation, entitled "Health, United States, 1983" [7]. The report documented that, while the overall health of the nation showed significant progress, major disparities existed in "the burden of death and illness experienced by blacks and other minority Americans as compared with the nation's population as a whole" [7].

In response to the disparities identified in the report, the secretary of the Department of Health and Human Services established a task force on black and minority health—the first time that the US government formed a group of experts to conduct a comprehensive study of minority health problems. In 1985, release of the "Report of the Secretary's Task Force on Black and Minority Health" significantly raised awareness of the disparate health of the country's minority groups as compared to the white majority population [8].

Large epidemiologic studies like the Harvard Medical Practice Study emerged, documenting that a significant portion of practice variability could be classified as substandard care and that there was a correlation between substandard care and health care centers treating substantial numbers of poor and minority patients [9-11].

The emerging problems of differential outcomes and health status were not limited however to minorities and the poor. The Whitehall studies of a large cohort of British civil servants had convincingly demonstrated that a social class–based health gradient existed even among the well educated and employed [12]. Additionally, it became increasingly recognized that certain community and societal level factors, including stress [13,14], early life experiences [15], social capital [16], and income inequality [17,18] seemed to exert significant effects on health and disease outcomes independent of personal behavior [3,19].

Soon, major philanthropic and advocacy organizations, including The Commonwealth Fund, The Kaiser Family Commission, the Kellogg Foundation, the Robert Wood Johnson Foundation, and The California Endowment began major initiatives designed to address issues related to disparities and health care quality [20].

By the late 1990s, the scientific evidence seemed to indicate that issues of disparity, practice variation, substandard care, and socioenvironmental determinants of health may all be related to the quality of health care experienced by patients. Friscella published his paper entitled "Inequality in Quality," in which he called attention to issues of health care quality and health care disparities as related issues of health care organizational capacity. He further contended that national efforts to eliminate racial and ethnic disparities in health care and national health care quality improvement initiatives represented two inseparable components of providing high-quality health care for all citizens [21].

Synthesizing the Scientific Evidence on Health Disparities

As the domestic evidence for population differences continued to accumulate, definitions of disparities were nonstandardized, and racial categorizations became increasingly criticized as being imprecise and biologically meaningless [22,23]. While multiple definitions are still in current use, disparities are generally held to be population differences in (1) environmental exposures, (2) health care access, utilization, or quality, (3) health status, or (4) health outcomes [24]. As alluded to above, within the US health care system these differences have most convincingly been demonstrated across racial and ethnic lines (whites vs minorities); however, disparities based on other categorizations have also been described, including geography (urban vs rural) [25], gender (male vs female) [26,27], socioeconomic status (poor vs nonpoor) [28,29], and age (nonelderly vs elderly) [30].

Health disparities are generally thought to be related to the health care system and other social factors. Several lines of investigation examining the socioenvironment and the clinical encounter give evidence of differences in the quality of care received by many racial and ethnic minorities. While these factors have been described as "causes" and are likely to be important in the genesis of disparities, scientifically validated evidence of definitive causal pathways and the underlying biologic mechanisms is largely lacking [31].

To help bring clarity to these issues, the Institute of Medicine (IOM) released the first of several reports highlighting and summarizing the scientific evidence concerning issues of differential health status, culture, behavior, communication, substandard care/medical errors, and health care quality [32-37]. The work of the IOM on disparity issues culminated with the 2003 release of a report entitled “Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care” [38]. In this report, the IOM Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care was charged with assessing the extent and potential sources of racial and ethnic disparities in health care that are not otherwise attributable to access to care, ability to pay, or insurance coverage. The committee was also to provide recommendations regarding potential interventions to eliminate health care disparities [38]. The committee found that, within the United States, even among individuals with access to care, significant racial and ethnic disparities indeed existed and were related to historic and contemporary social and economic inequality, discrimination, and a fragmented US system of health care [38]. While the release of this report has engendered significant public, media, and academic interest, likely ensuring that efforts to understand and eliminate disparities will continue at least into the foreseeable future, the magnitude and intransigence of the problem, the complexity of its causal pathways, and its resistance to intervention efforts are only beginning to be realized [32].

Digital Disparities

Since the mid-1990s when the World Wide Web became a powerful part of America’s communications and information culture, there has been great concern that the nation’s racial minorities would be further disadvantaged because Internet access was not spreading as quickly in the African-American community as it was in the white community. Former Assistant Secretary of Commerce Larry Irving said the following in his introduction to “Falling Through the Net,” the 1999 Department of Commerce Study on the digital divide (the divide between those with access to new information technologies and those without): “[The digital divide] is now one of America’s leading economic and civil rights issues” [39]. This report found that, although, overall, the number of Americans connected to the nation’s information infrastructure was soaring, a digital divide existed between whites and African-Americans in terms of their access to the Internet, and that, in many cases, the divide was *widening* over time. A follow-up study revealed a persistent but substantially narrowed gap, with large increases in computer ownership and Internet use across most major demographic populations [40]. The most recent survey, released in 2003, indicated a significant slowing in the growth of the number of Internet users since late 2001 [41]. Overall, 42% of surveyed individuals did not use the Internet, and significant utilization differences remained according to race, education, income, and geography (urban vs rural) [41]. Generally, whites are more connected than African-Americans and Hispanics. Even at equivalent levels of income, African-Americans are less likely to be online than whites or Hispanics. In fact, over the period of this study (mid-2000 to mid-2002), the composition of the non-Internet user group did not change substantially [41]. Interestingly, 56% of nonusers said they did not ever plan to go online and cited the cost of computers or Internet access, fear of fraud, credit card theft, or pornography as the major reasons for avoiding Internet use [41].

Recently, there has been a significant increase in the public availability of computers and Internet access at schools, public libraries, and workplaces [42]. Thus, conclusions regarding the extent of a digital disparity based on data considering only home-based access may be limited. Despite this reality, Internet availability in the home is accepted as an important indicator of equitable access among population groups [42]. In addition, access in public settings may be problematic because of computer monitoring in the workplace, privacy and confidentiality concerns, and the facilities’ hours of operation. Because of the potentially sensitive nature of health-related uses of the Internet, access at home is thought to be essential [42].

Several studies have shown that access to the Internet correlates with income level and educational attainment [39-42]. As with racial and ethnic differences, Internet utilization is increasing in all income brackets. The largest increases are seen in the higher income categories. All things considered, household incomes above US$50000 are positively associated with Internet utilization [41]. Beyond socioeconomic issues, some researchers have speculated that African-Americans have had less access to the Internet because they participate to a greater degree in entertainment-oriented technologies like television, rather than in information technologies. They argue that relatively high proportions of African-Americans use radio and television, but a relatively low proportion read newspapers [40]. As suggested above, the primary reasons why some groups have less access to information technology and resources are related to geography, literacy, disability, local infrastructure requirements, and cultural differences [43], some of which are not easily overcome simply by increasing personal computer ownership. Even if equity in personal computer and Internet access were achieved, emerging evidence suggests that online habits may vary by race and ethnicity. For example, online African-Americans are more likely than online whites to have (1) searched for information about major life issues such as researching new jobs and finding places to live, (2) used entertainment online, (3) used the Internet to obtain health information, and (4) searched for religious or spiritual information [40]. On the other hand, African-Americans with access to the Internet do not go online as often on a typical day as whites do, and they do not participate on a daily basis in most Web activities at the same level as online whites [40].

As information technology plays an ever-increasing role in Americans’ economic and social lives, the potential health implications of these findings need to be more clearly evaluated because the prospect that some people will be left behind in the information age may have serious repercussions [44]. Persistent digital disparities in access or utilization could leave some groups less able to take advantage of cutting edge innovations in population health technologies that enhance disease surveillance, environmental monitoring, food safety, emergency planning, disaster management, and geographic information systems–based tracking of environmental hazards [45].

The Role of Information Technology in Overcoming Health Disparities

One major domain of eHealth focuses on improving health communication through the use of technology. This notion of enhancing communication and understanding is a fundamental component of addressing health disparities. Among other things, the recommendations of the Institute of Medicine report call for initiatives designed to enhance patient-provider communication, trust, and cultural appropriateness of delivered care [38]. Similar goals are the basis for the Healthy People 2010 objective to increase the number of individuals with Internet access in the home. Providers, health care organizations, and public health agencies are increasingly using the Internet as a main source of information dissemination and communication [42]. This need for innovative improvements in communication should represent a significant opportunity for eHealth technologies, researchers, and interventionists, with many important implications for overcoming disparities in health and health care. Given that eHealth is currently understood as attempting to facilitate the utilization of information technologies, the Internet, and communication technology in order to facilitate behavior change, improve health care, and enhance health outcomes [46], eHealth researchers may become the catalysts needed to spur the development of transdisciplinary interventions to effectively address disparities in health and health care.

Recent advances in the computer sciences and information technology fields have spawned several methodological advances in the biological and molecular sciences (eg, DNA chip technology and microarray analysis), enabled quantum leaps in molecular and submolecular medicine, and catalyzed the emergence of whole new fields of study such as proteomics, phenomics, nutrigenomics, and pharmacogenetics. Perhaps, in like manner, with the emergence of eHealth, the behavioral and population sciences may be on the verge of a similar information technology–based scientific revolution. New eHealth solutions may soon permit the real-time integrative utilization of vast amounts of behavioral-, biological-, and community-level information in ways not previously possible. Behavioral algorithms and decision support tools for scientists could facilitate the analysis and interpretation of population level data to enable the development of "community (population) arrays" or community-wide risk profiles, which in turn could form the foundation of a new "populomics." This population-level risk characterization could potentially go beyond the limitations of typical geographic analyses and yield insights distinctly different from risk stratification based on current methodologies. Generically, these emerging technologies have been termed population health technologies and are believed to offer significant promise [45].

These assertions are not based on mere speculation. Encouraging early evidence suggests that multimedia health communication and behavior change efforts that include the use of computers and other eHealth technologies can improve health outcomes [47]. Among other factors, the evidence suggests that applications that are tailored to the individual, participatory, personally relevant, and contextually situated will be more likely to promote behavior change [47]. On the other hand, the Internet has been implicated in the causation or persistence of disparities because of the relative lack of access of some groups and because of its current inability to deliver content that is dynamically tailored to meet the cultural, language, or literacy needs of the individual user [48]. This may be particularly true of eHealth applications that are "Internet-enabled," requiring access to the Internet to provide the interventional content. It is conceivably possible, however, to conceptually divide eHealth applications into at least two genres: those that rely on the Internet to deliver the interventional content directly to patients (Internet-enabled), and those that only employ the Internet to facilitate transfer and utilization of data for or about content that is delivered to patients by an alternate approach. The content or interventions themselves can actually function without the Internet, but when used in the context of the Internet, they are potentially much more efficacious and far-reaching. These types of technologies could be termed "Internet-enhanced" eHealth solutions. Here the Internet would facilitate the transfer of data and information, but the tailored content could be delivered by trusted people from the users own culture or community. The actual intervention could also be administered to patients by print or multimedia applications. Thus, in terms of overcoming health disparities, issues of guaranteeing Internet access for every individual may prove to be less important than attempting to address health disparities via interventions and methodologies that lack cultural relevance. Indeed, those interventions and strategies that integrate behavioral interventions with emerging information technologies will likely be the interventions capable of cost-effectively enabling mass customization, interactivity, and convenience. Ultimately, though, the health disparities challenge for eHealth researchers remains to harness the technical capabilities of emerging information technologies in ways that support the social and cultural realities in which people work and live [47], while enhancing our ability to address the health needs of every patient [49].

Acknowledgments

The author wishes to acknowledge Earl Fox for his support of the author's work in this area.

Conflicts of Interest

None declared.

References

1. ; Institute of Medicine (U.S.). Cancer and the Environment; Gene-Environment Interaction. Washington, DC: National Academy Press; 2002.
2. Macintyre S. The Black Report and beyond: what are the issues? Soc Sci Med 1997 Mar;44(6):723-745. [Medline: 97235439] [doi: 10.1016/S0277-9536(96)00183-9]
3. Amick BC, Levine S, Tarlov AR, Walsh DC. Society and Health. New York: Oxford University Press; 1995.
4. Acheson D. Independent Inquiry into Inequalities in Health Report. London, UK: The Stationery Office; 1998.
5. Mcpherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. N Engl J Med 1982 Nov 18;307(21):1310-1314. [Medline: 83036821]
6. Barnes BA, O'brien E, Comstock C, D'arpa DG, Donahue CL. Report on variation in rates of utilization of surgical services in the Commonwealth of Massachusetts. JAMA 1985 Jul 19;254(3):371-375. [Medline: 85237817] [doi: 10.1001/jama.254.3.371]
7. NCHS: Health, United States, 1983; and Prevention Profile. [PHS] 84-1232. Washington, DC: US Government Printing Office; 1983.
8. Mayberry RM, Mili F, Ofili E. Racial and ethnic differences in access to medical care. Med Care Res Rev 2000;57 Suppl 1(1):108-145. [Medline: 20543653] [doi: 10.1177/107755800773743628]
9. Brennan TA, Leape LL, Laird NM, Localio AR, Hiatt HH. Incidence of adverse events and negligent care in hospitalized patients. Trans Assoc Am Physicians 1990;103:137-144. [Medline: 92087352]
10. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991 Feb 7;324(6):370-376. [Medline: 91101671]
11. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med 1991 Feb 7;324(6):377-384. [Medline: 91101672]
12. Marmot MG, Rose G, Shipley M, Hamilton PJ. Employment grade and coronary heart disease in British civil servants. J Epidemiol Community Health 1978 Dec;32(4):244-249. [Medline: 79130934]
13. Marmot MG. Does stress cause heart attacks? Postgrad Med J 1986 Jul;62(729):683-686. [Medline: 86313379]
14. Sapolsky RM, Mott GE. Social subordinance in wild baboons is associated with suppressed high density lipoprotein-cholesterol concentrations: the possible role of chronic social stress. Endocrinology 1987 Nov;121(5):1605-1610. [Medline: 88029227]
15. Tager IB, Weiss ST, Muñoz A, Rosner B, Speizer FE. Longitudinal study of the effects of maternal smoking on pulmonary function in children. N Engl J Med 1983 Sep 22;309(12):699-703. [Medline: 83297521]
16. Coleman JS. Social capital in the creation of human capital. Am J Sociol 1988;94(suppl):S95-S120.
17. Wilkinson RG. National mortality rates: the impact of inequality? Am J Public Health 1992 Aug;82(8):1082-1084. [Medline: 92343936]
18. Wilkinson RG. Income distribution and life expectancy. BMJ 1992 Jan 18;304(6820):165-168. [Medline: 92145215]
19. Wilkinson RG. Unhealthy Societies: The Afflictions of Inequalities. New York: Routledge; 1996.
20. ; Compendium of Cultural Competence Initiatives in Health Care. Washington, DC: The Henry J. Kaiser Family Foundation; 2003.
21. Fiscella K, Franks P, Gold MR, Clancy CM. Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. JAMA 2000 May 17;283(19):2579-2584 [FREE Full text] [Medline: 20276026] [doi: 10.1001/jama.283.19.2579]
22. Anderson NB, Nickerson KJ. Genes, race, and psychology in the genome era: an introduction. Am Psychol 2005 Jan;60(1):5-8. [Medline: 15641916] [doi: 10.1037/0003-066X.60.1.5]
23. Smedley A, Smedley BD. Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. Am Psychol 2005 Jan;60(1):16-26. [Medline: 15641918] [doi: 10.1037/0003-066X.60.1.16]
24. Carter-pokras O, Baquet C. What is a "health disparity"? Public Health Rep 2002;117(5):426-434. [Medline: 22389379] [doi: 10.1093/phr/117.5.426]
25. Handbook of Urban Health: Populations, Methods and Practice. New York: New York Academy of Medicine; 2005.
26. Quinn TC, Overbaugh J. HIV/AIDS in women: an expanding epidemic. Science 2005 Jun 10;308(5728):1582-1583. [Medline: 15947174] [doi: 10.1126/science.1112489]
27. Mcgrath BB, Puzan E. Gender disparities in health: attending to the particulars. Nurs Clin North Am 2004 Mar;39(1):37-51. [Medline: 15062726] [doi: 10.1016/j.cnur.2003.11.003]
28. Federico MJ, Liu AH. Overcoming childhood asthma disparities of the inner-city poor. Pediatr Clin North Am 2003 Jun;50(3):655-75, vii. [Medline: 22759288] [doi: 10.1016/S0031-3955(03)00045-2]
29. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. Ann N Y Acad Sci 1999;896:3-15. [Medline: 20146465]
30. Pyle MA, Stoller EP. Oral health disparities among the elderly: interdisciplinary challenges for the future. J Dent Educ 2003 Dec;67(12):1327-1336 [FREE Full text] [Medline: 23094267]

31. Evans RG, Barer ML, Marmor TR. Why Are Some People Healthy and Others Not? The Determinants of Population Health. New York: Aldine De Gruyter; 1994.
32. Speaking of Health: Assessing Health Communication Strategies for Diverse Populations. Washington, DC: National Academy Press; 2002.
33. Haynes M, Smedley BD. The Unequal Burden of Cancer. Washington, DC: National Academy Press; 1999.
34. Envisioning the National Health Care Quality Report. Washington, DC: National Academy Press; 2001.
35. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
36. Priority Areas for National Action: Transforming Health Care Quality. Washington, DC: National Academy Press; 2003.
37. Kohn L, Corrigan J, Donaldson M. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press; 2000.
38. Smedley BD, Stith AY, Nelson AR. Unequal Treatment; Confronting Racial and Ethnic Disparities in Healthcare. Washington, DC: National Academies Press; 2003.
39. ; National Telecommunications and Information Administration and US Department of Commerce. Falling Through the Net: Defining the Digital Divide. Washington, DC: Government Printing Office; 1999.
40. Spooner T, Rainie L. African-Americans and the Internet. Washington, DC: Pew Internet & American Life Project; 2005.
41. Lenhart A, Horrigan J, Allen K, Boyce A, Madden M, O'Grady E. The Ever-Shifting Internet Population: A New Look at Internet Access and the Digital Divide. Washington, DC: Pew Charitable Trusts; 2003.
42. ; Office of Disease Prevention and Health Promotion. Healthy People 2010. 2nd edition. Washington, DC: Government Printing Office; 2002.
43. Eng TR, Maxfield A, Patrick K, Deering MJ, Ratzan SC, Gustafson DH. Access to health information and support: a public highway or a private road? JAMA 1998 Oct 21;280(15):1371-1375 [FREE Full text] [Medline: 99008453] [doi: 10.1001/jama.280.15.1371]
44. ; US Department of Commerce. Digital Divide Summit, Washington, DC; December 9, 1999. URL: http://www.ntia.doc.gov/ntiahome/digitaldivide/summit/ [accessed 2005 Oct 5]
45. Eng TR. Population health technologies: emerging innovations for the health of the public. Am J Prev Med 2004 Apr;26(3):237-242. [Medline: 15026105] [doi: 10.1016/j.amepre.2003.12.004]
46. Eng TR. eHealth research and evaluation: challenges and opportunities. J Health Commun 2002;7(4):267-272. [Medline: 22243100] [doi: 10.1080/10810730290001747]
47. Neuhauser L, Kreps GL. Rethinking communication in the e-health era. J of Health Psychol 2003;8(1):7-22. [doi: 10.1177/1359105303008001426]
48. Cashen MS, Dykes P, Gerber B. eHealth technology and Internet resources: barriers for vulnerable populations. J Cardiovasc Nurs 2004;19(3):209-14; quiz 215-6. [Medline: 23363773]
49. Safran C. The collaborative edge: patient empowerment for vulnerable populations. Int J Med Inform 2003 Mar;69(2-3):185-190. [Medline: 22695138] [doi: 10.1016/S1386-5056(02)00130-2]

---

submitted 20.07.05; peer-reviewed by T Eng; comments to author 15.09.05; revised version received 19.09.05; accepted 20.09.05; published 04.10.05.

Please cite as:
Gibbons MC
A Historical Overview of Health Disparities and the Potential of eHealth Solutions
J Med Internet Res 2005;7(5):e50
URL: http://www.jmir.org/2005/5/e50/
doi:10.2196/jmir.7.5.e50
PMID:16403714

Viewpoint

Mobile eHealth Interventions for Obesity: A Timely Opportunity to Leverage Convergence Trends

James T Tufano[1], MHA; Bryant T Karras[2], MD

[1]Department of Medical Education and Biomedical Informatics, University of Washington School of Medicine, Seattle, WA, USA
[2]Department of Health Services, University of Washington, School of Public Health and Community Medicine, Seattle WA, USA

Corresponding Author:
James T Tufano, MHA
University of Washington School of Medicine
Biomedical & Health Informatics
1959 NE Pacific Street, HSC E-312
Box 357240
Seattle, Washington
98195-7240
Phone: +1 206 719 0332
Fax: +1206 338-2527
Email: jtufano@u.washington.edu

Abstract

Obesity is often cited as the most prevalent chronic health condition and highest priority public health problem in the United States. There is a limited but growing body of evidence suggesting that mobile eHealth behavioral interventions, if properly designed, may be effective in promoting and sustaining successful weight loss and weight maintenance behavior changes. This paper reviews the current literature on the successes and failures of public health, provider-administered, and self-managed behavioral health interventions for weight loss. The prevailing theories of health behavior change are discussed from the perspective of how this knowledge can serve as an evidence base to inform the design of mobile eHealth weight loss interventions. Tailored informational interventions, which, in recent years, have proven to be the most effective form of conventional health behavior intervention for weight loss, are discussed. Lessons learned from the success of conventional tailored informational interventions and the early successes of desktop computer–assisted self-help weight management interventions are presented, as are design principles suggested by Social Cognitive Theory and the Social Marketing Model. Relevant computing and communications technology convergence trends are also discussed. The recent trends in rapid advancement, convergence, and public adoption of Web-enabled cellular telephone and wireless personal digital assistant (PDA) devices provide timely opportunities to deliver the mass customization capabilities, reach, and interactivity required for the development, administration, and adoption of effective population-level eHealth tailored informational interventions for obesity.

(J Med Internet Res 2005;7(5):e58) doi:10.2196/jmir.7.5.e58

KEYWORDS

eHealth; obesity; intervention; mobile computing; cellular telephone; weight loss; health behavior; health communication; behavior modification; consumer health informatics

Introduction

Obesity is often cited as the most prevalent chronic health condition and highest priority public health problem in the United States, and it has also gained recognition as a global public health concern. The increasing prevalence of obesity-related diseases such as type 2 diabetes and cardiovascular disease has lead the medical and public health communities to declare an obesity epidemic and to conclude that more effective population-oriented interventions are needed [1].

There is a limited but growing body of evidence suggesting that eHealth interventions exhibit potential to address a variety of chronic illnesses that can be effectively treated via behavior modification [2]. Also, it appears that there is a significant and growing opportunity for eHealth obesity intervention designers to leverage the widespread public adoption of rapidly converging information and communication technologies—most notably the World Wide Web, wireless PDAs, and cellular telephones.

The Failures of Conventional Behavioral Interventions

An Institute of Medicine report published in 2000 concluded that “behavioral and social interventions offer great promise to reduce disease morbidity and mortality, but as yet their potential to improve the public’s health has been relatively poorly tapped” [3].

This claim is further validated by trends exhibited in US public health data collected via the Behavioral Risk Factor Surveillance System (BRFSS) program of the Centers for Disease Control and Prevention (CDC). The BRFSS is a cross-sectional telephone survey program administered by the CDC in conjunction with state health departments. Data collected from each state are pooled to produce nationally representative estimates of the prevalence of specific health behaviors that affect risk for one or more of the ten leading causes of death in the United States. According to the results of the 2000 BRFSS, which are based on responses from 184450 participants in 50 states, the prevalence of obesity was 19.8% among US adults, representing a 61% increase compared to the 1991 BRFSS data. Also, 56.4% of 2000 BRFSS survey respondents were overweight (body mass index ≥ 25 kg/m^2) compared to 45% in 1991 [4].

The 2000 BRFSS data also showed that 38.5% of US adults reported that they were attempting to lose weight, compared to 36.6% in 1991; 35.9% were trying to maintain weight, compared to 34.4% in 1991; and 25.6% were doing neither, compared to 29.0% in 1991 [5]. This would indicate that Americans are collectively putting more effort into losing or maintaining weight, or at least *perceive* that they are putting more effort into weight reduction or maintenance, yet as a whole are exhibiting decreasing levels of success.

Provider-Delivered Health Behavior Interventions

Behavior changes related to dietary intake and physical activity might be effective in preventing and treating obesity. Studies have shown that overweight and obese people are more likely to attempt weight loss by adopting lifestyle changes if they are advised to do so by a health care professional [6], yet the 2000 BFRSS data showed that only 42.8% of the obese respondents who had received a routine medical checkup in the past year had been advised to lose weight by their providers [4].

An explanation is needed for why 57.2% of American clinicians are failing to advise their obese patients to lose weight during routine medical checkups. This clinical window of opportunity for engaging patients in discussion about obesity and their health is being missed. Various explanations are offered in the literature, including lack of provider access to high-quality information about effective patient intervention strategies, lack of access to appropriate support services, and a lack of provider motivation to work with obese and overweight patients [7]. A compelling explanation is that the provider community is skeptical of the efficacy of conventional provider-delivered health behavior change interventions—an explanation supported by a systematic review of the existing literature on provider-delivered and health care organization–delivered behavioral health interventions that was conducted in 2002 [8]. In this review, Harvey et al concluded that relatively few obesity interventions have been evaluated rigorously, and “at present, there are few solid leads about improving obesity management.” The recent systematic reviews of the literature on dietary and physical activity counseling, used to inform the recommendations of the US Preventive Services Task Force, have yielded similar results [9].

Conventional Population-Oriented Public Health Interventions

Health communication campaigns provide the traditional means by which public health organizations have sought to promote health behavior change. Conventional health communication campaigns involve dissemination of messages from experts to the public through mass media channels, with the intent of motivating the public to adopt specific behaviors that have been proven to reduce the risk of disease. However, there is evidence that traditional health communication interventions exhibit a high rate of failure to promote behavior change [10] and that traditional methods of health communication are particularly ineffective at addressing weight-related health behavior change [11].

Two examples of the failure of public health communication interventions aimed at obesity are the California Five-A-Day for Better Health Campaign and the former Health Education Authority (HEA) of England’s Active for Life campaign. The California Five-A-Day Campaign was an intensive five-year statewide intervention aimed at promoting dietary behavior change by increasing the consumption of fresh fruits and vegetables. Program evaluation showed that knowledge and understanding of the importance of eating more fruits and vegetables increased substantially among Californians as a result, but this did not result in any measurable behavior change. There was no increase in fruit and vegetable consumption in any population group, and among Hispanic adults, consumption actually dropped by 18% [12].

The largest-scale public health intervention attempted in Europe during the latter part of the 1990s was the Active for Life campaign, a three-year intervention that aimed to promote moderate-intensity physical activity as a part of everyday life. The multi-faceted intervention ran from 1996 to 1999 and included a variety of mass media communication components, as well as a program of support to health care and public health professionals who worked to develop and promote localized community-based physical activity programs. Evaluation showed that, after three years, there was no evidence that the campaign improved physical activity, either overall or in any subgroup [13].

These studies highlight the need for the public health practitioners and policy makers to explore alternative methods of promoting health behavior change in the fight against the obesity epidemic.

Prevailing Theories of Health Behavior Change and Intervention Models

Prevailing theories of health behavior change suggest several root causes for the failure of conventional obesity interventions. These theories and their supporting empirics also imply key intervention design features that may increase their likelihood of promoting and sustaining the desired behavior changes.

Several established theories and models of health-related behavior have informed the design of successful health behavior change communication interventions. These theories and models are drawn mostly from the fields of psychology, sociology, communication, and medicine, and they draw heavily on research in persuasion, social marketing, and relational communication [14]. The most prevalent health behavior theories cited by intervention designers include the Health Belief Model, the Theory of Reasoned Action, the Transtheoretical (or Stages of Change) Model, Learning and Conditioning Theory, Decision-Making Theory, and the Diffusion of Innovations Model. These various theories and models share the common objective of attempting to explain and predict individual health-related behaviors, and they have formed the basis for the design, deployment, and evaluation of the majority of health behavior interventions at the individual, organizational/group, and population levels [15].

Of the dominant classic models, the Transtheoretical Model and the Health Belief Model have been drawn on most heavily in the design of tailored informational interventions, which, in recent years, have generally proven to be the most effective model of health behavior change intervention for a variety of diseases [16,17]. In their 2002 systematic review, Ryan and Lauver suggest that tailored informational interventions exhibit four key defining characteristics that differentiate them from other intervention types [16]. According to their definition, tailored informational interventions include (1) an assessment of key characteristics of each targeted person; (2) small units of content prepared to match each of these key characteristics, often stored in a message library; (3) a decision algorithm that provides the matching logic; and (4) a designated information delivery channel or mechanism (eg, print, email, telephone) [16].

Tailored informational interventions are based on the premise that the design of a behavioral communication intervention must be modified or “tailored” to accommodate varying states of readiness for behavior change among the targeted recipients. Under this design paradigm, delivering the right message at the right time often requires both multiple forms of communication delivery and variation in the content of messages delivered. They are typically used to support individual provider-to-patient intervention strategies and small-scale health promotion programs, often involve the use of computer-generated content construction, and are gaining broad acceptance as one of the only forms of health communication intervention that successfully yields behavior change [16-18].

The Transtheoretical (or Stages of Change) Model is based on the premise that people move across a continuum of readiness to change, moving from “pre-contemplation” (e.g., “I suppose that I *should* try to eat healthier and shed some pounds at some point…”) to “action” (e.g., “Today I am purging my fridge and cupboard of all junk food…”) to “maintenance” of a behavior change (e.g., “I have been planning my meals and buying only healthy foods for the past eight months…”)of a behavior change. Tailored informational interventions leverage the Transtheoretical Model by designing incorporating communication mechanisms and content specific to a targeted individual’s state of readiness for change as indicated by his or her position on this continuum [19].

The Health Belief Model estimates a person’s likelihood of adopting a healthy behavior based on his or her perceptions of the risk of becoming ill, anticipated benefits to be gained, and the barriers to adoption of the behavior change. Tailored informational interventions leverage this model primarily in the design of the content of the messages delivered [20].

These and most of the other classic theories and models of health behavior change emphasize the individual as the decision maker, which drives intervention designers to focus on delivering expert-driven, risk-based information to targeted at-risk individuals. Targeted individuals are expected to use this information to make rational decisions about discrete behavior change and then act on these rational decisions by changing their behaviors. Emmons cites this as a fundamental weakness and argues that improving the effectiveness of health behavior interventions will require models of behavior that account for how mediating variables of behavior change are influenced by sociocultural dynamics [21]. Growing agreement within the public health community has lead to exploration of interventional approaches that leverage Social Marketing Theory, Social Cognitive Theory, and other behavioral science and social epidemiology theories that place a greater emphasis on the social, institutional, and cultural contexts that impact an individual’s behaviors [22].

Social Cognitive Theory shares some attributes of both the Transtheoretical Model and the Health Belief Model, but it offers some unique and compelling contributions in its emphasis on the role of personal empowerment in behavior change. This theory suggests that individuals’ sense of “self-efficacy” or agency about a behavior and their perceived ability to cope with and control situations are core determinants of behavior change [23].

Successful Health Behavior Interventions for Obesity

In assessing the literature on behavioral obesity interventions, careful consideration must be given to the magnitude and duration of the primary outcomes. Many “successful” obesity interventions produce significant weight loss that is difficult to maintain or result in weight losses that are too small to yield a substantial health gain. There is a paucity of compelling evidence that any behavioral obesity interventions consistently yield both clinically significant and sustainable weight loss. Given this caveat, several studies nonetheless provide encouraging insights into intervention characteristics that, if considered in the design of eHealth obesity interventions, may improve their likelihood of success.

In their systematic review of studies of health behavior interventions aimed at increasing physical activity published from 1983 to 1997, Marcus et al identified a total of 28 qualifying papers [11]. Seven described studies of traditional mass media health communication campaigns conducted at the state or national level. None of these interventions were found to have affected behavior change. However, in this same review, the majority of the other 21 studies were found to have had some positive impact on exercise-related behaviors. They were all smaller-scale interventions using various forms and combinations of print and/or telephone media. The majority of these interventions exhibited design characteristics based heavily on the Transtheoretical Model and/or Social Cognitive Theory, and most could be classified as tailored informational interventions. Three of these studies reported that adherence rates were better for home-based exercise programs augmented with telephone contacts than for structured programs entailing face-to-face contact.

The findings of two systematic reviews conducted in 2002 and 2004 formed the foundation for physical activity intervention recommendations offered by the CDC-sponsored Guide to Community Preventive Services [24]. Both stressed the importance of tailoring to individual and/or targeted population characteristics.

Marcus and Heimendinger's randomized controlled trial of a tailored informational intervention targeting dietary behavior change demonstrated its effectiveness in improving eating behaviors that significantly improved fruit and vegetable consumption [25]. Further evidence of the promise offered by tailored informational interventions that incorporate information technologies is evident in Brug et al's 1999 review of the literature. The authors concluded that computer-tailored nutrition education is more likely to be read, remembered, and experienced as personally relevant compared to standard materials. They also found that interventions incorporating computer-generated personalized nutrition education delivered via tailored informational interventions are effective in promoting desired dietary behavior changes [18]. Similar results were published in 2001 by Bull et al, who found that tailored health education materials were significantly more effective than nontailored materials at changing dietary behaviors associated with weight loss interventions [26].

O'Neil concluded that self-monitoring on an ongoing basis is a key component of effective dietary behavior change, that self-monitoring enhances weight loss outcomes, and that information technology advances offer promise in improving compliance and effectiveness of self-monitoring [27]. Similarly, in their 2002 systematic review of tailored informational intervention outcomes studies, Ryan et al concluded that they are more effective when ipsative feedback (eg, comparing current to past behavior) was included as a feature of the intervention [16]. Brug et al reached similar conclusions in their 1998 study of the impact of computer-tailored iterative feedback on fat, fruit, and vegetable consumption [28].

Further encouraging findings for eHealth interventions are offered in Latner's 2001 review of obesity self-treatment interventions that included studies of computer-assisted obesity interventions [29]. Several interventions were identified that described pilot intervention studies demonstrating the effectiveness of computer-assisted self-monitoring of food intake and exercise, goal setting, response-contingent feedback, and regular auditory prompts reminding users to enter self-reports. One study was identified that showed no significant difference between the outcomes of computer-assisted self-therapy and a conventional weight loss program using therapist-conducted treatment. Although weight loss was modest in both groups, this study may indicate the potential for substitution of computer-assisted, self-therapy weight loss interventions for more costly and inconvenient provider-delivered interventions.

Of particular interest to eHealth intervention designers working with cellular telephone platforms are Kreps et al's conclusions that conventional (land-line) telephone-delivered tailored informational interventions are generally more effective at promoting health behavior change than printed media interventions [30]. Other studies have shown that tailored informational interventions that utilize combinations of print and telephone-delivered interventions can be highly effective [31].

Also worth noting is the recent trend among commercial weight loss programs such as Weight Watchers and Jenny Craig to incorporate Web-based, self-help tools into their programs [32,33]. Although there are no evaluation studies of these tools currently available in the published literature, there is limited evidence that these commercial weight loss programs in their totality may be effective in promoting behavior changes that yield sustained weight loss among their enrollees [34].

Lessons Learned for eHealth Intervention Development

The potential to leverage eHealth behavioral interventions to improve weight management behaviors appears to be significant. The successes of both computer-assisted, self-help interventions for obesity and tailored informational interventions for a variety of health conditions provide a limited but valuable evidence base that can be leveraged to inform the design of effective eHealth obesity interventions.

Revere and Dunbar's 2000 systematic review of computer-generated outpatient health behavior intervention studies included 37 eligible studies that were published from 1996 to 1999. Of these 37 clinical trials, 34 (91.9%) reported either statistically significant or improved outcomes, and 23 were classified as exhibiting the tailored informational intervention design [2]. These findings would imply that the tenets of conventional tailored informational intervention design translate well to the design of eHealth interventions.

Neuhauser and Kreps conducted an extensive review of the literature on health behavior theory and conventional health behavior intervention outcomes studies in 2003 [15]. Not surprisingly, they found that tailored communication is more effective than generic messages in promoting health behavior change and that health communication is more effective when it reaches people on an emotional as well as a rational level. They cite both social influence theory (eg, Social Cognitive Theory, Social Marketing Model) and evidence from intervention outcomes studies to build a compelling case that interactivity may be the most important trait of effective health behavior interventions and that the involvement of targeted recipients in the design and engagement of the health communication intervention increases the likelihood that they will adopt the desired behaviors. They also suggest that a combination of the effectiveness of interpersonal communication and the reach of mass media communication is needed to change population behavior. This conclusion would appear to have significant implications for eHealth intervention designers working with cell phone and voice over Internet protocol (VoIP) technologies.

The concept of active construction of information bundles, while not an established theory of health behavior change, is also worthy of consideration. Gorman et al's 2001 paper [35] discusses the relationships between individuals' information management activities and maintenance of situation awareness. They assert that as people actively and consciously engage in the creation and/or organization of information into "bundles" to support specific tasks, the act of actively processing and manipulating information improves their understanding and situational awareness. They also cite additional studies [36,37] that support the notion that "over automation" of data entry and information processing may diminish both the users' situation awareness and the usefulness of the information, lowering the probability of achieving the desired outcomes. Applied to eHealth intervention design, this concept would suggest that some data entry and application configuration tasks that could be automated should instead remain in the foreground and fully visible to the end user. The acts of consciously entering and organizing data (eg, estimating food portions, calculating and entering nutritional values for foods eaten) may be just as (or more?) relevant to promoting the desired behavior changes than retrospective use of the captured data.

Technology Adoption: Crossing the Chasm or Finding the Bridges?

As previously discussed, both conventional and eHealth tailored informational interventions have proven to be successful for a variety of health conditions when delivered as individual or small-scale, provider-to-patient interventions. But the difficulties of individual tailoring of message content and interactive delivery have hindered their application to the design of large-scale public health interventions. However, the trend toward rapid convergence of these technologies [38,39] enables a wide variety of desirable eHealth intervention design features that were previously not feasible. Three of the most significant features are interactivity, self-configuration and customization, and mass customization of organizationally sponsored informational interventions (eg, administered by public health organizations, health care providers, health maintenance organizations, or commercial weight loss programs).

Cell phones and networked PDAs enable interactive voice and text communication. However, convergence trends have added real- and near-time multimedia communication capabilities to both. As these device technologies converge, voice-, text-, and multimedia communication modalities are supported in a unified device. Furthermore, as the computing power and memory of these devices increases, users are becoming more empowered to self-configure applications. The content and timing of alerts, reminders, and text memos can be easily customized, and users are increasingly enabled to customize the look, feel, and organization of the user interface of applications running on their cell phones, PDAs, and desktop computers. This technology convergence is an opportunity to deliver organizationally sponsored eHealth obesity interventions.

Health behavior interventions must *reach* the public in order to succeed in promoting and maintaining health behavior change at the population level. eHealth behavioral interventions must therefore be designed and deployed using existing technology development and adoption trends rather than introducing new devices/technologies. By this line of reasoning, the widespread public adoption of cellular telephones [40], wireless PDAs, and use of the Web represent pervasive and rapidly expanding and converging technology adoption trends that should be leveraged in any population-oriented eHealth obesity intervention aiming to reach beyond the desktop. Also, the same newly enabled features of interactivity and self-configuration also provide public health officials with easy facilities for developing and administering flexible and tailored interventions to better meet the needs of specific targeted populations. The trend toward information technology–enabled mass customization of service design and delivery that revolutionized other industries during the 1990s [41] is now possible in public health, medicine, and the commercial weight loss industry.

Table 1 demonstrates, through scenarios, how an eHealth obesity intervention designer might employ the theories, empirical evidence, and technology convergence trends we have discussed.

Table 1. Scenarios illustrating how an eHealth obesity intervention designer might employ the theories, empirical evidence, and technology convergence trends discussed in this article

Convergence-Enabled Feature	Example Use Case
Interactivity	Before eating a meal, a user borrows her friend's Blackberry to access her Web diet journal. She checks the remaining balance in her daily calorie budget, enters the number of calories she wishes to "spend" on the given meal, and is then presented with a personalized list of her "favorite healthy foods" that fall within the range.
Self-Configuration and Customization	A user creates an alert to text message himself at 11:45 am daily with the message "drink your water before going to lunch."
Mass Customization	Weight Watchers clients enroll in a service that sends reminders and Web forms to their smart phones. When opened, they are automatically localized to the recipients' language, food preferences, and FlexPoints targets based on their unique configuration settings.

Conclusions

One study published in 2004 estimated that by the end of 2003, over 60% of all US adults owned an activated cellular telephone, with this statistic growing at slightly over 5% annually [42]. Thus, it would appear that cell phones would be the preferred hardware platform for eHealth obesity interventions for reasons of both enabling effective intervention design features and for promoting rapid public adoption and acceptance.

In their 2000 systematic review of mobile eHealth intervention studies, Revere and Dunbar concluded that “future studies need to identify which [eHealth intervention] models are best suited to which health behavior, whether certain delivery devices are more appropriate for different health behaviors, and how care can benefit from patients’ use of portable devices” [2]. We conclude that the appropriate model for obesity and weight management is the tailored informational intervention modified according to design principles suggested by Social Cognitive Theory and the Social Marketing Model. The health behaviors to target are self-monitoring of diet and physical activity. The devices are Web-enabled “smart” cellular telephones and wireless PDAs. Given the lack of effectiveness of other interventions to prevent or treat obesity in a sustainable matter, trials of these persuasive, ubiquitous technologies are required without delay.

Acknowledgments

This work was made possible by the National Library of Medicine’s Biomedical and Health Informatics Training Program (Grant #T15 LM07442).

Conflicts of Interest

None declared.

References

1. ; World Health Organization, International Obesity Task Force. Obesity: Preventing and Managing the Global Epidemic. Report of a WHO Consultation on Obesity, June 1997. Geneva: World Health Organization; 1998.
2. Revere D, Dunbar PJ. Review of computer-generated outpatient health behavior interventions: clinical encounters "in absentia". J Am Med Inform Assoc 2001;8(1):62-79. [PMC: 11141513] [Medline: 20581373]
3. Smedley BD, Syme SL. Promoting Health: Intervention Strategies from Social and Behavioral Research. Washington, DC: National Academy Press; 2000.
4. Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP. The continuing epidemics of obesity and diabetes in the United States. JAMA 2001;286(10):1195-1200 [FREE Full text] [Medline: 11559264] [doi: 10.1001/jama.286.10.1195]
5. Serdula MK, Mokdad AH, Williamson DF, Galuska DA, Mendlein JM, Heath GW. Prevalence of attempting weight loss and strategies for controlling weight. JAMA 1999;282(14):1353-1358 [FREE Full text] [Medline: 10527182] [doi: 10.1001/jama.282.14.1353]
6. Galuska DA, Will JC, Serdula MK, Ford ES. Are health care professionals advising obese patients to lose weight? JAMA 1999 Oct 27;282(16):1576-1578 [FREE Full text] [Medline: 20012189] [doi: 10.1001/jama.282.16.1576]
7. Frank A. Futility and avoidance. Medical professionals in the treatment of obesity. JAMA 1993 Apr 28;269(16):2132-2133. [Medline: 93225205] [doi: 10.1001/jama.269.16.2132]
8. Harvey EL, Glenny AM, Kirk SFL, Summerbell CD. An updated systematic review of interventions to improve health professionals' management of obesity. Obes Rev 2002 Feb;3(1):45-55. [Medline: 22115775] [doi: 10.1046/j.1467-789X.2002.00053.x]
9. Agency for Healthcare Research and Quality. URL: http://www.ahrq.gov/clinic/3rduspstf/diet/dietrr.pdf [accessed 2005 Aug 21] URL: http://www.ahrq.gov/clinic/3rduspstf/physactivity/physactrr.htm [accessed 2005 Aug 21]
10. Snyder LB, Hamilton MA. A meta-analysis of U.S. health campaign effects on behavior.. In: Public Health Communication: Evidence for Behavior Change. Mahwah, NJ: L. Erlbaum Associates; 2002:357-383.
11. Marcus BH, Owen N, Forsyth LH, Cavill NA, Fridinger F. Physical activity interventions using mass media, print media, and information technology. Am J Prev Med 1998 Nov;15(4):362-378. [Medline: 99056357] [doi: 10.1016/S0749-3797(98)00079-8]
12. Foerster SB, Hudes M. California Dietary Practices Survey: Focus on Fruits and Vegetables, Trends among Adults, 1989-1993, Topline Report. Sacramento, CA: California Department of Health Services and Public Health Foundation; 1994.
13. Hillsdon M, Cavill N, Nanchahal K, Diamond A, White IR. National level promotion of physical activity: results from England's ACTIVE for LIFE campaign. J Epidemiol Community Health 2001 Oct;55(10):755-761 [FREE Full text] [Medline: 21437775] [doi: 10.1136/jech.55.10.755]
14. Kreps GL. The history and development of the field of health communication.. In: Health Communication Research: A Guide to Developments and Directions. Westport, CT: Greenwood Press; 1998:1-15.
15. Neuhauser L, Kreps G. Rethinking communication in the e-Health era. J Health Psychol 2003;8(1):7-22. [doi: 10.1177/1359105303008001426]

16. Ryan P, Lauver DR. The efficacy of tailored interventions. J Nurs Scholarsh 2002;34(4):331-337. [Medline: 22390415] [doi: 10.1111/j.1547-5069.2002.00331.x]
17. Strecher V, Wang C, Derry H, Wildenhaus K, Johnson C. Tailored interventions for multiple risk behaviors. Health Educ Res 2002 Oct;17(5):619-626. [Medline: 22295382] [doi: 10.1093/her/17.5.619]
18. Brug J, Campbell M, Van Assema P. The application and impact of computer-generated personalized nutrition education: a review of the literature. Patient Educ Couns 1999 Feb;36(2):145-156. [Medline: 22891269] [doi: 10.1016/S0738-3991(98)00131-1]
19. Prochaska JO, Diclemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 83239271] [doi: 10.1037//0022-006X.51.3.390]
20. Glanz K, Lewis FM, Rimer BK. Health Behavior and Health Education: Theory, Research, and Practice. 2nd ed. San Francisco: Jossey-Bass; 1997.
21. Emmons KM. Behavioral and social science contributions to the health of adults in the United States. In: Smedley BD, Syme SL, editors. Promoting Health: Intervention Strategies from Social and Behavioral Research. Washington, DC: National Academy Press; 2000:254-321.
22. Stokals D. The social ecological paradigm of wellness promotion. In: Promoting Human Wellness: New Frontiers for Research, Practice, and Policy. Berkeley: University of California Press; 2000:21-37.
23. Baranowski T. How individuals, environments, and health behavior interact: social cognitive theory. In: Health Behavior and Health Education: Theory, Research, and Practice. 2nd ed. San Francisco: Jossey-Bass; 1997:153-178.
24. Guide to Community Preventive Services. URL: http://www.thecommunityguide.org/pa/pa-int-indiv-behav-change.pdf [accessed 2005 Aug 21] URL: http://www.thecommunityguide.org/pa/pa-int-decision-prompts.pdf [accessed 2005 Aug 21]
25. Marcus AC, Heimendinger J, Wolfe P, Rimer BK, Morra M, Cox D, et al. Increasing fruit and vegetable consumption among callers to the CIS: results from a randomized trial. Prev Med 1998;27(5 Pt 2):S16-S28. [Medline: 99030555] [doi: 10.1006/pmed.1998.0405]
26. Bull FC, Holt CL, Kreuter MW, Clark EM, Scharff D. Understanding the effects of printed health education materials: which features lead to which outcomes? J Health Commun 2001;6(3):265-279. [Medline: 21435382] [doi: 10.1080/108107301752384442]
27. O'neil PM. Assessing dietary intake in the management of obesity. Obes Res 2001 Dec;9 Suppl 5(suppl 5):361S-366S [FREE Full text] [Medline: 21915275]
28. Brug J, Glanz K, Van Assema P, Kok G, Van Breukelen GJ. The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. Health Educ Behav 1998 Aug;25(4):517-531. [Medline: 98354565]
29. Latner JD. Self-help in the long-term treatment of obesity. Obes Rev 2001 May;2(2):87-97. [Medline: 22115777] [doi: 10.1046/j.1467-789x.2001.00028.x]
30. Kreps GL. The role of interactive technology in cancer communications interventions: targeting key audience members by tailoring messages. Presented at: American Public Health Association Conference; November 1999; Boston, MA.
31. Rimer BK, Halabi S, Sugg Skinner C, Kaplan EB, Crawford Y, Samsa GP, et al. The short-term impact of tailored mammography decision-making interventions. Patient Educ Couns 2001 Jun;43(3):269-285. [Medline: 21279247]
32. Jenny Craig. URL: http://www.jennycraig.com/etools/ [accessed 2005 Jul 29]
33. Weight Watchers. URL: http://www.weightwatchers.com/plan/www/etools_01.aspx?navid=etoolsaag [accessed 2005 Jul 29]
34. Heshka S, Anderson JW, Atkinson RL, Greenway FL, Hill JO, Phinney SD, et al. Weight loss with self-help compared with a structured commercial program: a randomized trial. JAMA 2003 Apr 9;289(14):1792-1798 [FREE Full text] [doi: 10.1001/jama.289.14.1792] [Medline: 22571305]
35. Gorman PN, Ash JS, Lavelle M, Lyman J, Delcambre L, Maier D. Bundles in the wild: managing information to solve problems and maintain situation awareness. Library Trends 2000;49(2):266-289.
36. Ash JS, Gorman PN, Hersh WR, Lavelle M, Poulsen SB. Perceptions of house officers who use physician order entry. Proc AMIA Symp 1999:471-475 [FREE Full text] [Medline: 20032913]
37. Von Thaden T. Social informatics and aviation technology. Bulletin of the American Society for Information Science 2000;26(3):13-14. [doi: 10.1002/bult.152]
38. Fink F. Multi-purpose portable telephones with seamless mobility. Elektronik Praxis 2004;21:32-36.
39. Hacklin F, Raurich V, Marxt C. How incremental innovation becomes disruptive: the case of technology convergence. Presented at: IEEE International Engineering Management Conference; Oct 18-21, 2004; Singapore.
40. Rice RE, Katz JE. Comparing internet and mobile phone usage. Telecomm Policy 2003;27(8-9):597-623. [doi: 10.1016/S0308-5961(03)00068-5]
41. Tseng MM, Jiao J. Mass customization. In: Handbook of Industrial Engineering, Technology and Operations Management. 3rd ed. New York: Wiley; 2001:685.
42. Tuckel P, O'Neill H. Ownership and Usage Patterns of Cell Phones: 2000-2004. Presented at: Annual Meeting of the American Association for Public Opinion Research; May 13-16, 2004; Phoenix, Arizona.

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System
CDC: Centers for Disease Control and Prevention
PDA: personal digital assistant
VoIP: voice over Internet protocol

submitted 31.08.05; peer-reviewed by I Buchan; comments to author 21.10.05; revised version received 02.11.05; accepted 02.11.05; published 20.12.05.

Please cite as:
Tufano JT, Karras BT
Mobile eHealth Interventions for Obesity: A Timely Opportunity to Leverage Convergence Trends
J Med Internet Res 2005;7(5):e58
URL: http://www.jmir.org/2005/4/e1/
doi:10.2196/jmir.7.5.e58
PMID:16403722

Original Paper

Feasibility of Adding Enhanced Pedometer Feedback to Nutritional Counseling for Weight Loss

Caroline R Richardson[1], MD; Beverley B Brown[2], MD; Sharon Foley[3], MS, RD; Kathleen S Dial[4], MS, RD; Julie C Lowery[1], PhD

[1]Veterans Affairs Medical Center and University of Michigan Health Systems, Department of Family Medicine, Ann Arbor, MI, USA
[2]University of Pittsburgh Medical Center, Pittsburgh, PA, USA
[3]Veterans Affairs Medical Center, Chicago, IL, USA
[4]Veterans Affairs Medical Center, Tucson, AZ, USA

Corresponding Author:
Caroline R Richardson, MD
Department of Family Medicine
University of Michigan
1018 Fuller St
Ann Arbor, MI
48109-0708
USA
Phone: +1 734 998 7120 ext 316
Fax: +1 734 998 7335
Email: caroli@umich.edu

Abstract

Background: Intensive interventions targeting diet and physical activity are effective for weight reduction but are costly. Tailored, computer-generated, step-count feedback may provide an intensive and affordable way to increase the physical activity of people at high risk for cardiovascular disease.

Objective: The objective was to test the feasibility of adding tailored, computer-generated, step-count feedback to a face-to-face nutritional counseling weight loss intervention.

Methods: We recruited 12 participants, 4 from each of three Department of Veterans Affairs medical centers. There were 11 male participants and 1 female participant. Each had a body mass index of 30 or greater and at least one of the following cardiovascular disease risk factors: diabetes, hypertension, hypercholesterolemia, obesity, or coronary artery disease. Participants attended one-on-one counseling sessions with a registered dietitian for four sessions over three weeks. At the initial session, each participant received an enhanced pedometer to record time-stamped, step-count data. Participants wore the device daily throughout the intervention. At the three follow-up sessions, the dietitian uploaded the computer data, reviewed a Web-based graphical display of step-count feedback, and helped set new walking goals.

Results: All 12 participants completed the program (100% attendance). Initial mean weight was 255 lbs (SD = 49 lbs), and weight loss was just over 4 lbs (n = 12, paired *t* test, *P* = .004). Mean daily step counts during the first week averaged 6019 steps per day, increasing to an average of 7358 per day after the third week (average increase of 1339 steps per day, or 0.6 miles, or 12 minutes of walking, n = 10, paired *t* test, *P* = .04).

Conclusions: Enhanced pedometer feedback in conjunction with nutritional counseling is feasible and results in significant weight loss and increased walking among individuals at high risk for cardiovascular disease.

(J Med Internet Res 2005;7(5):e56) doi:10.2196/jmir.7.5.e56

KEYWORDS

obesity; diet; exercise; cardiovascular diseases; risk factors; Internet

Introduction

Being sedentary and eating a poor diet increases the risk of developing diabetes [1-3], heart disease [4-8], hypertension [9], depression [10-12], and many other chronic diseases [13]. We now know that with intensive monitoring, counseling, and encouragement, individuals can change their diet and exercise behaviors and that such changes result in dramatic improvements in health outcomes. In the Diabetes Prevention Program study, over 1000 individuals at high risk for developing diabetes who were randomized to an intensive diet and physical activity intervention decreased their risk of developing diabetes by almost 60% over three years compared to individuals in the usual care arm [2]. Unfortunately, brief clinical interventions designed to help patients change their diet and physical activity have not been effective [14]. The challenge is to find cost-effective and feasible ways to integrate intensive diet and exercise interventions into clinical practice.

One approach to decreasing the cost and increasing the intensity of lifestyle interventions is to automate the process of diet and activity counseling using tailored Internet-based systems. While such systems have shown promise, low participant adherence in terms of returning to the website for reinforcement has proven to be a major barrier [15-18]. Automated systems may work better to augment existing relationships with health care providers rather than replacing the health care provider all together.

Behavioral researchers have shown that people have poor recall of daily walking duration, intensity, and frequency [19]. This poor recall makes it difficult to set short-term, incremental behavioral goals or to reward successes [16,20]. Clear and objective feedback about daily walking patterns can be useful in overcoming barriers, setting goals, and rewarding successes. Pedometers are devices worn on the waist or belt that count steps taken throughout the day. Walking interventions that use pedometers and specific step-count goals are more effective than walking interventions that use time-based goals. Hultquist et al found that sedentary women who were given pedometers and who were instructed to walk 10000 steps a day walked almost 2000 steps per day more than women who were instructed to go for a brisk 30-minute walk each day [21].

Unlike simple pedometers that keep a single running total of step counts, enhanced pedometers record detailed, time-stamped, step-count data. Once these data are uploaded to a computer, they can be used to automatically generate individually tailored reports about the duration and intensity of walking bouts throughout the day. Users can then review the tailored step-count feedback in graphical or numerical form.

In this study, we examine the feasibility of integrating an Internet-based enhanced pedometer system into routine nutritional counseling for individuals with cardiovascular disease (CVD) risk factors. Our primary goal was to assess health service delivery parameters, including required face-to-face counseling time, participant satisfaction, and technical barriers to delivering this type of intervention. We also assessed weight loss and increased walking among the participants.

Methods

Study Design

This was a multi-center intervention feasibility study with a pre-post design. The intervention consisted of four individual face-to-face nutritional counseling sessions for weight loss over a three-week period. Dietitians incorporated computer-generated, tailored, step-count feedback and goal setting to increase walking into each of the four visits. We recruited 12 participants, 4 from each of three different Veterans Affairs (VA) medical centers to assess portability and feasibility of the intervention. This study was approved by the Investigational Review Boards at the participating sites and by the Ann Arbor Veterans Affairs Medical Center. All eligible participants signed written informed consent documents prior to enrollment.

Study Population

Participants were adult VA patients who were referred by a VA physician for outpatient nutritional counseling. Dietitians at each site offered patients the opportunity to participate if they were ambulatory, sedentary, overweight (body mass index [BMI] > 28), and were at high risk for CVD. Patients were considered ambulatory if they could comfortably walk one block. Sedentary was defined as less than 150 minutes of physical activity of at least moderate intensity each week by self-report. Individuals were considered at high CVD risk if they had one of the following diagnoses: (1) type 2 diabetes, (2) hypertension, (3) hypercholesterolemia, (4) obesity (BMI > 30), or (5) known coronary artery disease. CVD risk status was assessed by self-report and confirmed by chart review. Smoking was not included in this list because there is relatively little evidence that physical activity reduces CVD risk in smokers who have no other CVD risk factors. All participants were either in the contemplation or preparation stage for starting a walking program [22,23].

Participants were not required to get medical clearance to participate in this study. Because the study emphasized a gradual increase in walking at a moderate pace, the risk of adverse cardiac events while walking was low. However, participants who reported chest pain, shortness of breath, or lightheadedness while walking, who had been advised by a physician not to walk, or who were undergoing a cardiac work-up were excluded from the study. Individuals who had been enrolled in another nutritional counseling program within the past 28 days were excluded. Participants were also excluded if they could not comfortably communicate in English or if they were not competent to give informed consent.

The Pedometer

All participants were given a SportBrain First Step enhanced pedometer (www.sportbrain.com) to wear throughout the intervention period. The device we used uploaded step counts over a phone line to a central computer. The SportBrain website then displayed individually tailored graphs of step counts along with motivational messages. Participants wore the pedometer every day, all day, for approximately 22 days per participant (4 weekly sessions). While more recent versions of the enhanced pedometers display the total daily step count on the pedometer itself, the First Step device used in this trial did not have a display on the pedometer. The only way to review the step-count feedback was to upload it to the computer and review it online. Figure 1 shows a sample of the Internet-based feedback provided on the SportBrain website. A wide variety of graphs are available on the website. The graph in Figure 1 shows time on the x axis and miles per hour on the y axis for a period of high activity on one day. Time stamping in the step-count data allows estimation of physical activity intensity and duration of activity bouts. Figure 2 shows the SportBrain pedometer and upload dock that we used in this intervention.

Intervention

Each participant met with a dietitian during four individual face-to-face sessions for both diet and physical activity counseling. The diet intervention was a structured nutritional counseling intervention based on the American Dietetic Association’s Medical Nutrition Therapy Weight Management Protocol [24]. This is a detailed six-session program, which we modified to a more feasible four-session protocol.

During the first session, which lasted approximately one hour, the study was described, and patients who agreed to participate and met eligibility criteria completed a written informed consent process. Participants were given a SportBrain First Step pedometer and were taught how to position the pedometer on their waistband, attach the safety clip to their belt or clothes, and check to see if the pedometer was counting steps properly using a 20-step test walk. Participants also completed a brief nutritional assessment and received preliminary nutritional counseling.

During visits 2 through 4, participants met individually face-to-face with their study dietitian for both nutritional counseling and reinforcement of the walking program. Participants reviewed step-count data on the computer with the dietitian and negotiated new step-count goals for the upcoming week. Step-count goals were increased gradually. For example, a participant who averaged 3000 steps a day during week one might have set a new goal of 3500 steps per day for week two. Dietitians then discussed nutritional counseling topics using standardized handouts and set new nutritional goals for the week. Finally, participants filled out a brief questionnaire about attitudes toward the Internet-based step-count feedback, problems using the enhanced pedometer, and self-efficacy for achieving walking goals. Participants received US$10 for each of the four visits attended. At the end of the session, participants who returned the enhanced pedometers received a “Veterans Walk for Health” T-shirt valued at US$10.

Measures

Feasibility Outcomes

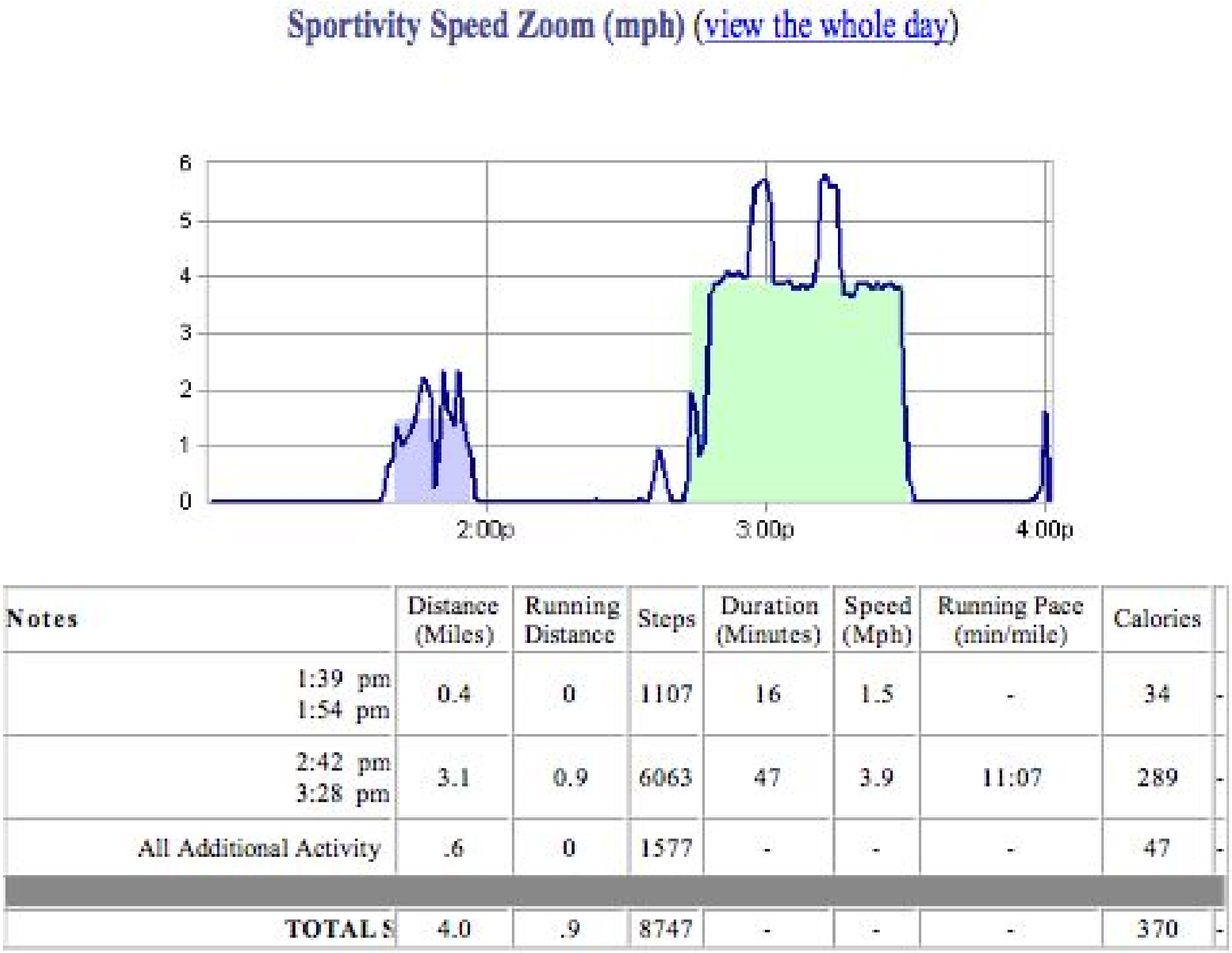
The following components of feasibility were measured explicitly:

1. **Participant satisfaction** with the intervention as measured by responses to a brief satisfaction survey. This survey included questions about overall satisfaction with the intervention and about willingness to pay to continue to have access to the intervention after the study was over. Finally, participants were asked what they liked and did not like about the enhanced pedometer component of the intervention in a series of open-ended questions.
2. **Participant adherence** to the intervention was assessed using three measures: (1) attendance at visits, (2) self-report of the number of study days that pedometer was worn, and (3) electronic step-count data indicating days that the pedometer was worn.
3. **Counseling time:** Mean face-to-face dietitian counseling time required to deliver the physical activity and nutritional counseling components of the intervention was recorded at each visit by the study dietitian on a study visit data collection form.
4. **Technical problems** with uploading the pedometer data to the website were logged by the study dietitian on the study visit data collection form at each visit.

Intervention Outcomes

1. **Participant weight:** A single clinical scale at each site was designated as the study scale. Scales were zeroed before each participant was weighed. Participants were weighed without shoes, and weight was recorded at each visit.
2. **Mean daily step count:** During the final week of the intervention, this was recorded electronically by uploading the data to the website.
3. **Participant self-efficacy** for adhering to a walking program as measured by participants' response to a single question: "I am confident that I can reach my new step-count goals next week." (possible responses were "Strongly Disagree", "Disagree", "Neutral", "Agree", and "Strongly Agree").

Figure 1. Sample graph and table from the SportBrain website



Notes	Distance (Miles)	Running Distance	Steps	Duration (Minutes)	Speed (Mph)	Running Pace (min/mile)	Calories	
1:39 pm 1:54 pm	0.4	0	1107	16	1.5	-	34	-
2:42 pm 3:28 pm	3.1	0.9	6063	47	3.9	11:07	289	-
All Additional Activity	.6	0	1577	-	-	-	47	-
TOTAL S	4.0	.9	8747	-	-	-	370	-

Figure 2. The SportBrain First Step enhanced pedometer and Sportport interface

Training for the dietitians who participated in the study took place in two phases. After an introductory conference call, all participating dietitians wore the pedometer for one week, uploaded their own walking data at least twice during this trial period, and learned to navigate around the SportBrain website by viewing different formats for step-count feedback and changing their step-count goals. After the trial period, we had a series of conference calls to discuss and troubleshoot technical problems with the devices and to review counseling guidelines.

Sample Size

We did not power this feasibility study to be able to detect clinically significant weight loss or increase in average daily step counts. Instead, we intentionally kept the overall sample size small but recruited from three different VA medical centers that were remote from the coordinating center. Recruiting from multiple sites increased the complexity of the study in that we were required to get human subjects committee approval from four different medical centers, and we had to train six different dietitians. However, recruiting from multiple sites increased the power to examine potential barriers to implementation and feasibility issues across multiple settings.

Data Analysis

Means and standard deviations were used to estimate time spent in both the step-count and the nutritional counseling components of the intervention. Mean participant satisfaction and percentage of sessions attended were also calculated. For pre-post comparisons, we used paired *t* tests. We compared mean daily step counts for the first 7 days of the intervention with mean daily step counts for the final 7 days of the intervention. Similarly, paired *t* tests were used to test for statistically significant weight loss.

We anticipated that, as the participant became more familiar with the pedometer and with the nutritional counseling component of the intervention, the time required for counseling might decrease. To test this hypothesis, we compared the counseling time difference between the second and fourth visits. We did not use the initial session for time comparison because of the substantial time required to enroll the participant in the study and explain the intervention. Finally, participant comments in response to open-ended questions about the intervention were tabulated and reviewed for common themes. All statistical analyses were performed using Stata version 8 [25].

Results

Quantitative Results

From September to December 2003, a total of 12 participants (11 male and 1 female) were enrolled from three different VA hospitals. Recruitment of four eligible and interested participants was completed within one month at each of the three sites for an average recruitment rate of one participant per week per site. Though obesity was not a criterion for enrollment, all 12 participants had a BMI greater than 30. Participants weighed an average of 255 pounds (SD = 49) at baseline. BMI averaged 37 (SD = 6.5). Ten of the 12 participants also had at least one other CVD risk factor in addition to their obesity. The participants were followed for an average of 21 days (range 18-27 days), including the initial day and last day. Table 1 provides details on the participants' baseline characteristics.

Table 1. Participant baseline characteristics (N = 12)

Chronic Conditions	**N (%)**
Diabetes	2 (17%)
Hypertension	8 (67%)
Hypercholesterolemia	6 (50%)
Obesity	12 (100%)
Coronary artery disease	1 (8%)
	Mean+SD
Height (inches)	69.4 + 2.5
Baseline weight (pounds)	254.9 + 49.4
Body mass index (kg/m^2)	37 + 6.5
Age (years)	52.7
	Mean+SD or (%)
Comfort with Computers (1 = not; 4 = very)	3.6 + 0.5
1	0
2	0
3	5 (41.7%)
4	7 (58.3%)
Initial Self-Efficacy to Increase Walking (1 = not; 5 = very)	4.5 + 0.7
1	0
2	0
3	1 (8.3%)
4	4 (33.3%)
5	7 (58.3%)

All 12 pedometers were returned in working condition at the end of the study. Of the 36 attempts to upload step-count data (12 participants times 3 sessions), 32 were successful. During 4 (11%) of the sessions, there were problems with uploading the data. Two of the upload attempts failed for administrative reasons related to access to the website, and one failed because the Internet connection was down. The remaining session failed because the pedometer was not recording steps properly.

Table 2 provides participant satisfaction measures as well as face-to-face time per visit. All 12 participants attended each of the four scheduled study visits for a 100% attendance rate and a 100% follow-up rate. Electronic monitoring data show that participants wore the pedometer for more days than they could recall. Participants reported remembering to wear the pedometer for the whole day for 95% (243 of 257) of patient days in the study. According to electronically recorded step-count data, the pedometers were worn at least part of the day for 99% (254 of 257) of patient days of the study.

Overall participant satisfaction at the final visit averaged 4.4 out of 5 on a 5-point scale, in which 5 indicated very satisfied; 11 out of 12 were very satisfied or satisfied. Satisfaction was also assessed by asking participants if they would be willing to pay for the enhanced pedometer. The cost participants were willing to assume ranged from US$0 to $100 per year. The average price a veteran was willing to pay for a year of pedometer service was US$20.50. The average total time per face-to-face counseling session was 33 minutes.

Table 2. Main results

	Mean+SD or (%)
Willingness to pay for service (US$)	20.50 + 30.8
Nutritional counselling session attendance, 4 total	4 + 0
Participant satisfaction with intervention (1 = very unsatisfied; 5 = very satisfied)	
1	1 (8.3%)
2	0
3	0
4	3 (25.0%)
5	8 (66.7%)
Post-Intervention Self-Efficacy to Increase Walking(1 = not; 5 = very)	
1	0
2	0
3	0
4	8 (66.7%)
5	4 (33.3%)
Counseling time for nutrition (min)	
Visit 1 (initial set-up)	19.49 + 12.15
Visit 2	21.25 + 10.47
Visit 3	14.00 + 6.99
Visit 4	13.64 + 8.97
Counseling time for step count (min)	
Visit 1 (initial set-up)	13.92 + 11.85
Visit 2	16.82 + 9.02
Visit 3	18.64 + 8.09
Visit 4	14.09 + 6.64
Total counseling time for both (min)	
Visit 1 (initial set-up)	33.41
Visit 2	38.07
Visit 3	32.64
Visit 4	27.73

The average weight loss was 4.1 pounds over three weeks (n = 12, paired *t* test, *P* = .004); 11 out of 12 participants lost weight. Weight loss ranged from –1.8 pounds to 9.7 pounds. Only one of the 12 participants gained weight. By the last week of the trial, participants significantly increased daily step counts. Mean daily step counts during the first week averaged 6019 steps per day, increasing to an average of 7358 per day after the third week (average increase of 1339 steps per day, n = 10, paired *t* test, *P* = .04). This increase translates into a daily increase of 0.6 miles, or 12 minutes of walking assuming that 2200 steps is a mile and that the average walking pace is 3 miles per hour. There was no significant change in self-efficacy between the first and final session. There were no major adverse events and, in particular, no adverse cardiovascular events during the intervention.

Qualitative Results

The most common problem with the pedometer was difficulty attaching it to the waistband or belt. Participants liked viewing graphs of step counts because it documented their successes and made it easy to see step-count targets or goals. One participant commented that “it is a new way to look at my walks…” and another liked “knowing what [he] accomplished for the week.” Others liked seeing the relationship between steps taken and calories burned, which is printed in a table on the website. Participants also liked being able to see the number of steps taken during a particular activity, such as on a specific walk or while shopping. In general, the participants did not like seeing that they were not doing enough walking to meet their step-count goals. Even participants who regularly met their step-count goals were worried about how they would feel if they failed.

Discussion

Our study demonstrates the feasibility of incorporating enhanced pedometers and Internet-based feedback into traditional nutritional counseling programs to promote weight loss. We were able to implement the enhanced pedometer intervention in three different medical centers with minimal training and technical support. While participants did have some trouble wearing the pedometer, they did not have trouble understanding the step-count feedback graphs, and satisfaction with the intervention was high. While we did not expect to see significant changes in either weight or step counts in this small and brief feasibility study, results for weight loss and step-count increases were both statistically and clinically significant.

Several recent studies have examined the effectiveness of Internet-based weight loss programs without face-to-face counseling [18,26,27]. The interventions tested in these studies have relied primarily on self-reported diet and physical activity logs and text-based behavioral counseling. Our intervention differs from these previously tested Internet-based weight loss programs in several ways. First, the enhanced pedometer and Internet-based, step-count feedback were used as an intensive and objective monitoring tool and not as a substitute for one-on-one counseling. By automating the process of uploading step-count data, enhanced pedometers reduce respondent burden involved in filling out logs and also increase the detail and accuracy of the data. Secondly, our intervention had a balanced approach, focusing on both diet and physical activity in approximately equal parts. While the exact role of physical activity in initial weight loss is still being debated, there is a consensus that physical activity is critical for long-term weight control [28-30].

The theory of self-regulation suggests that helping individuals learn how to more accurately and more intensively self-monitor the behavior they are trying to change should result in improved outcomes [16,20]. At least one study has shown that consistent self-monitoring of diet leads to better weight loss and control [33]. Similarly, success with simple pedometer interventions suggests that objective monitoring and step-count feedback can increase walking [21,34,35]. Using the existing information technology to optimize self-monitoring of both physical activity and diet in combination with one-on-one counseling, delivered either face-to-face, by telephone, or by email, may prove to be a cost-effective intervention for initially reducing weight and for subsequently maintaining weight loss. Detailed step-count feedback may augment the effect of face-to-face counseling sessions and may reduce the face-to-face counseling time required for effective behavior change interventions. More research will be needed to determine the optimal frequency and intensity of self-monitoring, group sessions, and one-on-one contact to maximize long-term maintenance of weight loss while keeping costs reasonable.

The cost of enhanced pedometer systems is dropping rapidly. When this feasibility study was conducted, the cost of a SportBrain enhanced pedometer, a Sportport computer interface dock, and a year-long subscription to the website was about US$200. Currently, the new and improved version of the SportBrain enhanced pedometer can be purchased for US$30 or less, a device-specific interface port is no longer needed, and basic Web access is free. There are a number of other companies on the market that are producing enhanced pedometers, and the cost is likely to continue to drop.

Limitations to this study include the small sample size, the short duration of the intervention, and the lack of a control group. Despite the small sample size and short duration, we did find significant reductions in weight and increases in walking. In addition, this study used enhanced pedometers with nutritional counseling for all participants, and hence, we are unable to estimate the component of weight loss that was attributable to the enhanced pedometer system alone.

There are many advantages to having detailed time-stamped data uploaded to a central computer rather than relying on the participant to log daily step counts on a paper calendar or type them into a Web page. The decreased respondent burden, the ability to automate the creation of tailored motivational messages based on previous step counts, as well as the ability to examine detailed walking patterns throughout the day, including duration and intensity data, may make enhanced pedometers well worth the small increase in cost over simple pedometers. The benefits of using an enhanced pedometer will depend on our ability to create sophisticated, reliable, and user-friendly, Web-based systems for clinical interventions. This approach may be best suited to a health care system using a team approach of dietitian, patient, and physician combined with Internet monitoring. Creating disease-specific walking websites that give enhanced pedometer feedback or websites that can tailor motivational messages and set goals based on current disease state may further increase the value of such interventions [36-38].

Advances in information technology have yielded low-cost, user-friendly, intensive physical activity monitoring and feedback systems. Results of this study demonstrate that enhanced pedometers with Web-based, step-count feedback can be integrated into a nutritional counseling program focusing on weight loss. This intervention is portable as it was successfully implemented at three different medical centers. Participants lost weight and increased their daily walking during the brief intervention. Further research and randomized controlled trials with assessment of long-term outcomes are needed to test the impact of incorporating enhanced pedometer step-count feedback into routine medical care for high-risk individuals.

Acknowledgments

Funding for this study was provided by the Health Services Research and Development Center for Practice Management and Outcomes Research, VA Medical Center, Ann Arbor, Michigan. This study was conducted at the request of, and with input from,

the Research and Professional Oversight Group of the Nutrition Advisory Board for the Veterans Health Administration Nutrition and Food Service, in an effort to address obesity, a major health problem for VA patients.

Conflicts of Interest

None declared.

References

1. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-parikka P, et al; Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001 May 3;344(18):1343-1350 [FREE Full text] [Medline: 21215436] [doi: 10.1056/NEJM200105033441801]
2. Knowler WC, Barrett-connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 21822558]
3. Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. N Engl J Med 2001 Sep 13;345(11):790-797 [FREE Full text] [Medline: 21418115] [doi: 10.1056/NEJMoa010492]
4. Hu FB, Stampfer MJ, Solomon C, et al. Physical activity and risk for cardiovascular events in diabetic women. Ann Intern Med 2001;134(2):96-105.
5. Manson JE, Hu FB, Rich-Edwards JW, Colditz GA, Stampfer MJ, Willett WC, et al. A prospective study of walking as compared with vigorous exercise in the prevention of coronary heart disease in women. N Engl J Med 1999 Aug 26;341(9):650-658 [FREE Full text] [Medline: 99383645] [doi: 10.1056/NEJM199908263410904]
6. Paffenbarger RS, Hyde RT, Wing AL, Hsieh CC. Physical activity, all-cause mortality, and longevity of college alumni. N Engl J Med 1986 Mar 6;314(10):605-613. [Medline: 86118507]
7. Williams PT. Physical fitness and activity as separate heart disease risk factors: a meta-analysis. Med Sci Sports Exerc 2001 May;33(5):754-761. [Medline: 21223197]
8. Richardson CR, Kriska AM, Lantz PM, Hayward RA. Physical activity and mortality across cardiovascular disease risk groups. Med Sci Sports Exerc 2004 Nov;36(11):1923-1929. [Medline: 15514508]
9. Hayashi T, Tsumura K, Suematsu C, Okada K, Fujii S, Endo G. Walking to work and the risk for hypertension in men: the Osaka Health Survey. Ann Intern Med 1999 Jul 6;131(1):21-26 [FREE Full text] [Medline: 99308804]
10. Lampinen P, Heikkinen RL, Ruoppila I. Changes in intensity of physical exercise as predictors of depressive symptoms among older adults: an eight-year follow-up. Prev Med 2000 May;30(5):371-380. [Medline: 20302369] [doi: 10.1006/pmed.2000.0641]
11. Weyerer S. Physical inactivity and depression in the community. Evidence from the Upper Bavarian Field Study. Int J Sports Med 1992 Aug;13(6):492-496. [Medline: 93053037]
12. Camacho TC, Roberts RE, Lazarus NB, Kaplan GA, Cohen RD. Physical activity and depression: evidence from the Alameda County Study. Am J Epidemiol 1991 Jul 15;134(2):220-231. [Medline: 91320894]
13. ; US Department of Health and Human Services. Physical Activity and Health: A Report of the Surgeon General. Atlanta, GA: Center for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996.
14. Whitlock EP, Orleans CT, Pender N, Allan J. Evaluating primary care behavioral counseling interventions: an evidence-based approach. Am J Prev Med 2002 May;22(4):267-284. [Medline: 21983806] [doi: 10.1016/S0749-3797(02)00415-4]
15. Anhøj J, Jensen AH. Using the internet for life style changes in diet and physical activity: a feasibility study. J Med Internet Res 2004 Sep 8;6(3):e28 [FREE Full text] [Medline: 15471754] [doi: 10.2196/jmir.6.3.e28]
16. Boekaerts M, Pintrich PR, Zeidner M. Handbook of Self-Regulation. San Diego, CA: Academic Press; 2000.
17. Leslie E, Marshall AL, Owen N, Bauman A. Engagement and retention of participants in a physical activity website. Prev Med 2005 Jan;40(1):54-59. [Medline: 15530581] [doi: 10.1016/j.ypmed.2004.05.002]
18. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA 2003 Apr 9;289(14):1833-1836 [FREE Full text] [doi: 10.1001/jama.289.14.1833] [Medline: 22571311]
19. Welk GJ. Physical activity assessments for health-related research. Alpha, NJ: Human Kinetics; 2002.
20. Cameron LD, Leventhal H. The Self-Regulation of Health and Illness Behavior. London: Routledge; 2003.
21. Hultquist CN, Albright C, Thompson DL. Comparison of walking recommendations in previously inactive women. Med Sci Sports Exerc 2005 Apr;37(4):676-683. [Medline: 15809569]
22. Marcus BH, Forsythe LH. Motivating people to be physically active. Champaign, Ill: Human Kinetics; 2003.
23. Prochaska J, DiClemente C. Transtheoretical therapy toward a more integrative model of change. Psychotherapy: Theory, Research and Practice 1982;19(3):276-287.
24. Medical Nutrition Therapy across the Continuum of Care: Client Protocols / Supplement 1. Chicago, IL: American Dietetic Association; 1997.

25. ; Stata Corp. Stata Statistical Software: Release 8.0. 2003.
26. Harvey-Berino J, Pintauro S, Buzzell P, Digiulio M, Casey Gold B, Moldovan C, et al. Does using the Internet facilitate the maintenance of weight loss? Int J Obes Relat Metab Disord 2002 Sep;26(9):1254-1260. [doi: 10.1038/sj.ijo.0802051] [Medline: 22174567]
27. Womble LG, Wadden TA, Mcguckin BG, Sargent SL, Rothman RA, Krauthamer-Ewing ES. A randomized controlled trial of a commercial internet weight loss program. Obes Res 2004 Jun;12(6):1011-1018 [FREE Full text] [Medline: 15229342]
28. Fogelholm M, Kukkonen-harjula K, Nenonen A, Pasanen M. Effects of walking training on weight maintenance after a very-low-energy diet in premenopausal obese women: a randomized controlled trial. Arch Intern Med 2000 Jul 24;160(14):2177-2184. [Medline: 20366601] [doi: 10.1001/archinte.160.14.2177]
29. Blair SN. Evidence for success of exercise in weight loss and control. Ann Intern Med 1993 Oct 1;119(7 Pt 2):702-706. [Medline: 93370824]
30. Votruba SB, Horvitz MA, Schoeller DA. The role of exercise in the treatment of obesity. Nutrition 2000 Mar;16(3):179-188. [Medline: 20171136] [doi: 10.1016/S0899-9007(99)00264-6]
31. King AC, Tribble DL. The role of exercise in weight regulation in nonathletes. Sports Med 1991 May;11(5):331-349. [Medline: 91297269]
32. Blair SN, Church TS. The fitness, obesity, and health equation: is physical activity the common denominator? JAMA 2004 Sep 8;292(10):1232-1234 [FREE Full text] [Medline: 15353537] [doi: 10.1001/jama.292.10.1232]
33. Baker RC, Kirschenbaum DS. Weight control during the holidays: highly consistent self-monitoring as a potentially useful coping mechanism. Health Psychol 1998 Jul;17(4):367-370. [Medline: 98361411] [doi: 10.1037/0278-6133.17.4.367]
34. Tudor-locke C, Bell RC, Myers AM, Harris SB, Ecclestone NA, Lauzon N, et al. Controlled outcome evaluation of the First Step Program: a daily physical activity intervention for individuals with type II diabetes. Int J Obes Relat Metab Disord 2004 Jan;28(1):113-119. [Medline: 14569279] [doi: 10.1038/sj.ijo.0802485]
35. Yamanouchi K, Shinozaki T, Chikada K, Nishikawa T, Ito K, Shimizu S, et al. Daily walking combined with diet therapy is a useful means for obese NIDDM patients not only to reduce body weight but also to improve insulin sensitivity. Diabetes Care 1995 Jun;18(6):775-778. [Medline: 96002685]
36. Marcus BH, Forsyth LH. Tailoring interventions to promote physically active lifestyles in women. Womens Health Issues 1998;8(2):104-111. [Medline: 98203592] [doi: 10.1016/S1049-3867(97)00102-3]
37. Marcus BH, Bock BC, Pinto BM, Forsyth LH, Roberts MB, Traficante RM. Efficacy of an individualized, motivationally-tailored physical activity intervention. Ann Behav Med 1998;20(3):174-180. [Medline: 99143869]
38. Strecher V, Wang C, Derry H, Wildenhaus K, Johnson C. Tailored interventions for multiple risk behaviors. Health Educ Res 2002 Oct;17(5):619-626. [Medline: 22295382] [doi: 10.1093/her/17.5.619]

submitted 14.06.05; peer-reviewed by D Tate, J VanWormer; comments to author 30.08.05; revised version received 31.10.05; accepted 31.10.05; published 17.11.05.

Please cite as:
Richardson CR, Brown BB, Foley S, Dial KS, Lowery JC
Feasibility of Adding Enhanced Pedometer Feedback to Nutritional Counseling for Weight Loss
J Med Internet Res 2005;7(5):e56
URL: http://www.jmir.org/2005/5/e56/
doi:10.2196/jmir.7.5.e56
PMID:16403720

Original Paper

Parent-Child Interaction Using a Mobile and Wireless System for Blood Glucose Monitoring

Deede Gammon[1], MA; Eirik Årsand[1], MSc; Ole Anders Walseth[1], MSc; Niklas Andersson[1], MSc; Martin Jenssen[1], MA; Ted Taylor[2], PhD

[1]Norwegian Centre for Telemedicine, Tromsø, Norway

[2]Oregon Research Institute, Oregon, USA

Corresponding Author:
Deede Gammon, MA
Norwegian Centre for Telemedicine
PO Box 35
9038 Tromsø
Norway
Phone: +47 777 54000
Fax: +47 777 54098
Email: deede.gammon@telemed.no

Abstract

Background: Children with type 1 diabetes and their parents face rigorous procedures for blood glucose monitoring and regulation. Mobile telecommunication systems show potential as an aid for families' self-management of diabetes.

Objective: A prototype designed to automatically transfer readings from a child's blood glucose monitor to their parent's mobile phone was tested. In this formative stage of development, we sought insights into the appropriateness of the concept, feasibility of use, and ideas for further development and research.

Methods: During four months, a self-selected sample of 15 children (aged 9 to 15 years) with type 1 diabetes and their parents (n = 30) used the prototype approximately three times daily. Parent and child experiences were collected through questionnaires and through interviews with 9 of the parents.

Results: System use was easily integrated into everyday life, and parents valued the sense of reassurance offered by the system. Parents' ongoing struggle to balance control of their children with allowing independence was evident. For children who measured regularly, use appeared to reduce parental intrusions. For those who measured irregularly, however, parental reminders (eg, "nagging") appeared to increase. Although increased reminders could be considered a positive outcome, they can potentially increase parent-child conflict and thus also undermine proper metabolic control. Parents felt that system appropriateness tapered off with the onset of adolescence, partly due to a potential sense of surveillance from the child's perspective that could fuel oppositional behavior. Parental suggestions for further developments included similar alerts of irregular insulin dosages and automatically generated dietary and insulin dosage advice.

Conclusions: User enthusiasm suggests that such systems might find a consumer market regardless of whether or not they ultimately improve health outcomes. Thus, more rigorous studies are warranted to inform guidelines for appropriate use. Potentially fruitful approaches include integrating such systems with theory-based parenting interventions and approaches that can aid in interpreting and responding to experiences of surveillance, virtual presence, and balances of power in e-mediated relationships.

(J Med Internet Res 2005;7(5):e57) doi:10.2196/jmir.7.5.e57

KEYWORDS

home blood glucose monitoring; diabetes; self-management; parent-child relations; wireless communication; mobile phones; pediatric psychology

Introduction

Parents of children with type 1 diabetes are often involved in aspects of their children's lives that they would ordinarily ignore [1]. Increased blood glucose monitoring is associated with improved glycemic control [2], which is essential in the prevention of serious, even life-threatening, complications of the disease. "Constant vigilance" is found to be a primary behavior strategy for parents coping with the responsibility of managing their child's disease [3]. They are faced with the difficult balancing act of helping their child learn and conform to the rigorous blood glucose monitoring (BGM) and regulatory regimens, while at the same time allowing their child the trust and independence necessary for developing their own sense of autonomy and coping skills [4].

This challenge changes in character as children grow into adolescence, and adolescents report parental worry and intrusive behavior as a major source of conflict [5]. Heightened levels of conflict and low levels of family cohesion and support are associated with poorer metabolic control [1,6], poorer quality of life [7], and poorer adherence among adolescents [3,8]. Thus, conflict reduction and management is a major concern in diabetes care. Facilitation of appropriate parent involvement is also crucial since less parental involvement in diabetes care has been associated with poorer diabetes outcomes [4]. Interventions designed to enhance BGM but that inadvertently exacerbate parent-child conflict or undermine parent involvement may thus do more long-term harm than good.

Information and communication technologies (ICT) show promise as a platform for facilitating evidence-based disease prevention and self-management interventions [9]. Much of this work has centered on Internet applications that facilitate information, decision support, and social support [10-12]. These types of applications may be enhanced by adaptation to mobile platforms. As cell phones increasingly approximate personal computer (PC) functionality with Internet access, they become a highly accessible platform for facilitating a wide range of health interventions, some of which appear particularly promising [13,14]. Cell phone–based text messaging has become a socially popular form of communication, particularly among adolescents [15]. Cell phone usage is comparatively higher in Norway (site for study) than other Western countries (eg, 96% of Norwegians vs 60% of US population), with Norwegians sending an average of 68 text messages per month [16]. However, since the dispersion of cell phones is expected to be considerably higher than PCs, both in Western and developing countries [17], the potentials for dispersion of disease prevention and management support through mobile devices are considerable. Still, mobile technologies coupled with specific behavioral health strategies have yet to be utilized effectively [14].

The working hypothesis that guided design of the prototype and concept tested in this study was as follows: automatic transfer of measures from the child’s blood glucose meter to the parent’s mobile phone could ease parental worries and tendencies to intrude in their children’s lives by unnecessary reminders and/or questions. Potentially, this could, in turn, aid in decreasing levels of conflict in parent-child interaction and thus also increase adherence. While research into the psychosocial dynamics of mobile communications among adolescent peers is increasing [15], we are unaware of any studies that address these issues in the context of family disease management. In light of the critical role parent-child interactions play in the monitoring and regulation of blood glucose, this issue warranted particular attention in the formative stages of these applications. As argued by others [18], a qualitative approach was considered most appropriate in this phase. Also, since these types of self-management technologies may well be used by consumers without the supervision of health care professionals, it is crucial to tap into the perspectives of potential users as early as possible.

The aim of this exploratory pilot study was twofold. First, we sought initial user insights into the appropriateness of the concept, feasibility of use in daily life, and desired system functionality. Second, we sought indications of relevant approaches for future developments of monitoring and messaging systems in disease self-management. In this formative stage of development, our concern was to better understand the psychosocial issues potentially involved in the use of this type of technology in families. This is useful as a basis for designing systems, as well as the process and outcome studies that are ultimately needed for determining appropriate roles for mobile ICT in family health management.

Methods

The prototype was developed at the Norwegian Centre for Telemedicine, a publicly owned non-profit national competence center. Using a Bluetooth connection, the prototype automatically transfers blood glucose readings from a blood glucose monitor (OneTouch Ultra from Lifescan) to a mobile phone after measurements are taken. The mobile phones (Nokia 7650) were programmed to automatically send the measurement results by means of the text messaging Short Message Service (SMS) to the parents’ mobile phone. The mobile phone sends the SMS without intervention from the user (in this case, the child) as long as the phone is within Bluetooth range (10 m) of the blood glucose monitor at the time of the blood glucose reading (Figure 1). When this range is exceeded, the blood glucose readings are sent in batches of five the next time the units are within range of each other.

Figure 1. Blood glucose readings are sent to parent's mobile phone

Prior to the project, three pilot users tested the prototype for a period of 3 months, during which improvements were made. Based on this, 17 prototypes were developed, 15 of which were provided to the participants in the project and 2 of which were kept as backup and reference.

Participants

Invitations to participate in the pilot study were sent to all 55 families of children with type 1 diabetes who, at the time, were being followed by the University Hospital of North Norway. The first 15 parent-child dyads (N = 30) who responded positively were accepted as participants (thus, we do not know how many nonresponders may have refused, had they been pursued). The group of children consisted of 11 boys and 4 girls, aged 9 to 15 years. The disproportionate number of boys may reflect a greater interest for technical trials among boys since gender distribution among the 55 approached parent-child dyads was fairly even. All of the children had received initial training courses for management of their diabetes, but they differed greatly in experience; duration of the disease ranged from 8 months to 6 years. For insulin injections, 13 reported using an insulin pen while 2 used an insulin pump. Before the trial, all of the parents and 11 of the children were frequent users of mobile phones, 3 of the children were nonfrequent users, and 1 child had not previously used a mobile phone.

Procedure and Instruments

All parents signed an informed consent form on behalf of themselves and their children, and the intervention and methods were approved by the Regional Ethics Committee. The children were provided with the prototype-enhanced mobile phones, while all parents used their existing mobile phones to receive the SMS. They were trained in use of the system during a routine hospital visit and used it for approximately four months between October 2003 and February 2004. The participants were requested to use the system a minimum of three times a day, but they were reimbursed for the equivalent of 10 messages per day (approximately US$1.40) regardless of use. At their own cost, they were free to use the phone for private purposes during the trial. A diabetic nurse at the hospital handled any questions from users, channelling technical problems to the project manager.

At the completion of the trial, all 15 children and their parents completed separate questionnaires about use and satisfaction. A semi-structured guide for parent interviews was designed to elicit experiences and ideas regarding the potential benefits and pitfalls and further system developments. The interview posed open-ended questions addressing three overriding issues: stress and coping, the parent/child relationship, and system functionality. Parents were encouraged to freely share experiences and thoughts. All but one of the interviews were conducted over the phone. The interviewer, who was unacquainted with the parents, interviewed those parents that were available during a limited period, that is, 9 mothers and 1 father of children ranging in age from 9-14 years. After interviewing 10 parents, no new information emerged ("data saturation" [19]), and further interviews were deemed unnecessary for our preliminary purposes. Of the 10 interviews, 1 was lost due to a faulty audio recorder.

Analysis and Presentation

The questionnaires were analyzed and reported as straightforward frequencies. For interviews, rigorous adherence to qualitative analysis procedures was deemed premature for our preliminary purposes, although efforts were made to comply with the basic principles of qualitative research [20]. Two co-authors (psychologists) independently read the transcripts, noting emerging themes and corresponding quotes that reoccurred across interviews thus allowing a broad range of possible interpretations and/or misunderstandings. These were then condensed to the nine themes presented as interview results. Presentations rely heavily on quotes from parents. These are edited to faithfully reflect what was said and meant, while at the same time ensuring readability [21].

Results

The families used the system as requested, on average 3 to 4 times daily, mostly when the children were at school or away from home. As indicated by questionnaire responses in Table 1, both children and parents reported that automatic transfer of blood glucose measures was definitely a good thing (80% and 93%, respectively), that living with diabetes was at least to some extent easier (73% and 80%, respectively), and that the system helped parents feel reassured (100%). While children were split as to whether they wanted to decide themselves about sending glucose measures rather than having them sent automatically (40% yes and 53% no), most parents (93%) did not think children should make the decision to send measures.

Table 1. Questionnaire responses regarding child and parent satisfaction

Questions to children (n = 15)	**Yes, definitely (%)**	**To some extent* (%)**	**No, not at all (%)**
Was it positive that your parents received your blood glucose measures?	12 (80)	3 (20)	0
Did living with diabetes become easier with the system?	2 (13)	9 (60)	4 (27)
Would you like to decide yourself whether the blood glucose measure should be sent or not?	6 (40)	Undecided 1 (7)	8 (53)
Questions to parents (n = 15)			
Was it positive to receive the blood glucose measures from your child?	14 (93)	1 (7)	
Did it become easier to manage your child's diabetes with the system?	6 (40)	6 (40)	3 (20)
Did receiving the blood glucose measure help you feel reassured?	14 (93)	1 (7)	0
Did receiving the blood glucose measures add to your worry?	1 (7)	1 (7)	13 (86)
Would you like for your child to decide whether the blood glucose measure should be sent or not?	1 (7)	Undecided 0	14 (93)

[*] For two questions, the middle category was formulated as "undecided."

The qualitative interviews among parents provided additional information and shed light on questionnaire responses. The nine interviews resulted in the following themes that reoccurred across interviews: (1) sense of security and reassurance, (2) nagging and scolding, (3) control, responsibility, and independence, (4) surveillance and opposition, (5) learning and age-phased appropriateness, (6) focus upon illness, (7) if it's not automatic, forget it, (8) system type and functionality, and (9) it depends on how you use it.

Sense of Security and Reassurance

The parents almost unanimously expressed appreciation for the security of knowing whether or not their children had measured their blood glucose and that they could intervene immediately if the values were alarming. Several expressed a wish to continue use. This was particularly apparent for newly diagnosed and younger children and when children were away from home or when the parent was traveling.

> *Both she and we slept better…. Without the system you can go around for hours saying to yourself, "What's happened; What hasn't happened; Should we call? No, maybe she's measured." It's stupid to call and hassle her, you know. And then she'll come home and hasn't measured...while [during the trial] if we didn't get an SMS we could just call right away.*

When parents did not receive an SMS measure as expected, various interpretations were triggered. These could range from technical failure, to their child's forgetfulness, to passiveness bordering on conscious sabotage on the part of their child. Parents describe this as a stress factor that had always been there on one level but was now dispersed throughout the day, as illustrated by two parents:

> *You didn't [before the system] get stressed when you didn't know she hadn't measured. Now, the minute you expect it to come and it doesn't, you start worrying, which maybe is negative, but not worse than you can live with.*
>
> *I went around checking my phone all the time. I didn't go with it in my pocket, but I checked my purse pretty often. That was pretty…what shall I say…but I guess it was tolerable.*

This increased vigilance may help explain the feeling of one parent who responded in the questionnaire that the system added to worry. In general, both questionnaire and interview responses suggested that parents were more inclined to view the SMS message as reassuring.

Nagging and Scolding

Some parents reported that their nagging increased, while others reported a decrease. This appeared to depend upon whether children monitored their blood glucose regularly. For example, one parent who received the updates on her child's blood sugar regularly reported that

> *I feel safer, so I pester him less.*

In contrast, another parent whose child did not monitor regularly reported that

> *Maybe I nag and scold more after seeing how negligent he can be…. I call him at school and say, "Why haven't you measured as we agreed?" which I'm sure annoys him…. Yes, I'm worse after we started with this system…. But, as I tell him, it's for his own good…. He can avoid calls from me by remembering to measure.*

Several parents described making conscientious efforts to remind their children in ways that would not be perceived as nagging, for example, by finding other excuses to call and thus triggering measures indirectly.

Control, Responsibility, and Independence

Regardless of system use, parents described their struggle to find a balance between the control they felt necessary to ensure the health of their child, while at the same time allowing for the child to develop their own sense of responsibility and independence. System use appeared to tip the balance in both directions. While one informant thought "…it clearly placed more responsibility with the child," another said, "It's obvious that we're the ones that have increased our responsibility."

Those who felt that their children's responsibility was enhanced through use argued that they were provided with a safe framework within which the child could learn to make their own decisions with parental guidance.

> *He can feel secure knowing that others are part of this and see, but that he can figure things out and do his own thing anyway.*

Another noted

> *There were a couple of times I forgot my phone, so he called me wondering whether or not I'd seen his measures, and if he should take something. So I think it worked well.*

Those who felt that the parents' responsibility was increased sensed that their child could be pacified. One parent described her child's likely thought process:

> *I don't need to bother [measuring], since Mom will start calling to hassle me pretty soon, and then I can do it.*

This parent was nevertheless uncertain about the degree to which the system was to blame, since her child was "sick and tired" of her illness long before the system was introduced. As she put it,

> *It could well be that the system undermines children's responsibility for their illness, but during the period we used it…in the situation she was in…it was a help for us.*

One parent alluded to a distant form of presence.

> *She always takes more responsibility when she's alone, than when she's with us.*

The meaning the system had for this child in this respect was, however, unclear for the parent.

> *She knew it happened [that parents saw her measures], but it happened so automatically, I don't think she thought about it.... At least it didn't pacify her.*

Surveillance and Opposition

One concern is that the system may create a negative sense of surveillance and thus fuel oppositional behavior. For older children, and those with preexisting levels of conflict, the system appeared to represent an additional source of tension.

This parent related episodes before the trial where their boy had consciously deceived them:

> *Of course we've wondered if he'd measure a buddy rather than get hassled by us. So who knows? But I don't think he's done it [sent a buddy's blood glucose reading].*

Another parent alluded to the possibility of her child consciously refraining from measuring when she knew her measures would trigger a reaction from the parent.

> *If she doesn't want me to know she's high or low, she'll refrain from measuring. She'd know she'd get a message or phone from me, but she could elude me anyway by not taking the phone or just turning it off. So she escapes me regardless.*

Parents differed in the degree to which they reported discussing this issue with their child.

Learning and Age-Phased Appropriateness

Parents indicated that the potential of the system to facilitate knowledge and skills about BGM and regulation was greatest at the onset of disease. Parents were also fairly consistent in their view that the appropriateness of the system tapered off with the onset of adolescence. One indicated that the system would have been particularly useful during the period when the child was newly diagnosed and in preschool.

However, parents differed in their perceptions of when the child became too old for the system. A parent with a 14-year-old stated

> *He hasn't really taken much responsibility himself…. I feel he's a little too young.... He's quite good at following up, but I still prefer to be in control…so this [system] has been positive.*

Another parent with a younger child said

> *The messages are positive for us [parents]…for our reassurance. But I don't feel it's right for him who's 11 years old and should manage himself. It gets too*

> *controlling, I think. But he'll have to speak for himself.*

Several parents appeared uncertain about age-appropriate expectations of their children, suggesting that it may be helpful to assist families in sharing experiences about how other parents deal with this issue.

Focus on Illness

Can system use exacerbate or reduce the dominance of illness as the focus of parent-child interaction? The parents who explicitly commented on this tended to think that their focus on illness was the same or less but that this depended on the way it was used.

One parent didn't think it made any difference since they had no choice but to constantly focus on the illness:

> *We have to hang over her all the time anyway or else things would fall apart. The system just helped us to follow up, it didn't make us more focused [on illness] than we were already.*

Another appreciated the system because it allowed focusing on other issues in their relationship:

> *It was wonderful! We started talking about other things—you know, mother-child dialogue—not blood sugar and setting insulin that had been the main content of our communication for a period, which was bad. It was absolutely fantastic, especially in the beginning [of the trial] when we noticed it so clearly. And I think he experienced it too.*

This parent underlined the importance of not responding to every SMS, which could trigger an unnecessary focus on disease.

If It's Not Automatic, Forget It

Parents were adamant about the measures needing to be transferred automatically, and they used expressions like "Alpha Omega," "extremely important," and "very thankful for it" to underline their view. They had no faith in a concept requiring a conscious effort on the part of their kids to trigger the transfer of measures to parents, arguing that they had enough to remember as it was. Also, although schools made exceptions to the mobile phone ban in class for these children, it was appreciated that the phones could remain unhandled and hidden in their bags.

Suggested Functionality

Suggestions for improving functionality included automatically generated dietary and insulin dosage advice. One parent argued for developing the same concept for the insulin unit, but only for transmitting irregular doses. One time his child had taken 20 insulin units in response to high blood sugar, but a few minutes later, he forgot he had done it and set 20 more, sending him into a coma.

> *If we'd known, we could have prevented him getting so traumatized…. These cases are about life and death.*

None of the parents thought the system would be useful in interaction with their health care provider, except if automatically generated historical graphics could be transferred to their provider in preparation for their ordinary quarterly checkups.

Depends on How You Use it

In varying ways, parents indicated that it was more the way they used the system than the system itself that was important. For example, one parent underlined how routines for system use could limit the dominance of the system in interaction with her child.

> *I think it's very important that you have it on all the time, even though it isn't economically smart. You get a message, and you just think "OK." Of course, if one talks about every single message, or sends a response, that's different. It [degree of interference by the system] has to do with how one uses the information and the system.*

Discussion

The ultimate objective of the system tested is to improve health and quality of life by supporting daily blood glucose monitoring and regulation processes in families. In this study, we sought an understanding of the tested system's potential role, feasibility of use in daily life, desired functionality, and approaches that may be relevant for future developments and research.

Potentials and Concerns

Knowledge of their children's blood sugar status eased parental concerns. This appeared to lessen parental intrusions in those cases where children measured regularly. For those who measured irregularly, parental reminders (eg, "nagging") appeared to increase. While an increase in parent reminders could be considered a positive outcome in light of studies underlining the importance of parent involvement in monitoring [2,4], it must be acknowledged that this can potentially increase parent-child conflict, particularly among adolescents and those with existing tensions. Discrepancies between parents and children as to whether or not the measures should be automatic underscore the need for more in-depth inclusion of children's perspectives in future studies.

Some parental responses could be construed to suggest that receiving automatic measures throughout the day could be experienced as an invasion in their lives that they would rather be without, but that they could not or would not admit it ("…but I guess it's tolerable," "...but it's not worse than you can live with"). Obviously, the disease itself is an unwanted "invasion" both in the lives of their children and themselves. Presumably, parents feel morally obligated to tolerate whatever is necessary to ensure the health of their child, thus denying themselves more "egoistical" reactions. If the virtual presence of children (through expectations of regular SMS messages demanding attention) exacerbates the existing burden of care among parents, it could inadvertently undermine their long-term involvement and/or fuel parent-child conflict. The concept of "constant vigilance" [3] associated with the burden of diabetes care may be useful in further exploring how system use may add to or relieve the burden of care. This should also include other caretakers' (eg, family, pre-school and school staff) and how system use may influence their willingness to accept responsibility for care.

Concept Suitability and System Functionality

Parents suggested that the system might be most beneficial for those who are younger and newly diagnosed. Our pilot users were aged 9 to 15, having lived with the diagnosis of diabetes from 8 months to 6 years, thus limiting our ability to shed light on the appropriateness for younger (eg, pre-school) children. One study found that mothers with pre-school children claim lack of confidence in other caregivers as a major reason for not having placed their children in day care [3]. The possibility that system use could support involvement of other caregivers in relieving the burden of care for parents of younger children may be worth pursuing.

Suggested functionality included a similar concept for alerting parents of irregular insulin dosages, as well as automatically generated dietary and insulin dosage advice. Combinations of information, decision support, and social support are core elements of existing Internet-based self-management applications [10,12]. Merging these types of applications with mobile monitoring and messaging functionality may enable relevant support whenever and wherever needed [14]. By building on daily life technologies (eg, blood glucose meters and mobile phones) the threshold for use should be relatively low, as our study suggests.

Historical graphs of measures were also suggested. We are currently working to enable such graphs, which may be used in conjunction with follow-up consultations. This can enhance both system evaluation and give health care providers an opportunity to supervise system use.

Future Developments and Research

The enthusiasm expressed by interviewed parents is worth noting, despite the weaknesses of our sample (see below). It may very well be that such systems find a consumer market regardless of whether or not they ultimately succeed in improving health outcomes. Parents of children with chronic health problems can face "life and death" issues, as one parent put it. Some may find that relief from some of the worries associated with chronically ill children is sufficient motivation for system use, even if the way they use it inadvertently exacerbates family tensions, and possibly also blood glucose control. This provides all the more reason to take the technology seriously and pursue more rigorous approaches to development and evaluation. As is typical of eHealth innovations [9], we are nowhere near reliable answers to very basic questions such as for whom, under what conditions, and how monitoring and messaging systems may relate to health outcomes.

Broader perspectives, two of which are suggested here, may be helpful in guiding future developments and evaluations. First, it may be useful to view these systems as an element of parenting, rather than a simple monitoring and messaging device. As illustrated by the parents in our study and others [1,4,8], the ordinary challenges of parenting are compounded when children have a chronic disease. Our findings suggest that system use may merely intensify ongoing parent-child interaction patterns—for better or worse—unless incorporated into conscientious efforts to improve such patterns. There is a growing body of evidence that behavioral family intervention approaches are effective at reducing parent-child conflict, increasing parents' competence and monitoring skills, and increasing child compliance, including families whose children have chronic health problems[22,23]. As such, evidence-based approaches to parenting practices may be useful in informing both system design choices and guidelines for use. These approaches incorporate various types of counseling and social support, which we believe our pilot users could have benefited from in conjunction with system use. As technologies such as those described in this paper become more available, it would seem important for health care providers to introduce them as an element in more holistic, empirically supported approaches to family interaction, coping skills, and child-rearing practices in disease self-management.

Another perspective suggested for guiding further developments relates to the dilemmas faced when introducing technology into interpersonal (eg, child-parent, doctor-patient) relationships. The parents in our study voiced some of these dilemmas. On the one hand, they valued the security of knowing that they could intervene if their child's blood glucose values were alarming. At the same time, they expressed concerns that the system could inadvertently undermine the child's independence and confidence in their own coping skills and that a sense of surveillance might fuel oppositional behavior, particularly among adolescents. One potentially fruitful approach to such dilemmas is outlined in Spears and Lea's SIDE (Social Identity and Deindividuation) model [24]. This model of computer-mediated communication suggests how we might anticipate and interpret experiences of surveillance, virtual presence, and balances of power in e-mediated interpersonal relationships. They remind us of the need to critically examine assumptions about the role of technology and the implications of use.

Study Limitations

The limitations of this study (beyond our exploratory and formative purposes) are obvious. Our sample is quite possibly biased since it is not random. Those who responded first to our invitation may, for example, be particularly technology savvy. The experiences of the five parents (one-third of pilot users) who were not interviewed may have deviated consistently from our respondents, although we have no reason to believe this was the case. However, even if the last five were all negative or all positive, our conclusions would remain the same.

Conclusion

The mobile monitoring and messaging system tested is feasible to use in daily life. While system use eases parental worries, and as such shows potential as an aid in disease self-management, it is unclear if or how this will improve child-parent interactions and/or health outcomes. User enthusiasm suggests that such systems might find a consumer market regardless of whether or not they ultimately improve health outcomes. Thus, more rigorous studies are warranted to clarify these issues and inform appropriate use. Potentially fruitful approaches include integrating such systems with theory-based parenting interventions and approaches that can aid in interpreting and responding to experiences of surveillance, virtual presence, and balances of power in e-mediated relationships.

Acknowledgments

This study received financial support from the Norwegian Directorate of Health and Social Affairs and from the Health and Rehabilitation Fund. We thank the personnel at the Department of Paediatrics at the University Hospital of North-Norway and the families who generously offered their time and experiences.

Conflicts of Interest

None declared.

References

1. Wysocki T. Associations among teen-parent relationships, metabolic control, and adjustment to diabetes in adolescents. J Pediatr Psychol 1993 Aug;18(4):441-452. [Medline: 94015753]
2. Levine BS, Anderson BJ, Butler DA, Antisdel JE, Brackett J, Laffel LM. Predictors of glycemic control and short-term adverse outcomes in youth with type 1 diabetes. J Pediatr 2001 Aug;139(2):197-203. [Medline: 21380327] [doi: 10.1067/mpd.2001.116283]
3. Sullivan-Bolyai S, Deatrick J, Gruppuso P, Tamborlane W, Grey M. Constant vigilance: mothers' work parenting young children with type 1 diabetes. J Pediatr Nurs 2003 Feb;18(1):21-29. [doi: 10.1053/jpdn.2003.4] [Medline: 22497485]
4. Anderson BJ, Brackett J, Ho J, Laffel LM. An office-based intervention to maintain parent-adolescent teamwork in diabetes management. Impact on parent involvement, family conflict, and subsequent glycemic control. Diabetes Care 1999 May;22(5):713-721 [FREE Full text] [Medline: 99265075]
5. Weinger K, O'donnell KA, Ritholz MD. Adolescent views of diabetes-related parent conflict and support: a focus group analysis. J Adolesc Health 2001 Nov;29(5):330-336. [Medline: 21546296] [doi: 10.1016/S1054-139X(01)00270-1]
6. Auslander WF, Anderson BJ, Bubb J, Jung KC, Santiago JV. Risk factors to health in diabetic children: a prospective study from diagnosis. Health Soc Work 1990 May;15(2):133-142. [Medline: 90306816]
7. Laffel LMB, Connell A, Vangsness L, Goebel-Fabbri A, Mansfield A, Anderson BJ. General quality of life in youth with type 1 diabetes: relationship to patient management and diabetes-specific family conflict. Diabetes Care 2003 Nov;26(11):3067-3073 [FREE Full text] [Medline: 22939618]
8. Cohen DM, Lumley MA, Naar-King S, Partridge T, Cakan N. Child behavior problems and family functioning as predictors of adherence and glycemic control in economically disadvantaged children with type 1 diabetes: a prospective study. J Pediatr Psychol 2004 Apr;29(3):171-184. [Medline: 15131135] [doi: 10.1093/jpepsy/jsh019]
9. Jadad AR, Delamothe T. What next for electronic communication and health care? BMJ 2004 May 15;328(7449):1143-1144 [FREE Full text] [Medline: 15142890] [doi: 10.1136/bmj.328.7449.1143]
10. 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 19;4:CD004274. [Medline: 16235356] [doi: doi:10.1002/14651858.CD004274.pub4]
11. Mckay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334 [FREE Full text] [Medline: 21365601]
12. Mckay HG, Glasgow RE, Feil EG, Boles SM, Barrera MJ. Internet-based diabetes self-management and support: initial outcomes from the diabetes network project. Rehabil Psychol 2002;47(1):31-48. [doi: 10.1037/0090-5550.47.1.31]
13. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261. [Medline: 16046689] [doi: 10.1136/tc.2005.011577]
14. Franklin V, Waller A, Pagliari C, Greene S. "Sweet Talk": text messaging support for intensive insulin therapy for young people with diabetes. Diabetes Technol Ther 2003;5(6):991-996. [Medline: 14709202] [doi: 10.1089/152091503322641042]
15. ; Norwegian Post and Telecommunication Authorities (NPT). Telestatistics 2004. URL: http://www.npt.no/pt_internet/venstremeny/publikasjoner/telestatistikk/statistikk2004/halvaar.pdf [accessed 2005 Aug 5]
16. Grinter R, Eldridge M. "y do tngrs luv 2 txt msg?". Presented at: ECSCW '01; 2001; Dordech, Netherlands.
17. ; Technology and development: the real digital divide. The Economist 2005 Mar;374(8417):12-18.
18. Pagliari C, Gregor P, Mcgillivray S, et al. Literature review and conceptual map of the area of E-health. NHS Service Development and Organisation Programme. URL: http://www.sdo.lshtm.ac.uk/ehealth.htm#pagliari [accessed 2005 Nov 3]
19. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 8;320(7227):114-116 [FREE Full text] [PMC: 10625273] [Medline: 20090738] [doi: 10.1136/bmj.320.7227.114]

20. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [Medline: 21405951] [doi: 10.1016/S0140-6736(01)05627-6]
21. Sandelowski M. The use of quotes in qualitative research. Res Nurs Health 1994 Dec;17(6):479-482. [Medline: 95063162]
22. Taylor TK, Biglan A. Behavioral family interventions for improving child-rearing: a review of the literature for clinicians and policy makers. Clin Child Fam Psychol Rev 1998 Mar;1(1):41-60. [Medline: 21222175] [doi: 10.1023/A:1021848315541]
23. Grey M. Interventions for children with diabetes and their families. Annu Rev Nurs Res 2000;18:149-170. [Medline: 20375508]
24. Spears R, Lea M. Panacea or Panopticon? The hidden power in computer-mediated communication. Communic Res 1994;21:427-459.

submitted 01.09.05; peer-reviewed by V Franklin, S Sullivan-Bolyai; comments to author 31.10.05; revised version received 09.11.05; accepted 31.10.05; published 21.11.05.

Please cite as:
Gammon D, Årsand E, Walseth OA, Andersson N, Jenssen M, Taylor T
Parent-Child Interaction Using a Mobile and Wireless System for Blood Glucose Monitoring
J Med Internet Res 2005;7(5):e57
URL: http://www.jmir.org/2005/5/e57/
doi:10.2196/jmir.7.5.e57
PMID:16403721

Original Paper

Website Quality Indicators for Consumers

Kathleen M Griffiths[1], PhD; Helen Christensen[2], PhD

[1]Depression & Anxiety Consumer Research Unit, Centre for Mental Health Research, The Australian National University, Canberra, Australia
[2]Centre for Mental Health Research, The Australian National University, Canberra, Australia

Corresponding Author:
Kathleen M Griffiths, PhD
Centre for Mental Health Research
The Australian National University
Canberra 0200
Australia
Phone: +61 2 6125 9723
Fax: +61 2 6125 0733
Email: kathy.griffiths@anu.edu.au

Abstract

Background: The rating tool DISCERN was designed for use by consumers without content expertise to evaluate the quality of health information. There is some evidence that DISCERN may be a valid indicator of evidence-based website quality when applied by health professionals. However, it is not known if the tool is a valid measure of evidence-based quality when used by consumers. Since it is a lengthy instrument requiring training in its use, DISCERN may prove impractical for use by the typical consumer. It is therefore important to explore the validity of other simpler potential indicators of site quality such as Google PageRank.

Objective: This study aimed to determine (1) whether the instrument DISCERN is a valid indicator of evidence-based Web content quality for consumers without specific mental health training, and (2) whether Google PageRank is an indicator of website content quality as measured by an evidence-based gold standard.

Methods: This was a cross-sectional survey of depression websites using consumer and health professional raters. The main outcome measures were (1) site characteristics, (2) evidence-based quality of content as measured by evidence-based depression guidelines, (3) DISCERN scores, (4) Google PageRank, and (5) user satisfaction.

Results: There was a significant association between evidence-based quality ratings and average DISCERN ratings both for consumers ($r = 0.62$, $P = .001$) and health professionals ($r = 0.80$, $P < .001$). Consumer and health professional DISCERN ratings were significantly correlated ($r = 0.77$, $P < .001$). The evidence-based quality score correlated with Google PageRank ($r = 0.59$, $P = .002$). However, the correlation between DISCERN scores and user satisfaction was higher than the correlation between Google PageRank and user satisfaction.

Conclusions: DISCERN has potential as an indicator of content quality when used either by experts or by consumers. Google PageRank shows some promise as an automatic indicator of quality.

(J Med Internet Res 2005;7(5):e55) doi:10.2196/jmir.7.5.e55

KEYWORDS

Depressive disorder; medical informatics; consumer participation; evaluation studies

Introduction

There has been widespread concern about the quality of Web-based health information designed for consumers [1]. In response to this, a number of initiatives have been developed to assist consumers in locating quality health information on the Web. These include the use of quality labels based on compliance with codes of conduct (eg, HON code), portals that provide a gateway to websites of “high quality” (eg, OMNI), and rating tools designed for consumer use [2].

One rating tool that shows particular promise is DISCERN, an instrument designed for use by consumers and providers “to judge the quality of written information about treatment choices” [3, p. 106]. This tool is widely recommended and used by authoritative sources for the evaluation of websites. However, it has not yet been convincingly established that DISCERN, particularly when used by consumers, is a valid indicator of quality when compared against an evidence-based gold standard.

Three studies have investigated the relationship between the DISCERN ratings of experts and “scientific” quality [4-6]. Two of the studies reported a significant association between DISCERN and scientific accuracy [4,5], but the authors of the third study found “no clear relationship between methodological (DISCERN) and medical-scientific quality” [6]. Unfortunately, except for the Griffiths and Christensen study [4], it is unclear if the standard against which the DISCERN ratings were compared was based on systematic reviews of the evidence. Moreover, in each study, ratings were made by health professionals. To date, to our knowledge, there has been no assessment of the validity of DISCERN as measured by an evidence-based gold standard when used by consumers without technical expertise.

Although the developers trialed DISCERN with self-help group users in a research context, it is a lengthy instrument, and it is not clear if individual consumers would use DISCERN in practice. Other simpler potential indicators of site quality include those based on the link structure of the World Wide Web. For example, Google PageRank is an automatically computed measure of the importance of a website based on the number and importance of Web pages linking to it. However, there is little evidence as to the validity of link structure as an indicator of quality.

The current study, therefore, sought to determine the following for depression information websites: (1) whether DISCERN is a valid indicator of evidence-based content quality for consumers without specific mental health training, and (2) whether Google PageRank is an indicator of content quality. Depression websites were selected because depression is a leading cause of disease burden [7], there is a high level of unmet need among people with depression [8], depression is one of the most common reasons consumers access health information on the Internet [9], and evidence-based guidelines for depression management are available.

Methods

Website Selection

Twenty-four depression websites with a Google PageRank were selected from the Depression Directory of the DMOZ Open Directory Project website (n = 127). Three sites for each Google PageRank score within the range 0 to 7 were randomly selected using the R Project statistical package [10] to ensure a range of sites were represented. Each of the selected sites was then captured (in April 2003) and electronically archived for assessment using purpose built software. External links from these sites were excluded.

Site Assessment

Sites were rated online by four researchers/health professionals with expertise in depression and three consumers with a history of depression but no professional experience in mental health or research. Two of the health professionals (KG, HC) rated the site using an evidence-based gold standard. They also rated the characteristics of each website. The other two health professionals (AJ, RK) and the three consumers rated the sites using the DISCERN measures. All raters provided satisfaction measures for each site. Sites were presented in a different random order for each rater, and each rater was supplied with a pro forma rating sheet. The consumer raters were employed as casual research assistants during the study.

Site Characteristics

Each site was rated on a range of attributes, including ownership structure, scope, editorial arrangement, and legal policies (Table 1).

Evidence-Based Guideline Score

Evidence-based quality was assessed using the depression guidelines produced by the Centre for Evidence Based Mental Health (CEBMH) at Oxford [11]. The guideline score was the number of CEBMH items (maximum 20) correctly endorsed by the website [4]. In the current study, the correlation between evidence-based guideline scores for the two health professional raters was 0.94 ($P < .001$). An average guideline score was therefore computed for the two raters.

DISCERN Scores

The DISCERN instrument comprises 15 items (each rated from 1 to 5) and an additional “overall quality” item (rated 1 to 5) [3,12]. Raters in the current study were informed that the DISCERN questionnaire was designed to assess the quality of information about medical treatments and that “In this study we are focusing on the quality of web sites related to the treatment of depression.” Each rater was provided with the DISCERN instrument, which includes hints for rating each item, and the DISCERN handbook, which contains detailed information about the scoring of DISCERN items. Items in DISCERN include questions about the reliability of the publication (eg, are information sources specified, is it clear where these information sources were produced, degree to which the discussion is balanced) and the quality of information on treatments (eg, description of the mechanism, benefits, risks of possible choices and inclusion of multiple treatment options).

Previous research has demonstrated acceptable inter-rater agreement on individual items of the instrument when used by expert health professionals and “fair” agreement among consumers [3]. The original version of the test used the overall quality score as the measure of quality. However, subsequently, a number of studies employing DISCERN have used a measure of quality based on a total DISCERN score derived by cumulating scores across the first 15 DISCERN items (minimum score = 15; maximum = 75) (eg, [4,5,13,14]). This measure shows acceptable inter-rater agreement ($r = 0.88$ [4], r = 0.82 [14]) and has been reported to correlate with the overall quality rating ($r = 0.8$ [14]). In the current study, the correlation between the total DISCERN score and the overall quality item score was 0.91 for consumers and 0.92 for experts. The DISCERN results reported in the primary analyses are therefore confined to the total DISCERN measure.

The correlation between the DISCERN ratings for the two health professionals was 0.86 ($P < .001$). Intercorrelations between DISCERN ratings for the three consumer raters were 0.78, 0.77, and 0.68 ($P < .001$). An average score was therefore computed for the health professionals and the consumers. The DISCERN ratings for the health professionals and consumers were significantly correlated ($r = 0.77$, $P < .001$). A paired t test demonstrated that mean DISCERN scores for the two types of rater did not differ significantly across the 24 websites (t_{23} = 0.64, *P* = .53).

Satisfaction

Website satisfaction was measured using a series of 9 items developed for the purpose of the study. Items included questions about the target website's perceived usefulness, relevance to people with depression, trustworthiness, author knowledge, esthetics, and whether the site could be easily understood, easily navigated, and would be recommended. A total satisfaction score was calculated by computing the total number of satisfaction items endorsed by the rater (minimum 0, maximum 9). The correlation between satisfaction ratings for the two evidence-based guideline health professional raters was 0.86 (*P* < .001) and for the two DISCERN health professional raters was 0.83. Intercorrelations between satisfaction measures for the three consumer raters were significant in two of the three cases (rater 1 vs 2: *r* = 0.60, *P* = .002; rater 2 vs 3: *r* = 0.58, *P* = .003; rater 1 vs 3: *r* = 0.26, *P* = .22). Therefore, although the satisfaction measure for the evidence-based guideline health professional raters was based on their average score, and an average score was also computed for the DISCERN health professional raters, the satisfaction measures for the three consumer raters were treated separately.

Google ToolBar PageRank

Google PageRank is employed by the Google search engine as a measure of the "importance" of a Web page. These PageRank values can range from 0 to 10, with higher values indicating greater importance. PageRanks are based on an iterative algorithm developed by Google founders Brin and Page [15] that takes into account the number and importance of pages which link to a website. The importance of pages linking to a site is assessed according to the number and importance of sites linking to those pages. The PageRank score on the Google toolbar is a transformed function (conjectured to be logarithmic or distributional) of a raw Google PageRank score. The latter are very small positive numbers which sum to 1.0 over the entire Web and are known to be power-law distributed [16]. Google PageRank differs from the ranking order in Google search results in that PageRank is query independent, whereas the ranking order in Google search results takes into account many other variables, such as frequency of occurrence of search terms on a page, anchor text used to link to sites, and a large number of other tuning variables not disclosed by the company, as well as PageRank.

The Google PageRank for each site was obtained by downloading the Google toolbar and recording the integer number attached to the toolbar. The lowest and highest identified page ranks in the DMOZ depression directory were 0 and 7, respectively.

Analyses

Intercorrelations between evidence-based scores, DISCERN, and overall satisfaction were computed using Pearson r tests. (Note that when these analyses were recomputed using non-parametric Spearman rho tests, similar patterns of results were observed.) Site quality was assessed as a function of site characteristic using independent *t* tests (with Levene's correction in the case of unequal variances). Differences between evidence-based scores as a function of individual satisfaction items were analyzed separately using independent *t* tests except that no analysis was performed for items for which the sample sizes in a cell were very small (less than 6 sites). Multiple independent *t* tests were also used in analyzing the effects of site characteristics and for individual satisfaction items because the data were not amenable to an overall multivariate analysis such as a multiple regression or a MANOVA followed by contrasts corrected for multiple comparisons. For example, there were insufficient websites given the number of independent predictors to apply multiple regression to the data. The probability values cited in the results tables and text therefore refer to error rate per comparison. Given that a large number of comparisons were conducted in this study, the chance of reporting one or more spuriously significant results is high. For this reason, patterns of results, rather than isolated findings, are emphasized in reporting and interpreting the study results, particularly with respect to the satisfaction items. With the exception of tests of the significance of differences between dependent correlations, which were carried out using the SISA online calculator [17], all analyses were carried out using SPSS version 13.0 [18].

Results

Site Characteristics

Site characteristics are summarized in Table 1. Site ownership was distributed relatively evenly between individuals and organizations. Only a minority of the sites had an editorial board, and a health professional was involved in fewer than 40% of the sites. The majority of the sites were focused specifically on the topic of depression as might be expected from sites selected from a depression directory, although one-third contained more general mental health or health content. Over 40% of the sites promoted some type of product or service. Just under one-third of the sites collected personal information, and one-quarter required registration in order to obtain all of the site's information. One-third of the sites did not publish a privacy policy. Surprisingly, over 40% failed to include a disclaimer (eg, a statement that the website was not intended as a substitute for medical advice).

Level of Quality and Satisfaction

Overall, the mean evidence-based score was low (3.6, SD = 3.9), and the mean DISCERN ratings for both the health professional and consumer raters fell in the poor to average range (health professionals: mean = 37.8, SD = 17.0; consumers: mean =36.3, SD = 10.6). Mean satisfaction scores were low for the evidence-based raters (mean = 2.8, SD = 2.1), were average for the health professional DISCERN raters (mean = 4.3, SD = 2.6), and average for the consumer DISCERN raters (rater 1: mean = 6.1, SD = 2.3; rater 2: mean = 4.7, SD = 3.0; rater 3: mean = 4.4, SD = 3.7).

Table 1. Site characteristics and evidence-based quality scores

Site Characteristic		**Number of Sites (%)**	**Mean (SD) Evidence-Based Guideline Score (max = 20)**
Ownership structure	Individual	13 (54.2%)	1.5 (2.1)
	Organization[*]	11 (45.8%)	6.1 (4.2)
			$t_{14.2}$[**] = −3.36, $P = .005$
Editorial board	Yes	6 (25%)	7.7 (3.2)
	No	18 (75%)	2.2 (3.1)
			$t_{22} = -3.67$, $P = .001$
Scope [†]	Depression specific	15 (62.5%)	2.4 (3.5)
	Broad scope	8 (33.3%)	6.2 (3.7)
			$t_{21} = -2.41$, $P = .03$
Health professional involved	Yes	9 (37.5%)	7.2 (3.0)
	No	15 (62.5%)	1.4 (2.6)
			$t_{22} = -4.94$, $P < .001$
Promotion of products/services	Yes	10 (41.7%)	4.1 (4.2)
	No	14 (58.3%)	3.2 (3.8)
			$t_{22} = -0.54$, $P = .596$
Privacy policy	Yes	9 (37.5%)	6.4 (3.6)
	No	15 (62.5%)	1.9 (3.1)
			$t_{22} = -3.23$, $P = .004$
Disclaimer	Yes	10 (58.3%)	7.0 (3.7)
	No	14 (41.7%)	1.2 (1.7)
			$t_{11.7}$[**] = −4.64, $P = .001$
Feedback mechanism	Yes	22 (91.7%)	N/A[‡]
	No	2 (8.3%)	N/A[‡]
Register for all information	Yes	6 (25%)	6.6 (4.0)
	No	18 (75%)	2.6 (3.4)
			$t_{22} = -2.38$, $P = .03$
Collect personal information	Yes	7 (29.2%)	4.6 (4.1)
	No	17 (70.8%)	3.1 (3.9)
			$t_{22} = -0.85$, $P = .41$
All sites		24	3.6 (3.9)

[*] Commercial, consumer, or other organized group

[**] Levene's correction applied

[†] One site not depression related

[‡] Not analyzed due to small sample size

Association Between Evidence-Based Quality and the Potential Indicators of Quality

DISCERN

There was a strong correlation between the average evidence-based score and the average DISCERN ratings for the health professionals ($r = 0.80$, $P < .001$) and a moderately high correlation for consumers ($r = 0.62$, $P = .002$). For health professionals, intercorrelations between DISCERN ratings and evidence-based scores for each of the items considered separately ranged from 0.37 ($P = .08$) for Item 5 (Is it clear when the information used or reported in the publication was produced?) to 0.88 for Item 3 ($P < .001$) (Is it relevant? eg, Does the publication address the questions readers might ask and are the treatment recommendations realistic or appropriate?). For consumers, this range was 0.18 ($P = .40$) for Item 5 to 0.68 ($P < .001$) for Item 3.

Google PageRank

There was a moderate correlation between the evidence-based guideline score and Google PageRank ($r = 0.59$, $P = .002$). The size of this correlation was almost the same as that between the consumer DISCERN ratings and evidence-based scores.

Site Characteristics

Table 1 shows the evidence-based guideline scores as a function of site characteristics. Evidence-based quality was significantly higher for organizations, sites with an editorial board, sites with broad health content, and sites involving a health professional than for their counterparts. Similarly, sites which posted a privacy policy, sites which included a disclaimer, and sites requiring registration to obtain all information were of significantly higher evidence-based quality. There was no significant difference in evidence-based guideline scores for sites that promoted products or services or that collected personal information on visitors.

Associations Between Quality Measures and Satisfaction

Evidence-based ratings were significantly correlated with overall rater satisfaction ($r = 0.85$, $P < .05$). Sites that were judged by consumers to have useful treatment information, to describe what a consumer might wish to know about depression, to be trustworthy, and to be written by people who knew about depression showed better evidence-based quality, at least for 2 of the 3 consumers (Table 2). There were no significant differences in evidence-based scores for consumers as a function of the judged attractiveness of the site or whether they would recommend it to someone else. Sites judged by health professional raters as useful, relevant, written by a knowledgeable author, and worthy of recommendation were of higher evidence-based quality. There were no significant differences in evidence-based scores as a function of whether the site was judged by health professionals to be navigable. The pattern of findings for the health professionals who provided evidence-based ratings was similar to the pattern of findings for health professionals who conducted DISCERN ratings.

Table 2. Mean DISCERN scores for consumers and mean evidence-based and DISCERN scores for health professionals, as a function of individual satisfaction items

	Consumer Raters			**Health Professional Raters**			
				Evidence-Based		**DISCERN**	
Item	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 1**	**Rater 2**	**Rater 1**	**Rater 2**
Useful treatment							
Yes	6.45 (n = 10)	6.17 (n = 9)	4.45 (n = 11)	8.93 (n = 7)	8.93 (n = 7)	7.50 (n = 9)	5.82 (n = 11)
No	1.54 (n = 14)	2.14 (n = 7)	2.85 (n = 13)	1.38 (n = 17)	1.38 (n = 17)	1.23 (n = 15)	1.09 (n = 11)
	t_{22} = −3.83	t_{14} = −2.48	t_{22} = −1.00	t_{22} = −9.46	t_{22} = −9.46	$t_{10.6}$* = −5.22	$t_{13.97}$* =−3.75
	P = .001	*P* = .03	*P* = .32	*P* < .001	*P* < .001	*P* < .001	*P* = .002
Useful overall							
Yes	4.63 (n = 16)	5.23 (n = 11)	4.04 (n = 12)	7.33 (n = 9)	8.93 (n = 7)	6.73 (n = 11)	6.89 (n = 9)
No	1.50 (n = 8)	2.19 (n = 13)	3.13 (n = 12)	1.33 (n = 15)	1.38 (n = 17)	.92 (n = 13)	1.08 (n = 13)
	t_{22} =−1.96	t_{22}= −2.02	t_{22} = −.57	$t_{9.93}$* = −4.57	t_{22} = −9.46	$t_{13.7}$* = −5.11	$t_{10.8}$* = −4.92
	P = .06	*P* = .06	*P* = .58	*P* = .001	*P* < .001	*P* < .001	*P* < .001
Relevant							
Yes	6.94 (n = 8)	6.71 (n = 7)	3.88 (n = 13)	10.0 (n = 1)	9.90 (n = 5)	8.75 (n = 6)	8.19 (n = 8)
No	1.90 (n = 16)	2.29 (n = 17)	3.22 (n = 11)	3.30 (n = 23)	1.92 (n = 19)	1.86 (n = 18)	1.28 (n = 16)
	t_{22} =−3.70	t_{22} = −2.89	$t_{15.71}$* = −.39	–	–	t_{22} = −5.83	t_{22} =−7.58
	P = .001	*P* = .008	*P* = .70			*P* < .001	*P* < .001
Trustworthy							
Yes	5.40 (n = 15)	6.71 (n = 7)	4.77 (n = 13)	7.20 (n = 5)	9.25 (n = 4)	6.22 (n = 9)	5.07 (n = 14)
No	.56 (n = 9)	2.29 (n = 12)	2.18 (n = 11)	2.63 (n = 19)	2.58 (n = 19)	2.00 (n = 15)	1.88 (n = 8)
	$t_{17.14}$* = −4.57	t_{17}= −3.14	t_{22} = −1.68			t_{22} = −2.96	t_{20} =−1.96
	P < .001	*P* = .006	*P* = .11			*P* = .007	*P* = .07
Knowledgeable							
Yes	4.58 (n = 18)	5.81 (n = 11)	4.93 (n = 15)	8.57 (n = 7)	8.57 (n = 7)	5.82 (n = 14)	5.34 (n = 16)
No	.58 (n = 6)	1.83 (n = 12)	1.50 (n = 8)	1.53 (n = 17)	1.53 (n = 17)	.45 (n = 10)	.08 (n = 6)
	$t_{21.97}$* = −3.76	t_{21} = −2.78	t_{21} = −2.16	t_{22} = −7.15	t_{22} = −7.15	$t_{15.4}$* = −5.23	$t_{15.2}$* = −5.70
	P = .001	*P* = .01	*P* = .04	*P* < .001	*P* < .001	*P* < .001	*P* < .001
Understandable							
Yes	3.58 (n = 24)	3.74 (n = 23)	4.55 (n = 10)	2.27 (n = 13)	4.08 (n = 19)	3.90 (n = 22)	4.38 (n = 17)
No	– (n = 0)	0 (n = 1)	2.89 (n = 14)	5.14 (n = 11)	1.70 (n = 5)	0 (n = 2)	1.64 (n = 7)
	–	–	t_{22} = −1.024	$t_{14.85}$* = 1.80	–	–	$t_{20.2}$* = −2.07
			P = .32	*P* = .09			*P* = .05
Navigable							
Yes	3.3 (n = 23)	3.95 (n = 21)	3.79 (n = 12)	3.2 (n = 16)	3.84 (n = 16)	3.08 (n = 18)	2.98 (n = 20)
No	10 (n = 1)	1 (n = 3)	3.38 (n = 12)	4.3 (n = 8)	3.50 (n = 7)	5.08 (n = 6)	6.63 (n = 4)
	–	–	t_{22} = −.26	t_{22} = .64	t_{21} = −.19	t_{22} = 1.09	t_{22} = 1.78
			P = .80	*P* = .53	*P* = .85	*P* = .29	*P* = .09
Attractive							
Yes	3.74 (n = 23)	3.92 (n = 13)	3.93 (n = 7)	6.80 (n = 5)	1.83 (n = 3)	6.42 (n = 7)	4.25 (n = 4)
No	0 (n = 1)	3.18 (n = 11)	3.44 (n = 17)	2.81 (n = 18)	3.92 (n = 19)	2.41 (n = 17)	3.45 (n = 20)
	–	t_{22} = −.46	t_{22} = −.27	–	–	t_{22} = −2.55	–
		P = .65	*P* = .79			*P* = .02	
Recommended							

	Consumer Raters			**Health Professional Raters**			
				Evidence-Based		**DISCERN**	
Item	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 1**	**Rater 2**	**Rater 1**	**Rater 2**
Yes	5.39 (n = 9)	4.65 (n = 10)	4.04 (n = 12)	10.83 (n = 3)	4.25 (n = 2)	8.75 (n = 6)	7.36 (n = 7)
No	2.5 (n = 15)	2.19 (n = 13)	3.13 (n = 12)	2.60 (n = 20)	3.39 (n = 19)	1.86 (n = 18)	1.83 (n = 15)
	t_{22} =−1.84	t_{21} =−1.66	t_{22} = −.57	–		t_{22} = −5.83	t_{20} = −4.00
	P = .08	*P* = .11	*P* = .58			*P* < .001	*P* = .001

[*] Levene's correction applied

Consumer DISCERN ratings were strongly correlated with satisfaction ratings (rater 1: *r* = 0.74, *P* < .001; rater 2: *r* = 0.85, *P* < .001) as were expert DISCERN ratings (*r* = 0.95, *P* < .001). By contrast, PageRank was correlated with consumer satisfaction for one rater only (rater 1: *r* = 0.45, *P* = .03; rater 2: *r* = 0.35, *P* > .05; rater 3: *r* = 0.21, *P* > .05) and was only moderately correlated with expert satisfaction ratings (*r* = 0.50, *P* = .01). This difference in correlation for the DISCERN and PageRank conditions was significant for two of the consumers and also for the health professionals (consumer rater 1: difference in *r* = 0.29, 95% CI = −0.02 to 0.60); rater 2: difference in *r* = 0.50, 95% CI = 0.18 to 0.82; rater 3: difference in *r* = 0.64, 95% CI = 0.26 to 1.02; health professionals: difference in *r* = 0.45, 95% CI = 0.23 to 0.67)

Discussion

This study provides the first published demonstration that DISCERN is an indicator of evidence-based website quality when used by consumers. It also confirms our previous finding [4] that DISCERN is an indicator of evidence-based quality when used by health professionals.

The finding that DISCERN may be a valid means for consumers to identify websites of high quality and satisfaction has practical implications for consumers. It is unlikely that individual consumers will invest the time required to use DISCERN solely for their own purposes. However, used with caution and an understanding that it is not a perfect predictor of evidence-based quality, DISCERN may be relevant to consumer organizations interested in assembling lists of links to high quality websites for their membership or for visitors to their website. Moreover, the finding that DISCERN may be useful for consumers raises the possibility that DISCERN might also be validly applied by other nontechnical experts, an observation of potential relevance to any organization or Web constructor interested in inexpensively assembling quality portals.

Interestingly, in the case of consumers, Google PageRank is as strong an indicator of evidence-based quality as DISCERN. Thus, this measure may be a simple and practical means by which individual consumers can evaluate, albeit imperfectly, the likely quality of mental health sites. Apart from the time required to download the Google toolbar in the first instance, its use requires minimal expertise and time. In addition, PageRank is likely to be convenient for users seeking health information on the Web since they typically do so by means of a search rather than via directories or portals [19,20]. Since the Google PageRank was correlated less highly with satisfaction than was DISCERN, the latter may be the preferred rating tool for organized groups for whom the overhead in learning to use DISCERN can be justified. However, even in this circumstance, it is possible that Google PageRank could be used as a screening device to eliminate likely sites of low quality and the more time consuming DISCERN instrument then applied to the remaining sites. Alternatively, the reduction in sites may render the task of assessment by a content expert feasible.

It is encouraging that sites regarded by consumers as more useful, trustworthy, and relevant are, on average, sites of higher evidence-based quality. This suggests that consumers' own judgment of and satisfaction with website content may be a useful indicator of appropriate sites. Finally, consumers might be guided by the finding from this and two of our previous studies [4,21] that sites produced by organizations and sites that have an editorial board are of above-average quality. In addition, consumers may be able to place more reliance on sites that pay attention to factors such as a privacy policy, a disclaimer, feedback mechanisms, and on sites that involve health professionals. By contrast, stylistic attributes (eg, judged attractiveness) do not appear to be a useful basis for identifying higher quality sites.

Limitations

This study suffers from several limitations. First, considerable caution is needed in applying the results given that the correlations between the evidence-based scores and DISCERN and Google PageRank were of the order of 0.6 for the consumers. Although considered a strong relationship in the behavioral sciences [22], correlations of this magnitude will result in misclassifications, including false positives and false negatives. Second, the number of consumers employed in the study was small. Third, the study was confined to the field of depression. Fourth, a study of the psychometric properties of the satisfaction measure has not been undertaken. It is therefore difficult to determine if the lower agreement between satisfaction and DISCERN among consumers reflects inadequate reliability of the measure for consumers or a greater variability among consumers than health professionals as to what constitutes satisfaction. In addition, consumer scores on this measure may have been influenced by their concurrent use of DISCERN. Similarly, satisfaction ratings provided by the evidence-based health professional raters may have been influenced by their prior coding of site characteristics and ratings of evidence-based quality. It would therefore be appropriate to repeat the study with a larger number of consumers and health professionals, to employ a design in which the ratings on different instruments, such as DISCERN and satisfaction, were each provided by different consumers and health professionals using a validated, reliable measure of satisfaction, and to determine if the findings are robust across a range of mental health and other health domains. In addition, although a number of site characteristics were associated with better evidence-based quality, the website sample size was insufficient to conduct analyses to identify the independent effects of these characteristics on quality. It is possible, for example, that organizations are more likely to both produce high-quality sites and incorporate a privacy policy, disclaimer, and feedback mechanism. Finally, none of the raters—professional or consumer—were experienced in the use of the DISCERN instrument. The findings may therefore underestimate the usefulness of DISCERN as an indicator of quality when used by an experienced rater.

Conclusions

These results represent a first step toward identifying tools that consumers who are not content experts can use as valid indicators of the evidence-based quality of websites. Further research is required to explore the utility of DISCERN and Google PageRank. In particular, it is important to determine optimal cutoff points for identifying higher quality sites and to explore the sensitivity and specificity of the measures. It is also of interest to document the relative utility of DISCERN for nontechnical raters of differing educational backgrounds, experience with the instrument, and Web experience. Finally, given that not one but many indicators may be useful in identifying high-quality sites, there may be value in identifying optimal combinations of multiple indicators of quality. There is also much to be gained by further identifying automatic indicators of the type that could be factored into the relevance algorithms of a specialized focused search engine.

Acknowledgments

This study was funded by a National Health and Medical Research Council Australia Program Grant to the Centre for Mental Health Research and by a grant from beyondblue: the national depression initiative. The work described here was carried out independently of the funders.

Authors' Contributions

KG conceived and designed the study, acted as a health professional rater (evidence-based), analyzed and interpreted the data, and wrote the paper. HC designed the study, acted as a health professional rater (evidence-based), and edited the paper. Dr. Simon Blomberg generated the list of websites, supervised the consumer research assistants, and collected the data. He and Kelly Blewitt set up the database and entered the data. Three consumers provided DISCERN and satisfaction ratings and were employed as research assistants for this purpose. Professor Anthony Jorm and Dr. Richard O'Kearney (health professionals) provided DISCERN and satisfaction ratings for each site. Anthony Bennett developed the Web-capture software.

Conflicts of Interest

None declared.

References

1. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 22016241] [doi: 10.1001/jama.287.20.2691]
2. ; Commission of the European Communities eEurope. Quality criteria for health related websites. J Med Internet Res 2002;4(3):e15. [doi: 10.2196/jmir.4.3.e15]
3. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Commun Health 1999 Feb;53(2):105-111 [FREE Full text] [Medline: 99324793]
4. Griffiths KM, Christensen H. The quality and accessibility of Australian depression sites on the World Wide Web. Med J Aust 2002 May 20;176 Suppl:S97-S104 [FREE Full text] [Medline: 22061104]
5. Bartels U, Hargrave D, Lau L, Esquembre C, Humpl T, Bouffet E. Analyse padiatrisch neuro-onkologischer Informationen in deutschsprachigen Internetseiten. Klin Padiatrie 2003;215(6):352-357. [doi: 10.1055/s-2003-45491]
6. Turp J, Gerds T, Neugebauer S. Myoarthropathien des Kausystems: Beurteilung der Qualitat von Patienteninformationen im Weltweiten Netz. Zeitschrift fur arztliche Fortbildung und Qualitatssicherung. In: Zusammenarbeit mit der Kaiserin-Friedrich-Stiftung fur das arztliche Fortbildungswesen 2001;95(8):539-547.
7. Murray CJL, Lopez A. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Cambridge, Mass: Harvard University Press; 1996.
8. Andrews G, Henderson S. Unmet Need in Psychiatry: Problems, Resources, Responses. Cambridge, UK: Cambridge University Press; 2000.
9. Taylor H. Explosive growth of 'cyberchondriacs' continues. URL: http://www.harrisinteractive.com/harris_poll/index.asp?PID=117

10. The R Project for Statistical Computing. URL: http://www.r-project.org/ [accessed 2005 Oct 24]
11. ; Centre for Evidence Based Mental Health, University of Oxford. A systematic guide for the management of depression in primary care [treatment]. URL: http://www.psychiatry.ox.ac.uk/cebmh/guidelines/depression/depression1.html [accessed 2005 Oct 26]
12. Charnock D. The DISCERN handbook. Quality criteria for consumer health information on treatment choices. Oxford, UK: Radcliffe Medical Press Ltd; 1998.
13. Hargrave D, Bartels U, Lau L, Esquembre C, Bouffet E. [Quality of childhood brain tumour information on the Internet in French language]. Bull Cancer 2003 Jul;90(7):650-655. [Medline: 22837922]
14. Ademiluyi G, Rees CE, Sheard CE. Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. Patient Educ Couns 2003 Jun;50(2):151-155. [Medline: 22667539]
15. Brin S, Page L. Anatomy of a large-scale hypertextual web search engine. Presented at: 7th International World Wide Web Conference; April 14-18, 1998; Brisbane, Australia.
16. Pandurangan G, Raghavan P, Upfal E. Using PageRank to Characterize Web Structure. Presented at: 8th Annual International Conference on Combinatorics and Computing (COCOON); 2002; Singapore.
17. SISA online statistical analysis. URL: http://home.clara.net/sisa/correl.htm [accessed 2005 Mar 11]
18. ; SPSS Inc. SPSS (12.0.1) for Windows. Chicago, Ill: SPSS. URL: http://www.spss.com/spss [accessed 2005 Mar 11]
19. 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] [PMC: 11884321] [Medline: 21881326] [doi: 10.1136/bmj.324.7337.573]
20. Fox S, Fallows D. Internet health resources. Washington, DC: Pew Internet & American Life Project; 2003.
21. Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. BMJ 2000 Dec 16;321(7275):1511-1515 [FREE Full text] [PMC: 11118181] [Medline: 20568134] [doi: 10.1136/bmj.321.7275.1511]
22. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

submitted 20.01.05; peer-reviewed by H Witteman, S Shepperd; comments to author 18.02.05; revised version received 09.03.05; accepted 21.10.05; published 15.11.05.

Please cite as:
Griffiths KM, Christensen H
Website Quality Indicators for Consumers
J Med Internet Res 2005;7(5):e55
URL: http://www.jmir.org/2005/5/e55/
doi:10.2196/jmir.7.5.e55
PMID:16403719

Original Paper

Automated Assessment of the Quality of Depression Websites

Kathleen M Griffiths[1], PhD; Thanh Tin Tang[2], B.IT (Hons); David Hawking[3], PhD; Helen Christensen[4], PhD

[1]Depression & Anxiety Consumer Research Unit, Centre for Mental Health Research, The Australian National University, Canberra, Australia
[2]Computer Science Department, The Australian National University, Canberra, Australia
[3]ICT Centre, CSIRO, GPO Box 664, Canberra, Australia
[4]Centre for Mental Health Research, The Australian National University, Canberra, Australia

Corresponding Author:
Kathleen M Griffiths, PhD
Centre for Mental Health Research
The Australian National University
Canberra 0200
Australia
Phone: +61 2 6125 9723
Fax: +61 2 6125 0733
Email: kathy.griffiths@anu.edu.au

Abstract

Background: Since health information on the World Wide Web is of variable quality, methods are needed to assist consumers to identify health websites containing evidence-based information. Manual assessment tools may assist consumers to evaluate the quality of sites. However, these tools are poorly validated and often impractical. There is a need to develop better consumer tools, and in particular to explore the potential of automated procedures for evaluating the quality of health information on the web.

Objective: This study (1) describes the development of an automated quality assessment procedure (AQA) designed to automatically rank depression websites according to their evidence-based quality; (2) evaluates the validity of the AQA relative to human rated evidence-based quality scores; and (3) compares the validity of Google PageRank and the AQA as indicators of evidence-based quality.

Method: The AQA was developed using a quality feedback technique and a set of training websites previously rated manually according to their concordance with statements in the Oxford University Centre for Evidence-Based Mental Health's guidelines for treating depression. The validation phase involved 30 websites compiled from the DMOZ, Yahoo! and LookSmart Depression Directories by randomly selecting six sites from each of the Google PageRank bands of 0, 1-2, 3-4, 5-6 and 7-8. Evidence-based ratings from two independent raters (based on concordance with the Oxford guidelines) were then compared with scores derived from the automated AQA and Google algorithms. There was no overlap in the websites used in the training and validation phases of the study.

Results: The correlation between the AQA score and the evidence-based ratings was high and significant (r=0.85, *P*<.001). Addition of a quadratic component improved the fit, the combined linear and quadratic model explaining 82 percent of the variance. The correlation between Google PageRank and the evidence-based score was lower than that for the AQA. When sites with zero PageRanks were included the association was weak and non-significant (r=0.23, *P*=.22). When sites with zero PageRanks were excluded, the correlation was moderate (r=.61, *P*=.002).

Conclusions: Depression websites of different evidence-based quality can be differentiated using an automated system. If replicable, generalizable to other health conditions and deployed in a consumer-friendly form, the automated procedure described here could represent an important advance for consumers of Internet medical information.

(J Med Internet Res 2005;7(5):e59) doi:10.2196/jmir.7.5.e59

KEYWORDS

Quality indicators; depression; Internet; World Wide Web; validity; information retrieval

Introduction

At least 50% of households in the United States, the United Kingdom and Australia are connected to the Internet [1-3]. In addition, many people have access to the Internet outside the home (eg, at work and in public libraries). A recent survey found that 79% of American Internet users have searched for health information online [4]. Moreover, there is evidence that online health information can improve health knowledge and health outcomes [5, 6].

To date, however, few health information websites have been subjected to rigorous assessment of their effectiveness in improving health outcomes. In the absence of such evidence, a minimum requirement for a high quality health information website should be that its content is evidence-based; that is, that its content and recommendations are consistent with evidence from a systematic review of the available medical literature. A non-evidence-based site would recommend treatments that are not supported by the evidence as effective, oppose the use of beneficial treatments of demonstrated benefit or fail to mention some effective treatments at all. For example, a depression webpage on the official site of the Church of Scientology claims that Dianetics is “the only proven effective technology of the human mind” [7] and fails to mention other medical, psychological and alternative treatments of demonstrated effectiveness. Elsewhere on the site, Zoloft and Prozac are described as “even more damaging than street drugs” [8].

Evidence-based health information is routinely disseminated to health professionals with the aim of assisting clinical decision making, improving healthcare and thereby improving health outcomes. Such evidence, when provided to consumers, has the potential to improve health outcomes by assisting consumers to select effective, rather than ineffective self-help techniques and by supporting shared decision making and consumer-provider collaborative care [9]. Unfortunately, reviews of the content of health websites have demonstrated that their quality, when assessed relative to evidence-based standards, is variable [10-13]. People seeking health information may therefore require assistance to filter out lower quality websites.

Various mechanisms have been proposed for assisting consumers to access high quality health websites [14]. These include the use of quality portals (such as OMNI in the United Kingdom and HealthInsite, Australia), pledges of webmasters to adhere to codes of conducts (such as the HON code [15]) and the use of consumer tools for assessing a site (for example DISCERN [16]). However, the criteria for inclusion in a quality portal typically do not include an evidence-based assessment, so their validity as guides to quality material is uncertain. Moreover, since such portals require time, effort and training to maintain, it may be difficult to update the database in a timely fashion. Similarly, codes of conduct and consumer tools employ accountability criteria (such as the identification of the author, their affiliations and qualifications) which are typically not validated against an evidence-based standard [11, 17, 18]. One exception is DISCERN, a tool designed to assist users without technical expertise to assess the quality of health information users. We have reported some preliminary evidence that DISCERN may be a valid indicator of the evidence-based quality of websites when used by consumers [19]. However, the tool may not be valid for all consumers. Moreover, DISCERN requires training and is lengthy, involving 15 items and requiring assessment of all the pages on the target website. In practice, individual consumers may lack the time and motivation to undergo the necessary training or to apply the tool to individual websites.

A potential solution to these problems is to develop assessment tools based on algorithms that automatically evaluate the quality of health information websites. To date there has been little work directed to this possibility. There is some evidence of a relationship between Google Page Rank and evidence-based quality from two recent studies [19, 20]. However, the association appears to be only of moderate strength, suggesting that a more valid automatic indicator of website quality may be required.

This paper describes the development of a computer algorithm, the Automatic Quality Assessment procedure (AQA), designed to automatically rank depression websites according to the evidence-based quality of their treatment information. In addition, it describes the results of an evaluation of the validity of the AQA as an indicator of human-rated evidence-based quality of treatment content. It also compares the validity of the AQA and Google PageRank as indicators of evidence-based quality.

The study focused on the evaluation of treatment information in depression websites since depression is a primary source of disability burden [21] and it has been reported to be a condition for which users commonly seek information on the Internet [4]. There is also a high degree of unmet need in the treatment of depression [22].

Methods

This section comprises two parts. The first describes the AQA and its development. The second describes the methodology used for the validation of the AQA as an indicator of evidence-based treatment quality. The evidence-based rating scale [12] employed in developing and validating the AQA was based on clinical practice depression guidelines developed by the Centre for Evidence-Based Mental Health in Oxford from a systematic review of the evidence [23].

The Automatic Quality Assessment Procedure (AQA)

In the following we present the procedure for calculating AQA scores and note its dependence upon two learned queries and three numerical parameters. We then describe the development phase during which the queries were learned and the parameter values chosen. The development phase employed websites/webpages not in the validation set but for which we had collected human-rated relevance or evidence-based quality measures from our previous studies [12, 19, 24].

The AQA assumes the availability of search engine software that incorporates a web crawler and has the ability to effectively score the relevance of documents to a query. The current study employed the Panoptic search engine for this purpose. However, we believe that other similar search engines could be substituted with similar results. A set of computer scripts were written to learn queries, to set values of tuning parameters and to collect and analyze output from the search engine. These scripts are not part of the Panoptic search engine.

The Procedure

The AQA procedure comprised six steps as follows:

1. The target websites were downloaded using web crawler software;
2. These downloaded pages were aggregated with a large set of arbitrarily chosen general English language web pages and the resulting collection indexed using the search engine. This was necessary to avoid the extremely biased term frequency distribution of a depression-only collection;
3. A previously learned relevance query (see below) was processed over the collection created in Step 2 using the search engine to produce relevance scores for all documents. The relevance query consisted of many words and phrases, each with a numerical importance weighting. Documents with non-zero scores were not retrieved. For each site to be evaluated, the number of retrieved documents |R| were counted and the mean relevance score (*r*) computed.
4. A previously learned quality query (see below) was processed in the same fashion as in Step 3, yielding |Q| and *q*.
5. Site relevance and site quality scores were computed using Equations (1) and (2). These scores were normalized such that the highest S_r became 1.0 and the highest S_q was also 1.0.
6. An overall site score was computed using Equation (3). Gamma is a scaling parameter designed to make scores comparable with those from the human rating scale.

Equation 1: $S_r = \alpha \times r + (1-\alpha) \times |R|$

Equation 2: $S_q = \alpha \times q + (1-\alpha) \times |Q|$

Equation 3: $S = \gamma \times (\beta \times S_q + (1-\beta) \times S_r)$

The following sections describe how relevance and quality queries were learned and the values of α, β and γ chosen.

Learning Relevance and Quality Queries

Relevance and quality queries were learned using an extension and novel application of the relevance feedback technique from the field of information retrieval. In the relevance feedback approach, a complex query consisting of weighted terms (words and phrases), is automatically generated by comparing the term frequency distributions of sets of relevant and irrelevant documents. Good terms occur frequently in relevant text but seldom otherwise. The resulting query is used by a text retrieval system to derive relevance scores for documents. We extended this method to learn a 'quality' query from sets of high and low quality webpages.

Relevance query: During development of the relevance query, query terms were selected by computing Term Selection Values (TSVs) [25] for each candidate term, ranking them in descending order and taking all the terms above a cutoff. Numerical weights were applied to the selected terms using the Robertson-Sparck Jones formula [26].

Using 347 documents previously judged relevant to the topic of depression [24] and 9000 documents with very low probability of relevance to that topic, we generated a relevance query consisting of the words with the 20 highest TSVs and the two-word phrases with the 20 highest TSVs. The cutoff of 20 was arbitrary but consistent with past information retrieval practice.

Quality query: We generated a quality query in the same fashion, using 110 documents judged to be relevant to depression and of high quality as the “relevant” set and 3002 documents which were judged either irrelevant or relevant but not of high quality [24]. In this case the number of words (29) and phrases (20) in the query was the minimum number needed to ensure the inclusion of the names of all the evidence-based depression treatments listed in our previously published systematic review of the effectiveness of medical, psychological and alternative treatments for depression [27].

Choosing Parameter Values

All the documents from 29 training sites for which we had human evidence-based (Oxford) ratings from previous studies [12, 19] were fetched using the Panoptic crawler and combined with 10000 documents from the Yahoo! Directory which were not in the depression category, as per Step 2 of the AQA procedure.

In following Steps 3 and 4 of the procedure during training, we computed |R|,*r*,|Q| and *q* based on scores obtained using the Okapi BM25 [28] relevance scoring mode of the Panoptic search engine. Okapi BM25 takes into account the frequency of occurrence of query terms in a document, the discriminating power of each query term, and the length of the document in calculating a relevance score.

The parameter adjusts the balance between the average document score and the coverage of a site. We then arbitrarily chose α = 0.75. The parameter β adjusts the balance between the relevance and quality scores for a site. We stepped through the range of values between 0.0 and 1.0 and chose the value which, when used in Equation 3, maximized the correlation between the computed site scores and the human-assigned quality scores. The best combination found, α = 0.75 and β = 0.70, yielded a correlation of 0.94 on the training data. It is possible that better values could be found with a more exhaustive optimization of parameters.

The parameter γ does not affect the correlation but scales the raw AQA scores to match the range of the human assigned scores. We chose γ = 17.27 which caused the highest AQA score to be the same as the highest human-rated score.

The values determined in training (α = 0.75, β = 0.70 and γ = 17.27) were used in the validity testing phase.

Validity of the AQA versus PageRank

Here we describe the methodology used in a comparative validation study of the AQA and the Google PageRank procedures as an indicator of evidence-based treatment website quality. Each of two judges provided evidence-based ratings of 30 new depression websites. These ratings were compared with automated scores derived from the AQA and Google PageRank.

Selection of Sites

The 30 depression information test websites were selected in the following manner.

First, we compiled a master list of all depression websites from the Open Directory (http://dmoz.org), Yahoo (http://www.yahoo.com) and LookSmart (http://www.looksmart.com) main and personal and treatment depression subdirectories as of September 2004. DMOZ, Yahoo and LookSmart are the three major human-compiled search engines on the World Wide Web. The human-compiled directories of many major crawler-based search engines such as Google are derived from Open Directory and the human-compiled content of the Lycos Directory is currently supplied by LookSmart.

After excluding websites that were no longer accessible, websites that were a subdirectory of an already included website, or "websites" that were actually links to an individual article, 208 websites remained.

Using the Google Toolbar, Google PageRank scores were recorded for each of these 208 websites. Sites were then pooled into 5 PageRank bands (0, 1-2, 3-4, 5-6 and 7-8) and, from each of the 5 PageRank bands, 6 websites were randomly selected (using a computer generated random number function) to form an initial set of 30 depression websites. Sites were stratified by PageRank prior to sampling to avoid generating a spuriously low correlation due to restricted range effects. A further 3 websites were excluded because the content was not free, there was no depression information on the primary site, or the site comprised only a single clinical tool for clinicians. These sites were each replaced from the equivalent PageRank band using the same computer generated random function.

Site content for each of the 30 websites was printed out in its entirety by systematically following all internal links. Any audio or video material content on a site was accessed online by the evidence-based raters and incorporated into the overall evidence-based score.

Content within a site was included for evaluation if it was free, written in the English language, comprised core informational material and focused on unipolar depression. Since the evidence-based rating scale employed in the current study was based on systematic guidelines for the treatment of major depressive disorder, pages in a site were excluded from evaluation if they focused on bipolar disorder, premenstrual syndrome, premenstrual dysphoric disorder or seasonal affective disorder. In addition, the following content on the target websites was excluded from the evaluation: news sections, videos of research conferences, book reviews, collections of PubMed abstracts, poetry, message board and chatroom content. This content was excluded because it was often unmanageably large (eg, poetry archives and chatrooms) and peripheral to the core educational material contained on the websites. General clinician assessment instructions and survey databases were excluded as they were not relevant to the study. Non-English content was excluded for practical reasons.

Site Assessment

Site Characteristics

Test sites were coded independently by 2 raters according to their ownership structure (individual vs organization); whether or not they had an editorial board; whether or not the site was depression specific, was somewhat broader in scope, or comprised a clearinghouse; whether a health professional was involved; and whether the site promoted products or services (see Table 1). Where the two coders disagreed, the final categorization was assigned by a third rater (KG).

Evidence-Based Score

Each test site was rated independently by 2 raters using a 20-item rating scale previously developed by us [12] (see Textbox) from statements in the evidence-based, systematically developed clinical practice guidelines for the management of depression in primary care published by the Oxford University's Centre for Evidence-Based Mental Health guidelines [23]. Only statements directly relevant to treatment were incorporated in the 20-item scale. The 30 test sites were rated in a different computer generated random order by the two raters. This rating scale has previously shown high interrater reliability [12, 19]. In the current study, interrater reliability was also very high (r=.93, P<.001) and there was no significant difference between scores for the two raters (mean difference= 0.17, 95% Confidence Interval (CI)= -0.96–0.62, P=.67).

Textbox 1. Evidence-Based Rating Scale for Human Raters

The evidence-based rating scale [12] was developed from statements in the treatment section of *A systematic guide for the management of depression in primary care* published by the Centre for Evidence-based mental health, Oxford [23]

- Antidepressant medication is an effective treatment for major depressive disorder.
- Antidepressants are all equally effective.
- The effectiveness of antidepressants is around 50 to 60%.
- Full psychosocial recovery can take several months.
- Drop out rate is same for different antidepressants.
- The side effect profile varies for different antidepressants.
- The choice of antidepressant should depend on individual patient factors (eg presence of co-morbid psychiatric or medical conditions, previous response to a particular drug, patient preference regarding the desirability of specific side-effects, concurrent drug therapy, suicidal risk)
- Antidepressants are not addictive.
- A trial of 6 weeks at full dose is needed before a drug can be considered to have failed and another tried.
- A second-line drug should probably be from a different class of antidepressant.
- Once improved continuation treatment at the same dose for at least 4-6 months should be considered.
- Discontinuation syndrome may occur with abrupt cessation of any antidepressant so antidepressants should not be stopped suddenly. Where possible antidepressants should be withdrawn over a 4 week period, unless there are urgent medical reasons to stop the drug more rapidly. [To score 1, need to make general points that abrupt cessation can cause discontinuation syndrome and that antidepressants should not be stopped suddenly]
- St John's Wort appears to be as effective as tricyclic antidepressants and causes fewer side effects, but little is known about any long term adverse effects.
- Cognitive therapy can be an effective treatment for depression.
- Cognitive behaviour therapy is at least as effective as drug treatment in mild-to-moderate depression.
- Cognitive behaviour therapy may be valuable for people who respond to the concept of Cognitive behaviour therapy, prefer psychological to antidepressant treatment, have not responded to antidepressant therapy. [Score 1 if mention at least one of these]
- Problem-solving may be effective for depression.
- [Generic] counselling is probably no more effective than treatment as usual from the GP for depression.
- Written information (usually based on a cognitive model of depression) can improve mild-to-moderate depression. [Score 1 if cognitive model]
- Exercise can be effective - alone or as an adjunct to other treatments.

For each item, score 1 if the site information is consistent with the statement. Cumulate item scores across the scale to yield a total evidence-based score for the site.

Computing AQA Scores for the Test Sites

AQA scores were computed for the test websites by following Steps 1 to 6 of the AQA procedure using the relevance and quality queries and the values of α and β that were derived during training.

Google PageRank

The Google PageRank was recorded for each home page. Google PageRank is a measure employed by the Google Search Engine company to evaluate the reputation of a webpage. The PageRank is based on a computer algorithm that computes iteratively the number and importance of links to a webpage and in turn the number and importance of links to these linking pages [29]. As noted above, we identified the PageRank for each test site by downloading the Google Toolbar and recording the integer value (range 0 to 10) on the toolbar for the homepage of the site.

Statistical Analysis

The sample size was considered sufficient to justify meaningful parametric analysis of the data. Intercorrelations between variables were computed using Pearson correlation tests. The validity of the automatic measure as an indicator of evidence-based quality was evaluated using hierarchical multiple regression. These analyses were performed using SPSS 12.0.1 [30]. Tests of the significance of differences between dependent correlations were computed using the SISA online calculator [31].

Results

Site characteristics

The characteristics of the 30 depression test sites used in the validation are summarized in Table 1. Two-thirds of the sites were depression-specific, a little over one-half were owned by an individual, and a health professional was involved in approximately half of the sites. One-fifth of the sites had an editorial board and over half of the sites promoted products or services or both.

Table 1. Characteristics of the websites employed in the test phase of the study

Site characteristic	n (%) of sites (N=30)
Ownership structure	
Individual	17 (56.7%)
Organization*	12 (40.0%)
Unknown	1 (3.3%)
Editorial board	
Yes	6 (20%)
No	24 (80%)
Scope	
Depression specific	20 (66.7%)
Broad scope	9 (30.0%)
Clearing house	1 (3.3%)
Health professional involved	
Yes	16 (53.3%)
No	14 (46.7%)
Promotion of products/ services	
Yes	19 (63.3%)
No	11 (36.7%)

* Commercial, consumer or other organized group

Quality scores

The mean (and standard deviation) of the evidence-based, Google PageRank and AQA scores were 5.92 (SD = 5.46; n = 30), 3.67 (SD = 2.59; n = 30) and 8.07 (SD = 5.22; n = 29) respectively. AQA scores were available for 29 test sites only as one website included a robots.txt exclusion, an indicator that the administrator of the site prohibited external crawlers from accessing the website.

The relationship between the AQA score and the evidence-based ratings is shown in Figure 1. The linear correlation between these two measures was high and significant (r=0.85, $P < .001$).

Figure 1. The relationship between the AQA and evidence-based scores

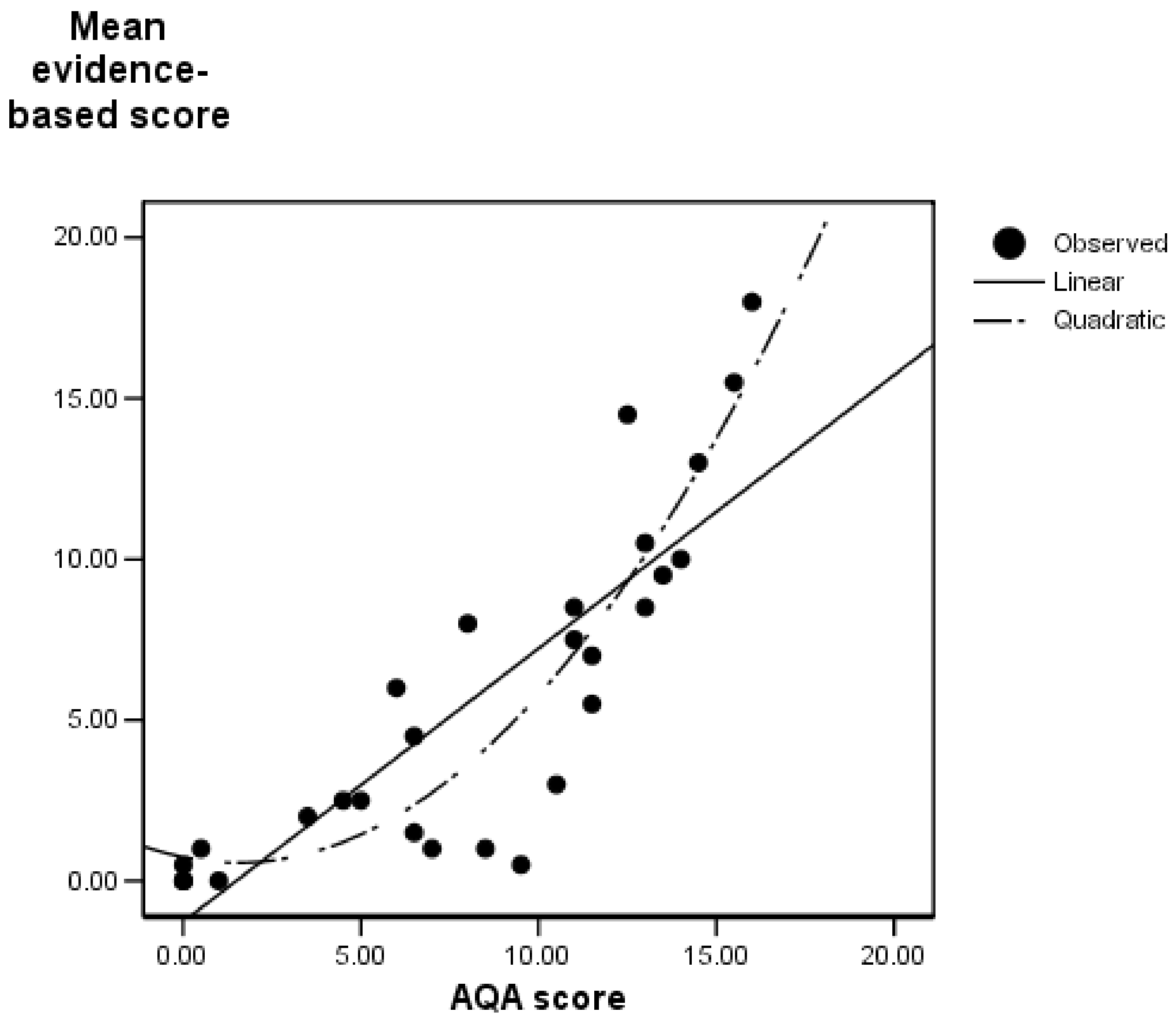


Table 2. Summary of hierarchical regression analysis for predicting evidence-based quality from automatic quality

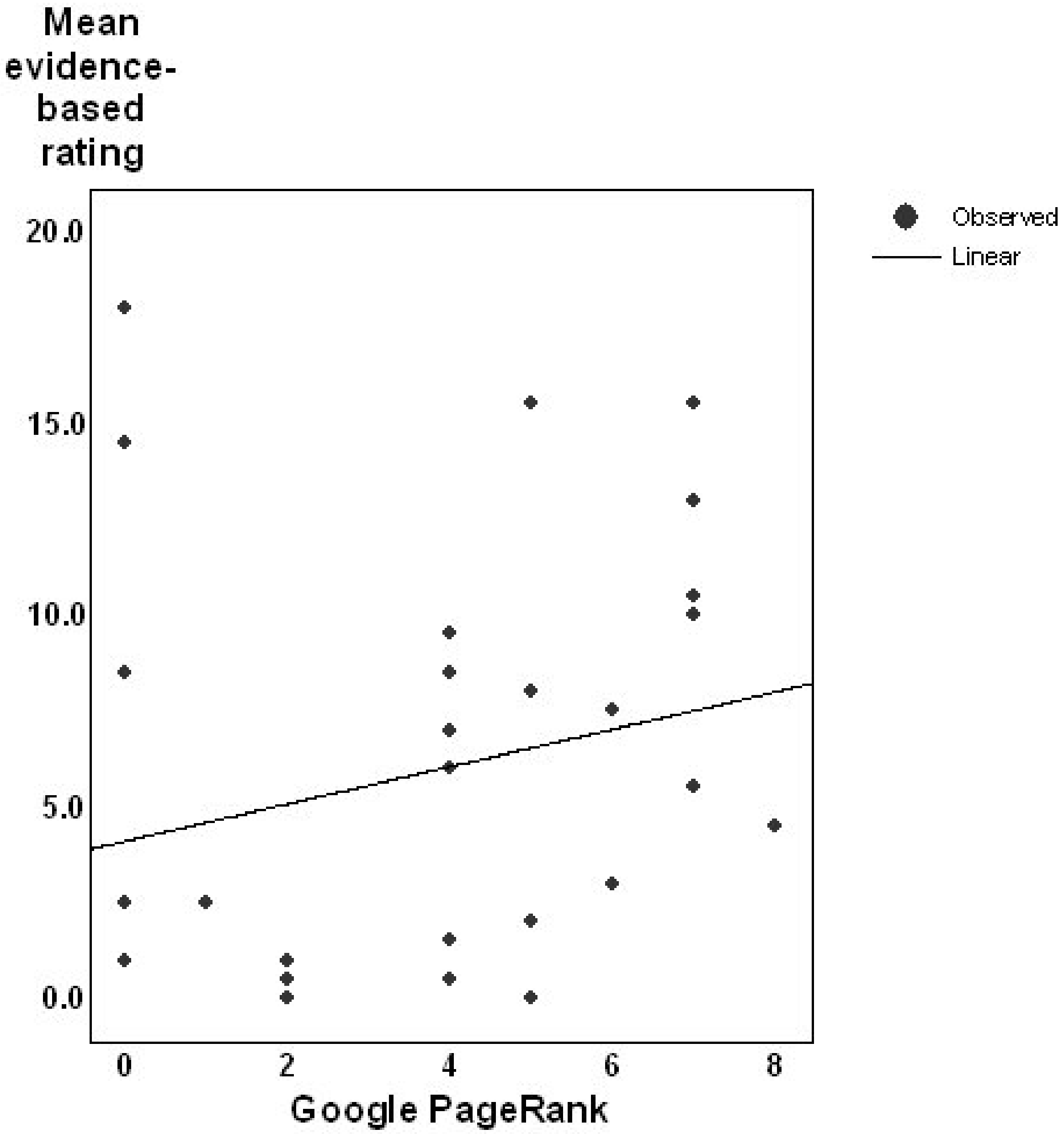
Variable	B	SE(B)	β	P
Model 1				
AQA	.85	.10	.85	*P*<.001
Model 2				
AQA	-.22	.30	-.22	*P* = .45
AQA^2	.07	.02	1.11	*P* = .001

Note: Model 1: R^2 = .71 ; Model 2: ΔR^2 = .10 (*P* = .001)

Visual inspection of the scatterplot in Figure 1 indicated a possible quadratic component to the relationship. A hierarchical multiple regression analysis was therefore performed to determine if adding the square of the AQA score to the linear solution improved the prediction (see Table 2). A substantial 71.4% of the variance in the evidence-based quality score was explained by the automatic quality score alone. Addition of the quadratic component significantly improved the fit ($\Delta R^2 = 0.10$, $F\Delta$ (1,26) = 14.3, *P* = .001), the combined linear and quadratic model explaining 82 percent of the variance.

By contrast, the correlation between Google PageRank and the evidence-based score was small and non-significant (r = 0.23, *P* = .22, n = 30; see Figure 2). Excluding the missing case for which AQA could not be computed, this association between Google PageRank and the evidence-based score was significantly lower than the association between the AQA score and the evidence-based score (r(difference) (df = 26) = .64, t = 4.82, *P* = .0001).

Figure 2. The relationship between the Google PageRank and evidence-based scores.



Since it has been argued by some members of the search engine optimisation community that PageRanks of 0 may constitute a special subset of PageRank values (see Comment below), the above analyses were recomputed after excluding sites with PageRank of 0. The association between the PageRank and evidence-based scores for remaining sites was significant (r = 0.61, *P* = .002, n = 24), but remained significantly lower than the association between the AQA and evidence-based quality scores (r(difference) (df = 20) = 0.22, t = 2.61, *P* = .02).

Discussion

Principal Results

A recent article concluded that “quality benchmarking of health-related resources will always depend on a human assessor …” [32]. We have demonstrated that an algorithm based on relevance feedback (and involving no human judgment) is a valid indicator of evidence-based quality of the treatment content of depression sites. To our knowledge, this is the first published study of the validity of a custom designed automated tool for identifying the evidence-based quality of health information. If replicable and generalizable to other health conditions, the current findings may have major practical implications for e-health, consumer empowerment and self-managed healthcare.

Previous researchers have developed search systems designed to identify medicine-specific Web-based information [33]. However, these systems focus on identifying material that is relevant to the medical domain rather than selecting sites of high content quality.

One published study has described a prototype system for rank ordering Web-based health information by quality [34]. However, this paper used accountability criteria (eg, presence of authorship information, detection of an HONcode logo, detection that the page included information about the date it was last updated) rather than content accuracy as a benchmark of quality. There is little or no evidence that these accountability measures singly or together correlate with evidence-based content quality [10, 11, 17, 18]. Moreover, the researchers did not evaluate the content quality of the retrieved pages in order to validate their system of ranking against an evidence-based standard. Finally, in contrast to the procedure described in the current study, the system focused on individual pages rather than on the entire website on a topic. It may be that only by examining all the content of a site is it possible to gain a comprehensive picture of its quality.

The finding in this study that websites can be automatically evaluated for content quality is of considerable practical significance. Suitably adapted, refined and integrated into or used to post-process websites retrieved by a general search engine, this system could assist consumers to identify websites of higher quality. In the shorter term, the system can be used to compile lists of websites for use in a focused search engine for depression, such as that used on the BluePages Depression Information website (http://bluepages.anu.edu.au). In addition, the system may prove useful as a screening device for the use of web developers interested in maintaining quality health portals or links of high quality. Once filtered by the automatic quality evaluator, developers could evaluate the remaining sites manually to confirm accuracy of content, and to assess sites according to other dimensions of quality (eg, usability) and according to the needs of the organization and its users.

Consistent with previous studies [19, 20] we have demonstrated in the current study that Google PageRank may be an indicator of evidence-based quality. However, the fact that Google PageRank was unable to provide a meaningful quality assessment for sites with a zero score is a significant impediment to its practical use. Moreover, even with zero PageRanks excluded, the association between PageRank and content quality is less strong than the association between the AQA score and content quality. This suggests that relevance feedback (employed by our AQA) may be superior to link structure (employed by Google PageRank) as a method for identifying evidence-based quality for a specific health domain. It also provides evidence that a specialized tool such as the AQA is warranted. It might be argued that a combination of relevance feedback and link structure methods could improve the validity of the AQA further. However, the preliminary indications are that this is not the case. Supplementary analyses showed that when the Google PageRank was added into the regression model along with the AQA scores, it did not improve the fit of the model. This was the case whether zero page ranks were included (explained 71% of the variance compared to 82% in the original model) or excluded (explained 65% compared to 82%).

The current tool is relevant only for identifying the quality of depression websites. However, the relevance feedback method used in the AQA is likely to generalize to other areas of mental health and medicine. To apply the AQA in alternative health domains requires that the procedure is trained to learn new terms and parameter values specific to the new domain. The validity of the technique in these other domains is a matter for empirical investigation.

Limitations

There are some limitations of the system described here. First, some minor changes to the AQA scaling procedure are necessary before the system is used in practice (particularly for evaluating a single website at a time). Applying a non-linear transformation to the raw AQA scores (rather than linear scaling) might result in a better prediction of the values of the evidence-based scores. However, given that the base (linear) correlation coefficient is already very high (0.85), the scope for improvement is limited. Secondly, the AQA scores could be compromised if publishers use “spamming” methods for optimizing their automatic quality scores. General search engines are faced with a similar problem when website developers attempt to artificially inflate the relevance rankings for their websites. This problem is not peculiar to automated methods for processing websites. It is likely that website developers also use strategies to maximize their scores on manually applied quality evaluation tools while leaving unchanged the substantive content. Public search engines incorporate algorithms for detecting attempts to distort rankings. It is likely that the AQA could be refined to do the same. The third limitation of the AQA system is that it is focused on treatment information (as indeed is DISCERN) and does not necessarily reflect the quality of other information on a site. In addition, the system may not adequately rate sites which present only one treatment type. The system will be most useful in identifying sites containing high quality, comprehensive treatment information. Another limitation relates to the gold standard employed for rating the evidence-based quality of the sites. It might be argued that to be considered valid, the rating system should be validated against health outcomes or another scale that has been thus validated. No such scales exist, and given the paucity of efficacy studies of websites, such validation exercises are not currently practical. Finally, one of the evidence-based raters had been involved in the initial selection of the sites which may have led to bias in the findings from the validity studies. However, the pattern of results and conclusions is identical if the findings are recomputed using the blinded rater’s data only. This is not surprising given the higher interrater reliability in this study for the evidence-based scale.

Conclusion and Future Work

The time has come to acknowledge that consumers do and will continue to use the Internet as a source of health information. We need to provide them with convenient, effective tools that optimize the usefulness of this process. This study demonstrates that automated methods offer considerable promise in this respect. The task before us now is to refine these methods, and to test the usability, robustness and generalizability of the systems we develop. In the process we need to test alternative strategies for quality filtering, to identify if multiple methods when combined improve the validity of the automatic algorithm, and to evaluate whether the techniques generalize across health domains. We must also construct a user interface for the procedure, and conduct consumer user and satisfaction studies on the resulting system.

Acknowledgments

This study was funded by the National Health and Medical Research Council Australia Program Grant to the Centre for Mental Health Research (Chief investigators: A F Jorm, H Christensen, B Rodgers, K Dear, and K Anstey). The work described here was carried out independently of the funders. The authors would like to thank the following people for their contributions to the study: Michelle Banfield who systematically identified and printed out the webpages for the test collection, provided evidence-based ratings of the test websites, set up the database for the study and conducted some preliminary analyses of the results and Kerry Sargent-Cox who provided evidence-based ratings of the test data.

Conflicts of Interest

DH and TT are employed by CSIRO which owns and has commercialised the Panoptic search technology used in the study. DH is leader of the Panoptic R&D team. The AQA is not currently part of any commercial product and no patents have been applied for in respect of it.

References

1. ; National Telecommunications and Information Administration. A nation online: Entering the broadband age. US Department of Commerce: Washington; 2004. URL: http://www.ntia.doc.gov/reports/anol/NationOnlineBroadband04.htm [WebCite Cache ID 95792]
2. ; Office for National Statistics. Individuals accessing the Internet - National Statistics Omnibus Survey. Access to Internet from home. Family Expenditure Survey 2005. Office for National Statistics: London URL: http://www.statistics.gov.uk/CCI/nugget.asp?ID=8&Pos=1&ColRank=1&Rank=192 [accessed 24 December, 2005] [WebCite Cache ID 95781]
3. ; Australian Bureau of Statistics. Household Use of Information Technology, Australia, 2002 and 2003. Canberra, Australia: Australian Bureau of Statistics; 2005.
4. Fox S. Health Information Online. URL: http://www.pewinternet.org/pdfs/PIP_Healthtopics_May05.pdf [accessed July 21, 2005] [WebCite Cache ID 95819]
5. Christensen HK, Griffiths KM, Jorm A. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004;328(7434):265. [doi: 10.1136/bmj.37945.566632.EE]
6. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185(7434):342-349 [FREE Full text] [Medline: 15458995] [doi: 10.1192/bjp.185.4.342]
7. Church of Scientology International. URL: http://www.scientology.org/en_US/directory/helping-people/depression/help/ [accessed October 30, 2005] [WebCite Cache ID 95711]
8. Church of Scientology International. URL: http://www.scientology.org/en_US/religion/think-clearly/pg001.html [accessed October 30, 2005] [WebCite Cache ID 95731]
9. Hibbard JH. Engaging health care consumers to improve the quality of care. Med Care 2003 Jan;41(1 Suppl):I61-I70. [Medline: 22432576] [doi: 10.1097/00005650-200301001-00007]
10. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [Medline: 22016241] [doi: 10.1001/jama.287.20.2691]
11. Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. BMJ 2000 Dec 16;321(7275):1511-1515 [FREE Full text] [PMC: 11118181] [Medline: 20568134] [doi: 10.1136/bmj.321.7275.1511]
12. Griffiths KM, Christensen H. The quality and accessibility of Australian depression sites on the World Wide Web. Med J Aust 2002 May 20;176 Suppl:S97-S104 [FREE Full text] [Medline: 22061104]
13. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. JAMA 2001;285(20):2612-2621 [FREE Full text] [Medline: 21262346] [doi: 10.1001/jama.285.20.2612]
14. Risk A, Petersen C. Health information on the internet: quality issues and international initiatives. JAMA 2002;287(20):2713-2715 [FREE Full text] [Medline: 22016244] [doi: 10.1001/jama.287.20.2713]
15. ; Health On the Net Foundation. HON Code of Conduct (HONcode) for medical and health web sites. URL: http://www.hon.ch/HONcode/Conduct.html [accessed 16 July, 2005] [WebCite Cache ID 95865]

16. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [Medline: 99324793]
17. Fallis D, Frické M. Indicators of accuracy of consumer health information on the Internet: a study of indicators relating to information for managing fever in children in the home. J Am Med Inform Assoc 2002;9(1):73-79. [PMC: 11751805] [Medline: 21623452] [doi: 10.1197/jamia.M1131]
18. Kunst H, Groot D, Latthe PM, Latthe M, Khan KS. Accuracy of information on apparently credible websites: survey of five common health topics. BMJ 2002 Mar 9;324(7337):581-582 [FREE Full text] [PMC: 11884323] [Medline: 21881328] [doi: 10.1136/bmj.324.7337.581]
19. Griffiths KM, Christensen H. Website Quality Indicators for Consumers. J Med Internet Res 2005;7(5):e55. [doi: 10.2196/jmir.7.5.e55]
20. Frické M, Fallis D, Jones M, Luszko GM. Consumer health information on the Internet about carpal tunnel syndrome: indicators of accuracy. Am J Med 2005 Feb;118(2):168-174. [Medline: 15694903] [doi: 10.1016/j.amjmed.2004.04.032]
21. Lopez A, Murray CJL. The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Cambridge, MA: Harvard University Press; 1996.
22. 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]
23. ; Centre for Evidence-Based Mental Health. A systematic guide for the management of depression in primary care: treatment. URL: http://www.psychiatry.ox.ac.uk/cebmh/guidelines/depression/depression1.html [accessed 26 October, 2005] [WebCite Cache ID 95881]
24. Tang TT. Quality and relevance of domain-specific search: a case study in mental health. J Inform Retrieval (In press).
25. Robertson SE. On term selection for query expansion. J Doc 1990;46(4):359-364.
26. Robertson SE, Jones KS. Relevance weighting of search terms. J Doc 1976;27(3):129-146.
27. Jorm AF, Christensen H, Griffiths KM, Korten AE, Rodgers B. Help for depression: What works (and what doesn't). Canberra, Australia : Centre for Mental Health Research; 2001.
28. Robertson SE. Okapi at TREC-3. In: Proceedings of TREC-3. NIST Special publications 500-225. Gaithesburg: NIST; 1994.
29. Brin S, Page L. Anatomy of a large-scale hypertextual web search engine. Proceedings of the 7th International World Wide Web conference 1998;7:107-117.
30. SPSS for Windows, Release 12.0.1 SPSS: Chicago; 2003.
31. SISA online statistical analysis. URL: http://home.clara.net/sisa/correl.htm [accessed May, 2005] [WebCite Cache ID 95896]
32. Currò V, Buonuomo PS, Onesimo R, De Rose P, Vituzzi A, Di Tanna GL, et al. A quality evaluation methodology of health web-pages for non-professionals. Med Inform Internet Med 2004 Jun;29(2):95-107. [Medline: 15370990] [doi: 10.1080/14639230410001684396]
33. Suarez HH, Hao X, Chang IF. Searching for information on the Internet using the UMLS and Medical World Search. Proc AMIA Annu Fall Symp 1997:824-828. [Medline: 9357740]
34. Price SL, Hersh WR. Filtering Web pages for quality indicators: an empirical approach to finding high quality consumer health information on the World Wide Web. Proc AMIA Symp 1999:911-915 [FREE Full text] [Medline: 20033003]

Abbreviations

AQA: automated quality assessment procedure
TSVs: Term Selection Values

submitted 04.10.05; peer-reviewed by K Douw, E Bernstam, D Fallis; comments to author 14.10.05; revised version received 31.10.05; accepted 03.12.05; published 30.12.05.

Please cite as:
Griffiths KM, Tang TT, Hawking D, Christensen H
Automated Assessment of the Quality of Depression Websites
J Med Internet Res 2005;7(5):e59
URL: http://www.jmir.org/2005/5/e59/
doi:10.2196/jmir.7.5.e59
PMID:16403723

Original Paper

Comparative Case Study of Two Biomedical Research Collaboratories

Titus KL Schleyer[1], DMD, PhD; Stephanie D Teasley[2], PhD; Rishi Bhatnagar[3], MHCI

[1]Center for Dental Informatics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA
[2]Collaboratory for Research on Electronic Work, School of Information, University of Michigan, Ann Arbor, MI, USA
[3]True Commerce Inc, Wexford, PA, USA

Corresponding Author:
Titus KL Schleyer, DMD, PhD
Center for Dental Informatics
School of Dental Medicine
University of Pittsburgh
3501 Terrace St
Pittsburgh, PA 15261
USA
Phone: +1 412 648 8886
Fax: +1 412 648 9960
Email: titus@pitt.edu

Abstract

Background: Working together efficiently and effectively presents a significant challenge in large-scale, complex, interdisciplinary research projects. Collaboratories are a nascent method to help meet this challenge. However, formal collaboratories in biomedical research centers are the exception rather than the rule.

Objective: The main purpose of this paper is to compare and describe two collaboratories that used off-the-shelf tools and relatively modest resources to support the scientific activity of two biomedical research centers. The two centers were the Great Lakes Regional Center for AIDS Research (HIV/AIDS Center) and the New York University Oral Cancer Research for Adolescent and Adult Health Promotion Center (Oral Cancer Center).

Methods: In each collaboratory, we used semistructured interviews, surveys, and contextual inquiry to assess user needs and define the technology requirements. We evaluated and selected commercial software applications by comparing their feature sets with requirements and then pilot-testing the applications. Local and remote support staff cooperated in the implementation and end user training for the collaborative tools. Collaboratory staff evaluated each implementation by analyzing utilization data, administering user surveys, and functioning as participant observers.

Results: The HIV/AIDS Center primarily required real-time interaction for developing projects and attracting new participants to the center; the Oral Cancer Center, on the other hand, mainly needed tools to support distributed and asynchronous work in small research groups. The HIV/AIDS Center's collaboratory included a center-wide website that also served as the launch point for collaboratory applications, such as NetMeeting, Timbuktu Conference, PlaceWare Auditorium, and iVisit. The collaboratory of the Oral Cancer Center used Groove and Genesys Web conferencing. The HIV/AIDS Center was successful in attracting new scientists to HIV/AIDS research, and members used the collaboratory for developing and implementing new research studies. The Oral Cancer Center successfully supported highly distributed and asynchronous research, and the collaboratory facilitated real-time interaction for analyzing data and preparing publications.

Conclusions: The two collaboratory implementations demonstrated the feasibility of supporting biomedical research centers using off-the-shelf commercial tools, but they also identified several barriers to successful collaboration. These barriers included computing platform incompatibilities, network infrastructure complexity, variable availability of local versus remote IT support, low computer and collaborative software literacy, and insufficient maturity of available collaborative software. Factors enabling collaboratory use included collaboration incentives through funding mechanism, a collaborative versus competitive relationship of researchers, leadership by example, and tools well matched to tasks and technical progress. Integrating electronic collaborative tools into routine scientific practice can be successful but requires further research on the technical, social, and behavioral factors influencing the adoption and use of collaboratories.

(J Med Internet Res 2005;7(5):e53) doi:10.2196/jmir.7.5.e53

KEYWORDS

Biomedical research; collaboration; collaboratories; community networks; information technology; interdisciplinary research

Introduction

Collaborating across geographic and disciplinary boundaries presents research initiatives with an unprecedented challenge in terms of communication and collaboration [1-3]. Meeting these challenges successfully requires unconventional tools and novel scientific work practices [4,5]. The "collaboratory" [6-8] has emerged as a concept for an infrastructure that supports new methods for collaboration using electronic communication networks. In a National Science Foundation workshop held in 1989, William Wulf proposed that "integrated, tool-oriented computing and communications systems to support scientific collaboration...can be called 'collaboratories.' Collaboratories [are]...centers without walls, in which the nation's researchers can perform their research without regard to geographical location, interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries" [7].

Since the concept was initially proposed, collaboratories have become more widely known and adopted in biomedical research [9,10]. Examples of current and past biomedical research and development projects that included formal collaboratories include the following:

- the Biomedical Informatics Research Network (BIRN), which enables data sharing across neuroimaging databases throughout the United States [11]
- the Biological Collaborative Environment (BioCoRE), which is a collaborative research environment for molecular modeling and simulation [12]
- the Molecular Modeling Collaboratory, which is centered around the development, deployment, and use of a highly extensible, interactive molecular modeling software [13]
- the National Laboratory for the Study of Rural Telemedicine, which established the Virtual Hospital, a digital health sciences library and multimedia information integrator providing just-in-time access to information for medical practice, continuing education, and patient education [14]
- the Visible Human Project (VHP), which created complete, anatomically detailed, three-dimensional representations of the normal male and female human bodies [15]

In this paper, we report on the implementation and evaluation of two collaboratories funded by the National Institutes of Health (NIH) that were created to support distributed research centers in biomedicine. We use the term "center" to refer to the center grant as a whole, and the term "collaboratory" to denote the electronic infrastructure that supports communication and collaboration within each center. In contrast to the projects referenced above, most biomedical research centers do not include a formal collaboratory. Many center directors, being unfamiliar with electronic collaborative tools, simply expect traditional methods, such as phone, fax, email, and occasional face-to-face meetings, to support effective and efficient work toward the project objectives. While centers using more traditional communication methods reduce the technical complexity of operating the center, opportunities for more efficient, effective, and novel collaboration through new electronic tools are lost.

The main purpose of this paper is to evaluate comparatively two collaboratories that used off-the-shelf tools and relatively modest resources to support the scientific activity of two biomedical research centers. We first describe the two centers, their goals, and their institutional participants and personnel. Next, we discuss the requirements for collaboration and communication within each center and how we supported these requirements using commercially available electronic tools. Finally, we present selected utilization and outcomes data and conclude with a discussion of barriers and enablers that affected the technology adoption within each collaboratory.

The main goal of this report is to help stakeholders in geographically distributed research centers understand the potential applications of a collaboratory and how to implement one using off-the-shelf tools. Two secondary goals are (1) to provide collaboratory architects with guidance on requirements definition, tool selection, implementation, and evaluation, and (2) to contribute to the growing literature on collaboratory design and implementation [16-18].

The collaboratories we describe were funded as part of the Great Lakes Regional Center for AIDS Research (HIV/AIDS Center) and the New York University (NYU) Oral Cancer Research for Adolescent and Adult Health Promotion Center (Oral Cancer Center). Both centers were large-scale, cooperative research projects funded by the NIH that focused on a single, complex, biomedical research problem. The HIV/AIDS Center, which is no longer operating, focused on HIV biology, immunology, vaccines, therapeutic trials, and behavioral science, and it included four academic institutions in the Midwestern United States. The Oral Cancer Center is currently addressing the reduction of health disparities in oral cancer and encompasses ten institutions. Table 1 provides an overview of both research projects.

Table 1. Summary of the Great Lakes Regional Center for AIDS Research (HIV/AIDS Center) and the NYU Oral Cancer Research for Adolescent and Adult Health Promotion Center (Oral Cancer Center)

Center	HIV/AIDS Center	Oral Cancer Center
Research topic	HIV/AIDS	Health disparities in oral cancer
Major organizational components	• 7 research areas • 8 cores (administrative, clinical research, nonhuman primate model, genomics and proteomics, single-cell imaging and analysis, immunology resource, biocomplexity, and collaboratory)	• 4 research studies • 3 cores (administrative, biostatistics, and collaboratory)
Research studies	1. HIV molecular biology 2. HIV/AIDS pathogenesis research 3. Epidemiology and natural history 4. Opportunistic infections and AIDS-related malignancies 5. Vaccine and other prevention research and development 6. Therapeutic research and development 7. Disease manifestations and metabolic complications	1. Risk factors for oral epithelial dysplasia 2. Oral cancer detection: current and emerging technologies 3. Cancer screening and research subject participation by minorities 4. Personalized risk feedback in dental clinic smokers
Number of		
• principal investigators (PIs)	12 (cores)	7 (4 studies and 3 cores; one study and one core are directed by the same PI)
• research personnel (including PIs)	117	15
• administrative personnel	4	9
Participating institutions	1. Northwestern University 2. University of Michigan 3. University of Minnesota 4. University of Wisconsin	1. Boston University 2. Howard University 3. Johns Hopkins University 4. Memorial Sloan-Kettering Cancer Center 5. New York University 6. Puerto Rico Health Department 7. Tuskegee University 8. University of Alabama/Birmingham 9. University of Pittsburgh 10. University of Puerto Rico
Project duration	9/1998 to 8/2003	8/2001 to 7/2008
Budget	$6.75 million (including $559000 for collaboratory)	$8.3 million (including $604000 for collaboratory)
Funded by	National Cancer Institute and National Institute of Allergy and Infectious Disease	National Institute of Dental and Craniofacial Research

The HIV/AIDS Center was comprised of eight cores engaged in seven research programs. The center's mission was to promote multidisciplinary AIDS research and to engage more scientists in developing more effective measures to prevent and treat HIV infection. Based on several proposed research areas, the center created an infrastructure in which new projects were developed and supported. The program was originally funded for four years (1998-2002) and received an additional year of bridging funds in 2003. Competitive renewal applications were unsuccessful, leading to the dissolution of the center in September 2003.

In contrast to the more developmental focus of the HIV/AIDS Center, the Oral Cancer Center clearly defined four research studies prior to the start of the project; a fifth study will be developed later in the project period. Each of the study proposals clearly framed research questions and methods and described participating research personnel, infrastructure, and budgets. The four research studies are supported by the administrative, biostatistics, and collaboratory cores. While some of those institutions are geographically close (such as the Memorial Sloan-Kettering Cancer Center and NYU), others are quite remote (such as the University of Puerto Rico). The project is funded until 2008.

Both centers had a similar governance structure. They were administered by a lead institution (Northwestern University for the HIV/AIDS Center and NYU for the Oral Cancer Center) and were guided and managed by the group of principal investigators. Each center was advised by an external advisory committee composed of leading scientists in the field and a representative of the funding agency. The principal investigators were responsible for the day-to-day operations of their respective projects.

Unlike most research centers, both the HIV/AIDS Center and the Oral Cancer Center proposed a formal collaboratory in their grant application. The principal investigator for the HIV/AIDS Collaboratory was S. Teasley (University of Michigan), and for the Oral Cancer Collaboratory, T. Schleyer (University of Pittsburgh). Both individuals participated in their respective centers as full members of the scientific and administrative leadership.

In summary, the centers resembled each other in the following ways: their multidisciplinary approach to a single, complex research question; the involvement of several geographically distributed institutional participants; the inclusion of a dedicated collaboratory core; the size of the budget; and the length of the funding period. The major difference was that the HIV/AIDS Center built a platform to develop projects, while the Oral Cancer Center is focusing on the completion of predefined projects.

Methods

Needs Assessment and Initial Requirements Definition

Both collaboratories were developed by conducting a needs assessment and defining requirements; researching, evaluating and selecting off-the-shelf collaboration tools; creating custom resources, such as websites, when needed; and implementing and evaluating the collaboratory. To understand the specific needs of investigators and projects, the collaboratory staff in each center reviewed the grant application in detail and interviewed each principal investigator and key research personnel. These semistructured interviews addressed questions about tasks related to projects, interaction between project teams and center members, the project-related information generated and/or managed, and other center characteristics. In addition, we assessed the local computing infrastructure as well as the software applications used by each investigator, both for desktop computers and personal digital assistants. In both centers, we also conducted contextual inquiry [19] sessions with selected personnel.

Tool Evaluation and Implementation

Once the requirements for a collaborative activity were sufficiently defined, technical staff researched and evaluated existing tools. This typically involved compiling lists of commercial and open-source applications, matching product features against requirements, and testing selected products.

Once a product that satisfied a set of requirements was found, we implemented it in a pilot installation with selected center members. This approach allowed us to address most implementation and functionality issues before a large-scale rollout. Typically, the technical staff of each collaboratory worked with remote technical support to install and configure the tools on scientists' desktops, conduct site-to-site pilot sessions or tests to ensure smooth functioning, and to train the research personnel. In this phase, technical staff members used collaborative tools, such as Timbuktu Pro (Netopia, Inc., Emeryville, CA, USA) and NetMeeting and Remote Desktop Connection (both Microsoft, Inc., Redmond, WA, USA). Scientists were given enough practice and on-site technical support to become comfortable with each tool as it was rolled out.

Collaboratory Evaluation

The evaluation of the collaboratory implementation used three main approaches. First, we collected utilization data through manual and automatic data collection methods (such as Web logs). For example, for Web conferencing, we tracked parameters such as the number of sessions, participants, and participating computers; for collaborative applications, the utilization data included the number of participants, their usage of tools (such as calendars and meeting tools), and the number of shared files. Second, we conducted brief surveys to assess participants' experience with and attitudes toward the tools. Third, we evaluated each collaboratory as participant observers [20] and engaged in repeated discussions with principal investigators and other research personnel, gaining valuable contextual data. As a final step in the evaluation, we identified barriers and enablers that affected the outcomes of the respective collaboratory implementations. The behavioral research within the HIV/AIDS Center was approved by the University of Michigan (IRB Protocol #B03-00001782) and, within the Oral Cancer Center, by the University of Pittsburgh (IRB Protocols #020722 and #0309076).

Results

Needs Assessment and Initial Requirements Definition

None of the investigators, research staff, or administrators in either center had experience with collaborative tools except email and locally shared data stores. Some individuals had participated in videoconferences, typically using PolyCom (Polycom Inc., Pleasanton, CA, USA). Therefore, early in Year 1 in both centers, we conducted educational sessions for the principal investigators and key research staff on the collaboratory concept and corresponding tools.

Within the HIV/AIDS Center, the needs assessment identified two activities as most important for supporting existing collaborations and getting new collaborations started. First, the scientists needed to run distributed lab meetings that would allow conversation over shared data, including, for example, images from a specialized microscope located at only one of the sites. The expectation for this activity was that it be fully interactive so that participants, from few to many, could interact with each other in real time. The importance of high-quality, real-time interactions has been shown to be important for scientific research [21]. Second, the scientists wanted a way to broadcast seminars to share information from experts inside and outside the center.

The needs assessment of the Oral Cancer Center suggested that the requirements for its collaboratory were different. One main objective was to facilitate interaction and tasks among the center participants at large, mostly from an administrative perspective. The second objective was to support the work in each project group. In this case, the requirements centered on facilitating small group communication; sharing protocols, raw research data, and analyses; and aiding workflow. The four research projects, however, differed significantly in their goals and objectives, operations, and personnel roles within the groups. For instance, in the research project on cancer screening and research subject participation by minorities, the initial work was highly sequenced and was either performed by one or two individuals at a time or by a group of research personnel (such as telephone interviewers) who required no support with collaborative tools. By contrast, the research project on personalized risk feedback in dental clinic smokers was highly interactive and data intensive. In this project, the research personnel at Memorial Sloan-Kettering Cancer Center (MSKCC) (who designed the study and analyzed the data) and the clinical personnel at NYU (who handled all patient interactions) interacted frequently and intensively through email, telephone, and face-to-face meetings. The other two groups suffered operational delays, partially due to several regulations of the Health InsurancePortability and Accountability Act coming into effect, and were therefore less active during the first phase of requirements definition.

Our contextual inquiry sessions with the project participants at MSKCC, NYU, and the University of Puerto Rico provided a detailed picture of information management across participating institutions. In general, information was managed in a highly fragmented fashion and in several different computing environments/applications. Members typically worked on several computers (such as home, office, and laptop) and maintained project-related and other work-related information in several places (such as Yahoo Calendar [Yahoo! Inc., Sunnyvale, CA, USA], personal digital assistants, application programs such as Netscape Communicator [AOL Inc., Dulles, VA, USA], and corporate email servers). While all subjects used MS Windows, they did not all use the same applications for the same tasks (eg, Netscape Communicator and Outlook were both used for email). We observed several breakdowns in the way information was shared among individuals. For instance, group meetings were not always recorded in all personal calendars, creating scheduling and coordination problems. In Puerto Rico, the unreliability and limited bandwidth of network connections made it routinely necessary to work around these obstacles.

It is important to note that needs assessment and requirements definition extended (and, in the case of the Oral Cancer Center, are extending) throughout the centers' funding period. Collaboratory staff closely monitored how the activities in their respective center evolved and continually evaluated opportunities for support through collaborative software.

Tool Evaluation and Implementation

Table 2 provides a summary of representative collaborative requirements and the selected tools.

Table 2. Collaborative requirements, sample tasks, and corresponding collaborative tools for the centers

Requirement	Center		Representative Tasks	Products Implemented
	HIV/AIDS	Oral Cancer		
Provide a center-wide website	X	X*	• publish reports, announcements, and member database • launch collaborative applications	• HTML • PHP • mySQL
Install and support collaboratory tools	X	X	• install software applications remotely • train end users one-on-one • troubleshoot end user problems	• Microsoft NetMeeting, Timbuktu Pro • Microsoft Remote Desktop Connection (Oral Cancer Center only)
Train groups in use of collaboratory tools	X	X	• introduce participants to tool concept and functionality • practice using features of tools	• Microsoft NetMeeting, Timbuktu Pro • CentraNow, Genesys Web/telephone conferencing (Oral Cancer Center only)
Manage group meetings and associated information		X	• schedule meetings • record agenda and minutes	• Groove Meetings tool
Conduct real-time, small group meeting	X	X	• share textual and numerical data • share analysis results (eg, SPSS files) • view histology slides through telemicroscopy (HIV/AIDS Center only)	• NetMeeting, Timbuktu Conferencing, Virtual PC (HIV only) • Genesys Web/telephone conferencing (Oral Cancer Center only)
Manage study patient appointments		X	• schedule patients for multiple clinical appointments	• Groove Calendar tool
Access research database and study-related documents		X	• manage research data in MS Access database • co-author study protocols	• Groove File Sharing tool
Broadcast seminars	X		• Share prepublication data and study progress	• Placeware Auditorium • iVisit

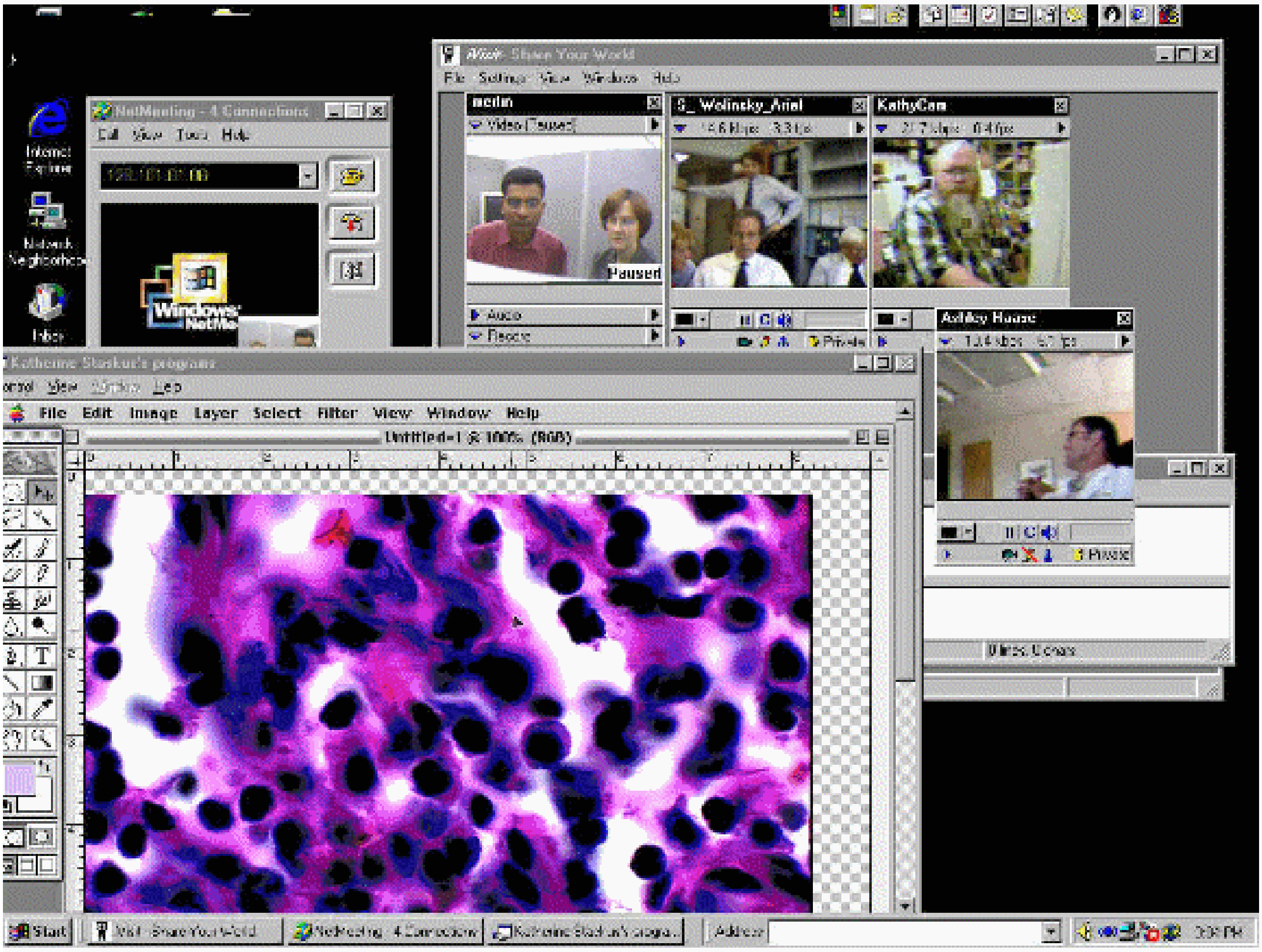
[*] Implementation of requirement suspended

HIV/AIDS Center

Based on the initial set of requirements, the HIV/AIDS Center implemented a comprehensive public/private website, several tools for synchronous collaboration, and Web conferencing for virtual seminars. The website offered progress reports about the research collaborations, descriptions of core services, and a searchable database of all existing members. The website was also used for administrative tasks (eg, registering members, making announcements of upcoming events, distributing applications for developmental grants, archiving center presentations, and providing help documents for the collaboratory tools), launching collaboratory applications for meetings and presentations, and evaluating the center’s activity (eg, collecting survey data, recording observations, and creating usage logs).

Point-to-point, real-time document, image, and equipment sharing was supported by Microsoft NetMeeting (Microsoft, Inc., Redmond, WA, USA). Because NetMeeting was not available on the Macintosh platform, Macintosh users used Timbuktu Conference (Netopia, Inc., Emeryville, CA, USA), and later, Virtual PC (Microsoft, Inc., Redmond, WA, USA). PlaceWare Auditorium (now owned by Microsoft, Inc. and marketed as Live Meeting), a Web-based presentation tool, was selected for virtual presentations. Since the quality of voice over IP (VoIP) connections was insufficient at the time of the HIV/AIDS project, we used telephones for audio during online sessions. On occasion, we used iVisit (iVisit, Santa Monica, CA, USA) to provide video in conjunction with NetMeeting and PlaceWare (Figure 1).

Figure 1. Screenshot of a virtual lab meeting in the HIV/AIDS collaboratory (the bottom left window displays the tissue sample being discussed by the participants shown in the four video feeds in the window at the right)



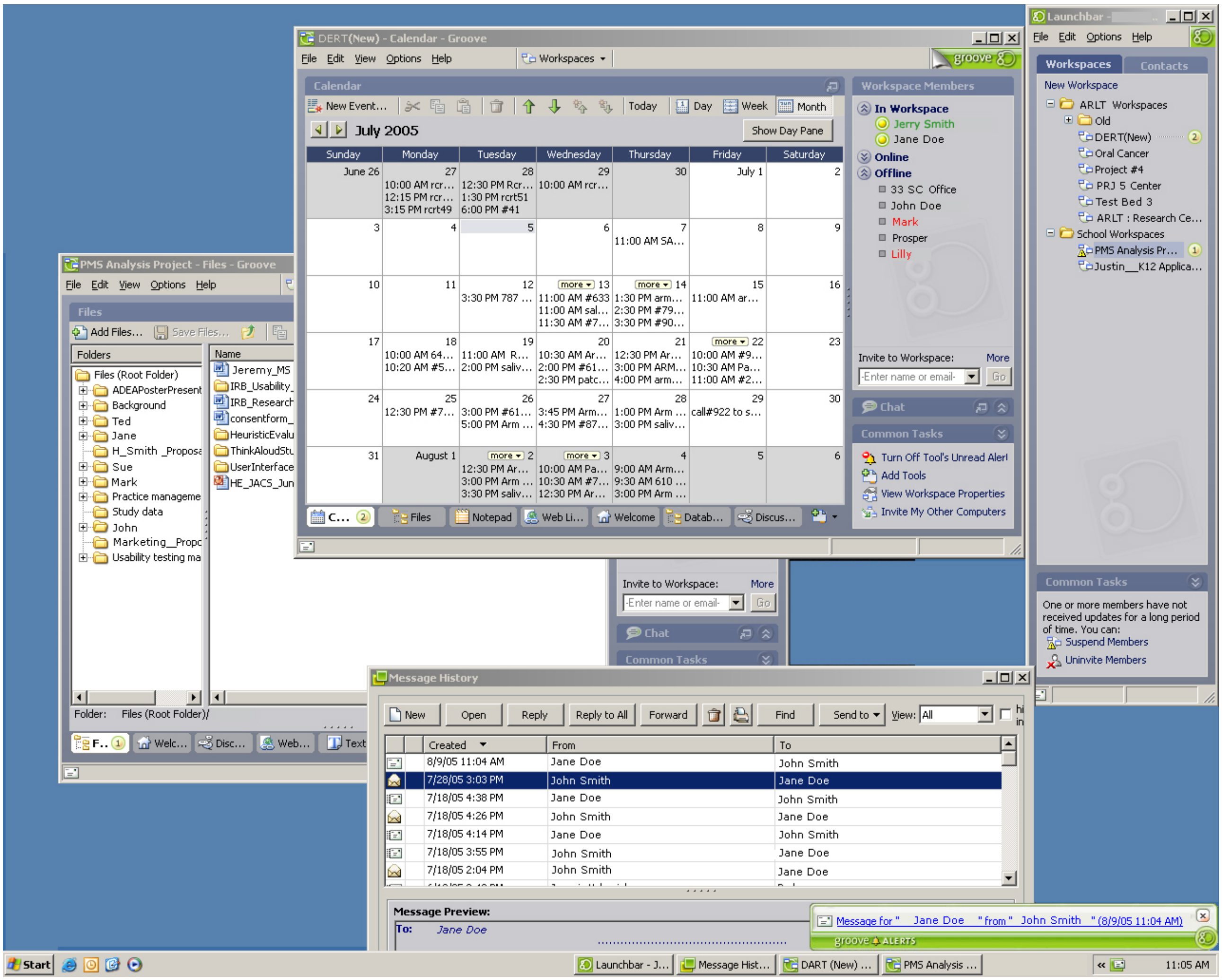
In order to implement these tools, the HIV/AIDS Collaboratory staff identified and trained a local support person at each of the four participating institutions. Although there was great variability in the expertise of the local support staff, the tools were successfully installed and tested before the project principal investigators used them. In addition, the regional nature of the HIV/AIDS collaboratory (including the fact that the collaboratory staff resided at one of the four member institutions) allowed collaboratory personnel to visit sites relatively easily when needed.

Oral Cancer Center

Just like the HIV/AIDS Center, the Oral Cancer Center initially focused on supporting communication and collaboration among members of the Oral Cancer Collaboratory as a whole. Based on the discussions with project personnel and the review of the Oral Cancer Collaboratory grant application, the collaboratory core personnel had a good understanding of the structure and workings of the center at large. On the other hand, gaining an understanding of each research project required a much longer period of time. In addition, because the individual research projects were just starting, their personnel, operational procedures, and infrastructure were still in development. We therefore developed a prototypical website that could function as the administrative “hub” for the center. After an initial period of high interest, it became clear that interaction among all center participants became less important than the increasingly intensive work on the research projects. We therefore suspended work on the center-wide website in order to focus on supporting each individual research group.

We evaluated several groupware applications, such as Lotus Notes (IBM, Armonk, NY, USA), eRoom (EMC Documentum, Pleasanton, CA, USA), Groove (Microsoft, Inc., Redmond, WA, USA), and Hyperwave (Hyperwave AG, München, Germany). We chose Groove (Figure 2) for a pilot implementation with the MSKCC/NYU research group for several reasons. Groove is secure, peer-to-peer collaborative software that integrates a wide variety of collaborative tools (such as file sharing, threaded discussions, Web links, document review, and calendar) into a single workspace (see Figure 2). An administrator can choose which tools are available in a particular workspace and therefore match the feature set to group requirements. Groove is relatively well integrated with the MS Office suite (Microsoft, Inc., Redmond, WA, USA), the application environment used by all of the Oral Cancer Center members. Due to its peer-to-peer architecture, Groove’s administrative overhead is much lower than that of some server-based applications (eg, Lotus Notes). Using NetMeeting, we trained the study personnel remotely in Groove functions relevant to their project tasks. The shared file area was the repository for patient data, and both MSKCC and NYU accessed and modified the same database. The shared calendar served as a tool to record past, current, and future clinical appointments for study participants. Groove’s meeting tool was intended to provide a facility for organizing and recording meetings. A member of the collaboratory core managed the group workspace in Groove, monitored feature usage, and took weekly snapshots of the workspace.

Figure 2. Screenshot of a desktop with two open Groove workspaces (the window on the right shows all available workspaces on this computer; the top window displays the calendar tool of the DERT(New) workspace; the background window shows the file sharing tool of the PMS Analysis Group workspace; the bottom center window contains the Groove Message History; messaging allows users to exchange short messages and can be configured to display notifications such as the one displayed in the lower right corner of the screen)



Created	From	To
8/9/05 11:04 AM	Jane Doe	John Smith
7/28/05 3:03 PM	John Smith	Jane Doe
7/18/05 4:38 PM	Jane Doe	John Smith
7/18/05 4:26 PM	John Smith	Jane Doe
7/18/05 4:14 PM	Jane Doe	John Smith
7/18/05 3:55 PM	John Smith	Jane Doe
7/18/05 2:04 PM	John Smith	Jane Doe

To: Jane Doe

Message for " Jane Doe " from " John Smith " (8/9/05 11:04 AM)

In Year 3, we evaluated Web conferencing software applications for use in joint project meetings for research groups, as well as for monthly updates of project and core principal investigators, and NIDCR personnel. We evaluated and pilot-tested CentraNow (Centra, Inc., Lexington, MA, USA), Elluminate (Elluminate, Inc., Calgary, AB, Canada) and Genesys Meeting Center (Genesys Conferencing, Inc., Reston, VA, USA). All three tools adequately satisfied our synchronous collaboration requirements, which included the ability to present PowerPoint slides, share applications, manage participation (eg, turn-taking for question-and-answer periods), and record online sessions. However, only Genesys could work through a firewall configured with standard ports, and thus was the default choice. To date, the research group for cancer screening and research subject participation by minorities has used Genesys for collaborating on survey data analysis.

In contrast to the HIV/AIDS Center, local support for collaboratory tools was problematic in the Oral Cancer Center. Since the collaboratory personnel were not co-located with any of the project principal investigators or research groups, they often had to interact directly with the principal investigators and research personnel. At some institutions, local IT support was not available; at others, IT support staff did not consider supporting the collaboratory tools as falling within their purview. Support issues frequently had to be solved remotely, often through screen-sharing tools. Occasionally, we provided support on-site during visits to member institutions.

Collaboratory Evaluation

HIV/AIDS Center

Membership in the HIV/AIDS Center was open to anyone engaged in AIDS and AIDS-related research at the four participating institutions. By Year 4, there were 117 registered members of the HIV/AIDS Collaboratory (ranging from 16 to 42 members per site), representing significant growth in membership from the approximate 11 scientists involved in writing the center grant application. During the total funding period, the center sponsored the development of seven major research studies and funded pilot projects for nine junior-level scientists. The research portfolio of the center members increased by 64% in a period when the overall NIH budget for HIV/AIDS increased by 33%.

Virtual Lab Meetings

In the HIV/AIDS Center, a series of virtual lab meetings was established after the initial technical demonstration in the third quarter of Year 1. In the first six months of operation, the collaboratory was used seven times for virtual lab meetings, a rate of about once per month. The average attendance was 4.6 principal investigators (range 3-6) located at three to five computers spread over three of the sites. The principal investigators were typically joined by a number of members from their labs, as well as occasional guests, including NIH administrators, scientific advisory board members, and members of the press.

The scientists valued the virtual lab meetings for the ability to have real-time discussion accompanied by a shared view of a screen and a shared pointer. Specifically, the scientists discussed tissue sample images that were broadcast from a microscope located at one of the sites (see Figure 1), other summary patient data represented as graphics, research protocols, and co-authored documents. The tissue samples and patient data were gathered at several sites, but the expertise for analyzing these samples was located at only one site. As a novel form of collaboration for these scientists, viewing the data together gave them the opportunity to see the data collected at all sites, discuss analyses in real time and, for the tissue images, in the presence of the pathology expert. The scientists also used these meetings to initiate joint studies. The director of the center characterized this as a change from “little science to big science.” By bringing in members of their lab groups to these meetings, senior scientists made their scientific practice more accessible to the junior members of their research teams [22,23].

After the initial use of NetMeeting in the first six months after collaboratory deployment, larger group meetings occurred with less frequency. Specifically, there were five meetings over a 16-month period, organized when individuals generated research results that they wanted to share. However, more one-on-one use of the collaboratory tools emerged to support specific research projects, representing several new cross-site collaborations between pairs of scientists who had not worked together before the center grant began.

Virtual Seminars

The method of broadcasting seminars using the PlaceWare Auditorium software combined with a conference call was also used as a mechanism for sharing pre-published data among the center members. The first virtual seminar occurred at the beginning of Year 2 of the grant. In total, there were nine seminars, four in Year 2 and five in Year 3. Two HIV/AIDS Center members and seven speakers from outside the center presented these seminars. There were an average of 13 computers (range 5-19) logged into each presentation, located at three to four sites. This figure greatly underestimates the number of participants because people were typically assembled in groups around monitors and projection screens. A more accurate picture of seminar participation is derived from a survey administered to the full membership, showing that 73% of HIV/AIDS Collaboratory members who responded to the survey attended at least one, and on average three, virtual seminars. The primary reason for nonattendance was scheduling conflicts (78% of survey respondents), and only 5% reported not attending due to technical difficulties.

The collaboratory personnel in the HIV/AIDS Center attempted to introduce asynchronous data sharing (such as file sharing) to support group work. This effort was not successful, however, because the application deployed, Docushare (Microsoft, Inc., Redmond, WA, USA), was agnostic about content, and the scientists primarily wanted to use a common clinical database.

Oral Cancer Center

Shared Workspaces

The Personalized Risk Feedback in Dental Clinic Smokers study performed by MSKCC and NYU offered multiple opportunities for collaboratory support. A utilization analysis of the Groove workspace after 20 months of use (Table 3) showed that group members used the workspace to collaborate on files and to coordinate clinical appointments. After rapid initial growth, the increase in Word and Excel files leveled off. Recorded patient appointments grew at a steady pace because the clinical personnel used them to closely coordinate their day-to-day work. The group discontinued its use of the meeting tool to organize and record meetings because it did not integrate with the calendar tool in Groove or with the group members' personal calendars. The number of tools used within Groove changed periodically as group members explored new tools and adopted only those that provided value. The number of members in the workspace fluctuated in the first five months, mainly because of technical problems (which led two senior group members to discontinue their use of Groove) and staff turnover. Subsequent to the pilot implementation with the MSKCC/NYU research group, the implementation of Groove with two other research groups is now in progress.

Table 3. Summary analysis of the use of the Groove workspace in the Personalized Risk Feedback in Dental Clinic Smokers study (all numbers represent the total number of objects at the time the workspace snapshot was taken)

Year	Month	Number of Members	Word Files	Excel Files	Patient Appointments	Other Appointments	Meetings	Tools Used
2003	May	3	7	0	0	0	2	4
	June	5	14	0	15	0	4	4
	July	2	22	4	64	1	4	6
	August	7	41	10	106	9	5	7
	September	4	45	11	143	10	6	7
	October	6	46	11	228	14	7	7
	November	6	46	11	318	19	8	6
	December	6	52	11	419	25	8	6
2004	January	6	54	11	512	25	8	6
	February	6	60	13	582	25	8	6
	March	7	63	15	647	26	8	6
	April	7	64	15	661	28	8	6
	May	7	66	15	674	28	8	6
	June	6	65	15	685	30	8	6
	July	6	60	17	701	30	8	6
	August	6	62	17	712	30	8	6
	September	6	64	17	719	30	8	6
	October*	4	40	7	722	30	8	6
	November	7	42	7	725	30	8	6
	December	7	45	9	729	30	8	6

[*] In October 2004, the loss of a password resulted in the temporary loss of members and Word and Excel files.

Web Conferencing

So far, the Cancer Screening and Research Subject Participation by Minorities project has used Genesys five times with six participants each for biweekly meetings. The group is working on analyzing a survey data set and has been using Genesys for sharing analysis strategies and results. To date, the hour-long meetings have typically included three activities. Technical startup, which includes the time until all attendees have logged into the meeting and are ready to participate, initially took between 15 and 20 minutes, and has declined to between 5 and 10 minutes. The time dedicated to discussing shared visual artifacts has increased from about 10 minutes in the first meeting to an average of about 30 minutes. Telephone conference–only phases, which address the analysis strategies, the work plan, and organizational matters, consume the remainder of the time.

The lengthy technical startup phase is due in part to the low general computer literacy of some participants, limited facility with the Genesys client software, and software usability problems. However, the added value of Web conferencing outweighs the current technical drawbacks. The project director commented, “I don't care that it takes us ten minutes to connect—the tool still allows us to do something which we could not do otherwise.”

Discussion

As this evaluation has shown, the two collaboratories described in this paper exhibited some similarities, but they also differed in fundamental ways both in terms of organizational issues and technical needs. Table 4 briefly summarizes those aspects.

Table 4. Comparative overview of the HIV/AIDS Center and the Oral Cancer Center, as well as their collaboratories

	HIV/AIDS Center	Oral Cancer Center
Membership	open	closed
Specification of research projects	loose, developmental	predefined
Collaboration emphasis	general, cross-site	individual project group
Need for		
• instrument sharing	high	none
• real-time collaboration	high	low
• data sharing	emerged late	high
Central collaboratory support	1 research assistant, 75% of collaboratory PI	1 research assistant, 20% of collaboratory PI
Remote IT support	predefined and dedicated	sporadic
Computing platforms	MS Windows, Macintosh, UNIX	MS Windows, Macintosh

The open membership and developmental nature of the HIV/AIDS Center were the primary reasons for the collaboratory focus on enabling general, cross-site collaborations with the capability of both one-on-one and group interactions. In contrast, the Oral Cancer Center was initiated with a much more specific work plan, and, therefore, the collaboratory emphasis was on supporting group work within individual projects. Real-time collaboration in the HIV/AIDS Center used a rich array of tools, resulting in types of collaboration that would not have been possible otherwise (for instance, the real-time discussion of tissue samples among pathologists and clinicians). For the Oral Cancer Center, making sure that the information for working on a particular project was available and up-to-date was initially more important than real-time interaction between co-principal investigators. The need for real-time collaboration only emerged when the first project transitioned to data analysis and interpretation.

Barriers to and Enablers of Collaboration

The comparison of collaboratories also identified several barriers and enablers that affected the outcomes of the respective implementations. These aspects should be addressed through further research. The barriers included the following:

- **Multiple computing platforms:** Cross-platform issues were more problematic in the HIV/AIDS Center (with MS Windows, Macintosh, and UNIX platforms) than in the Oral Cancer Center (MS Windows and Macintosh only), but the collaboratory staff of both centers had to use various workarounds (eg, Virtual PC on the Macintosh) to allow certain members to participate.
- **Network infrastructure complexity:** A major hurdle for the Oral Cancer Center was to find Web conferencing software that worked with the firewall configurations of all participants. For the HIV/AIDS Center, firewalls were less of an issue as the local technical support staff could negotiate with systems administrators to provide access as needed.
- **Variable availability of local versus remote IT support:** The availability of local IT support personnel facilitated the installation and use of collaboratory tools in the HIV/AIDS Center. On the other hand, limited remote support was a major impediment for the Oral Cancer Center.
- **Low computer and collaborative software literacy:** Limited computer literacy with groupware tools hindered participants' collaboratory adoption and use in both centers. While many scientists had some experience collaborating with distant colleagues, these collaborations typically relied on face-to-face meetings and email. Scientists in both centers needed strong incentives and low risk for adopting new ways of conducting their work.
- **Insufficient maturity of collaborative software:** Many collaborative software applications are relatively new products, and sometimes functional limitations, poor interface design, and bugs negatively affected the scientists' perceptions of the value of these tools.
- **Lack of integration with existing application environments:** Collaborative tools should, as much as possible, integrate seamlessly with a user's existing application environment [17]. This barrier was especially obvious for users of Groove in the Oral Cancer Center, as Groove provided stand-alone calendar and messaging functions which did not integrate with other applications. Similarly, in the HIV/AIDS Center, the need to use Virtual PC significantly decreased Macintosh users' enthusiasm for several collaboratory applications.

Despite the problems described above, the comparison of the two collaboratories also identified several factors that promoted collaboratory adoption:

- **Collaboration incentives through funding mechanism:** In both centers, the funding mechanism promoted collaboration, albeit in two different forms. For the HIV/AIDS Center, funding was predicated on the development of projects and new collaborations, while for the Oral Cancer Center, it depended on adequate progress of predefined research projects.
- **Collaborative versus competitive relationship of researchers:** In neither center did competitive pressures among researchers inhibit their readiness to collaborate

with other center members. The HIV/AIDS Center involved researchers with complementary expertise, and the Oral Cancer Center funded research projects with non-overlapping scientific questions. This structure ensured that each scientist's own individual work did not threaten to "scoop" the work of a center colleague.

- **Leadership by example:** In the HIV/AIDS Center, the director led by example as he was an early adopter and one of the highest users of the collaboratory technology in his own center. In addition, several senior scientists not only quickly adopted the technology for their work within the center, but also began to use the tools for other collaborations as well. In the case of the Oral Cancer Center, the director actively sought out opportunities for the use of collaborative tools and strongly encouraged members to participate.
- **Tools matched to tasks:** In general, tools in both collaboratories were relatively well matched with project tasks. For instance, Groove provided the capability to reduce or expand the feature set of a workspace depending on the current needs of a project. On the other hand, in the HIV/AIDS Center, the general functionality of the document sharing application did not match the specific clinical needs, and the tool was therefore not adopted.
- **Technical progress:** During the lifetime of the HIV/AIDS Center, VoIP had not matured sufficiently to be a viable option for multicast audio of acceptable quality. By the start of the Oral Cancer Center, however, VoIP applications were feasible. Conversely, the bandwidth of Internet connections was sufficient to satisfy the performance demands of the collaboratory applications in the HIV/AIDS Center where the research sites were interconnected via Internet2. Members who suffered from "the last mile problem" [24] (eg, wiring in their buildings was not modern enough to capitalize on the bandwidth enabled by Internet2) often solved the problem by participating in the virtual meeting in a colleague's office or in their lab located in a newer facility on campus.

Judging Success of a Collaboratory

Applying the collaboratory model for distributed biomedical research will require further research on the factors related to successful application of the tools to the scientific activity. It is clear from the failure of the HIV/AIDS Center to be refunded that the presence of a collaboratory does not ensure collaboration between all participants. The success of this center in leveraging Wulf's "collaboratory opportunity" [7] was judged differently by the NIH review panel and the center participants. A number of center members felt that their research benefited tremendously from the collaboratory and that they produced work with others with whom they would otherwise not have collaborated. An analysis of the scholarly output of 10 of the scientists who were the original principal investigators for the grant showed that 7 of the 8 new grants funded in the first three years of the center involved collaborators who had not previously been funded together. However, the NIH reviewers believed that the collaborations produced were not enough to merit the level of funding provided by a center. This question of what constitutes enough productivity to justify center-level funding is one that is very much unresolved at the current time and that has been articulated most recently by the National Science Foundation [25]. It is also unclear if the presence of a collaboratory increases the expectation for productivity that may lead to such centers being evaluated with more stringent success criteria than centers that employ more traditional mechanisms for collaboration.

Conclusions

As collaboration technology continues to mature and become more commonplace in scientists' everyday lives, the challenge will be to figure out how to integrate these tools into routine scientific practice in order to increase scientific efficiency and productivity. Disciplinary social norms will undoubtedly drive the pace and breadth of adoption of collaboratory tools [26]. For example, the rise in popularity of bioinformatics tools and the emphasis on exploiting cyber infrastructure for data archiving and management suggest that the capacity for sharing data is an important functionality for collaboration tools. However, without changes in the current reward structure for scientific advancement, incentives for contributing to and using such applications are unclear [27].

Acknowledgments

We gratefully appreciate the assistance of Jeannie Yuhaniak, Andrea Hyde, and Kimberlee Barnhart in the preparation of this manuscript. This work was in part supported by award numbers P30 CA79458, 1U54DE14257-01, and LM-01-001 of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Olson J, Teaseley S, Covi L, Olson . The (currently) unique advantages of collocated work. In: Distributed Work. Cambridge, Mass: MIT Press; 2002:113-135.
2. Olson G, Olson J. Distance matters. Hum-Comput Interact 2001;15(2/3):139-179. [doi: 10.1207/S15327051HCI1523_4]

3. Kraut RE, Egido C, Galegher J. Patterns of contact and communication in scientific collaboration. In: Lawrence Erlbaum Associates, editor. Intellectual Teamwork: Social and Technological Foundations of Cooperative Work. Hillsdale, NJ: Lawrence Erlbaum Associates; 1990:149-171.
4. Nass S, Stillman B. Large-Scale Biomedical Science: Exploring Strategies for Future Research. Washington, DC: The National Academies Press; 2003.
5. Teasley S, Wolinsky S. Communication. Scientific collaborations at a distance. Science 2001 Jun 22;292:2254-2255. [doi: 10.1126/science.1061619] [Medline: 21317431]
6. Finholt TA, Olson GM. From laboratories to collaboratories: a new organizational form for scientific collaboration. Psychol Sci 1997;8:28-36.
7. Wulf WA. The collaboratory opportunity. Science 1993 Aug 13;261(5123):854-855. [Medline: 93348485]
8. Finholt T. Collaboratories. In: Annual Review of Information Science and Technology. Washington, DC: American Society for Information Science; 2002:74-107.
9. Finholt T. Collaboratories as a new form of scientific organization. Econ Innovation New Tech 2003;12(1):5-25. [doi: 10.1080/10438590303119]
10. Schleyer TK. Collaboratories: leveraging information technology for cooperative research. J Dent Res 2001;80(6):1508-1512. [Medline: 21390341]
11. Biomedical Informatics Research Network (BIRN). URL: http://www.nbirn.net/ [accessed 2005 Aug 24]
12. Biological Collaborative Environment (BioCoRE). URL: http://www.ks.uiuc.edu/Research/biocore/ [accessed 2005 Aug 24]
13. Molecular Modeling Collaboratory. URL: http://www.cgl.ucsf.edu/Research/collaboratory/ [accessed 2005 Aug 24]
14. National Laboratory for the Study of Rural Telemedicine. URL: http://www.vh.org/ [accessed 2005 Aug 24]
15. Visible Human Project (VHP). URL: http://vhp.med.umich.edu/ [accessed 2005 Aug 24]
16. Olson GM, Finholt TA, Teasley SD. Behavioral aspects of collaboratories. In: Electronic Collaboration in Science. Mahwah, NJ: Erlbaum; 2000:1-14.
17. Mandviwalla M, Olfman L. What do groups need? A proposed set of generic groupware requirements. ACM Transactions on Computer-Human Interaction 1994;1(3):245-268. [doi: 10.1145/196699.196715]
18. Sonnenwald D, Whitton M, Maglaughlin K. Evaluation of a scientific collaboratory: research of a controlled experiment. ACM Transactions on Human Computer Interaction 2003;10(2):150-176. [doi: 10.1145/772047.772051]
19. Beyer H, Holtzblatt K. Contextual Design: Defining Customer-Centered Systems. San Francisco, Calif: Morgan Kaufmann Publishers, Inc.; 1998.
20. Dezin N, Lincoln Y. Handbook of Qualitative Research. Thousand Oaks, Calif: Sage; 1994.
21. Shortliffe EH, Patel VL, Cimino JJ, Barnett GO, Greenes RA. A study of collaboration among medical informatics research laboratories. Artif Intell Med 1998;12(2):97-123. [Medline: 98179041] [doi: 10.1016/S0933-3657(97)00045-6]
22. Bell G, Gray JN. The revolution yet to happen. New York: Copernicus; 1997.
23. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation. Cambridge, UK: Cambridge University Press; 1991.
24. Bell G, Gemmell J. On-ramp prospects for the information superhighway dream. Commun ACM 1996;39(7):55-60. [doi: 10.1145/233977.233995]
25. Mervis J. Research funding. Centers of attention: NSF takes fresh look at their proliferation. Science 2005 May 13;308(5724):943-945. [Medline: 15890857] [doi: 10.1126/science.308.5724.943]
26. Who'd want to work in a team? Nature 2003;424(6944):1-1. [doi: 10.1038/424001a] [Medline: 22724493]
27. Arzberger P, Finholt TA. Data and Collaboratories in the Biomedical Community. Report of a Panel of Experts Meeting held September 16-18, 2002 in Ballston, VA.

Abbreviations

AIDS: acquired immunodeficiency syndrome
HIV: human immunodeficiency virus
IT: information technology
MSKCC: Memorial Sloan-Kettering Cancer Center
NIH: National Institutes of Health
NYU: New York University
PI: principal investigator
VoIP: voice over Internet protocol

submitted 10.06.05; peer-reviewed by R Appleyard, G Olson; comments to author 09.07.05; revised version received 25.08.05; accepted 25.08.05; published 25.10.05.

Please cite as:
Schleyer TKL, Teasley SD, Bhatnagar R
Comparative Case Study of Two Biomedical Research Collaboratories
J Med Internet Res 2005;7(5):e53
URL: http://www.jmir.org/2005/5/e53/
doi:10.2196/jmir.7.5.e53
PMID:16403717

Original Paper

Architecture for Knowledge-Based and Federated Search of Online Clinical Evidence

Enrico Coiera[1], MB, BS, PhD; Martin Walther[1], BCompSci, MAI; Ken Nguyen[1], BEng; Nigel H Lovell[2], PhD

[1]Centre for Health Informatics, University of New South Wales, Sydney, Australia

[2]Graduate School of Biomedical Engineering, University of New South Wales, Sydney, Australia

Corresponding Author:
Enrico Coiera, MB, BS, PhD
Centre for Health Informatics
University of New South Wales
Sydney
UNSW 2052 NSW
Australia
Phone: +61 (2) 9385 9026
Fax: +61 (2) 9385 9006
Email: e.coiera@unsw.edu.au

Abstract

Background: It is increasingly difficult for clinicians to keep up-to-date with the rapidly growing biomedical literature. Online evidence retrieval methods are now seen as a core tool to support evidence-based health practice. However, standard search engine technology is not designed to manage the many different types of evidence sources that are available or to handle the very different information needs of various clinical groups, who often work in widely different settings.

Objectives: The objectives of this paper are (1) to describe the design considerations and system architecture of a wrapper-mediator approach to federate search system design, including the use of knowledge-based, meta-search filters, and (2) to analyze the implications of system design choices on performance measurements.

Methods: A trial was performed to evaluate the technical performance of a federated evidence retrieval system, which provided access to eight distinct online resources, including e-journals, PubMed, and electronic guidelines. The Quick Clinical system architecture utilized a universal query language to reformulate queries internally and utilized meta-search filters to optimize search strategies across resources. We recruited 227 family physicians from across Australia who used the system to retrieve evidence in a routine clinical setting over a 4-week period. The total search time for a query was recorded, along with the duration of individual queries sent to different online resources.

Results: Clinicians performed 1662 searches over the trial. The average search duration was 4.9 ± 3.2 s (N = 1662 searches). Mean search duration to the individual sources was between 0.05 s and 4.55 s. Average system time (ie, system overhead) was 0.12 s.

Conclusions: The relatively small system overhead compared to the average time it takes to perform a search for an individual source shows that the system achieves a good trade-off between performance and reliability. Furthermore, despite the additional effort required to incorporate the capabilities of each individual source (to improve the quality of search results), system maintenance requires only a small additional overhead.

(J Med Internet Res 2005;7(5):e52) doi:10.2196/jmir.7.5.e52

KEYWORDS

Evidence-based medicine; clinical decision support systems; information retrieval; meta-search filters

Introduction

Clinicians need to keep up-to-date with the biomedical literature in order to practice according to the best available evidence. However, this has become increasingly difficult as the amount of medical literature a clinician needs to consider grows exponentially [1,2]. As a result, the effort required to find a specific piece of evidence increases year after year [3]. Clinicians typically work under time pressure, which compounds the problem. The need to develop robust methods and tools to support evidence access is now widely recognized. Online evidence retrieval methods are increasingly seen as a core tool in support of evidence-based health care [4]. In the traditional model of online evidence services, clinicians have access to a number of online information sources, such as journals, databases, and Medline, each with its own idiosyncrasies and search interfaces. This means users need to know which resources are most suitable for their current question and how the search query must be formulated for a given resource. Interoperability standards for the efficient dissemination of content are being developed (eg, the Open Archive Initiative [5]), but until the majority of content adheres to such standards, there is still a need to search through heterogenous data sources.

The meta-search engine approach [6,7] addresses many of the limitations of these models by providing a mechanism to search all the available resources at one time and by translating user queries into the respective query languages of each resource. This typically uses a least-common-denominator approach, directly passing on user keywords to different information sources without regard for the specific capabilities or limitations of these resources. For example, a meta-search engine often disregards the rich query language available with some resources in order to simplify the overall meta-search process. Consequently, while the user expects the meta-search to return an integrated set of search results, the reality is that some resources would have been able to perform much better had they been queried individually; the user is unaware of the variations in search quality across the different resources that have been queried for them. Variants of the standard meta-search engine approach have been shown [8] to provide search capabilities beyond the least common denominator but still require users to select the resources they wish to search. One solution to this problem is to "federate" the different resources so that they more genuinely behave as one uniform data source. A federated search system may perform a *syntactic reformulation* of a user query, translating it into queries that have been optimized for the native query language of individual evidence sources. *Semantic reformulation* is also possible [9]. For example, user keywords may be translated into equivalent keywords or phrases using a terminological system.

However, a federated search can still produce an excessive number of candidate documents, or hits, many of them failing reasonable tests of relevance. One way to improve the chance of retrieving clinically relevant information is to pre-program a search system with specialist bibliographic knowledge using *search filters*. Search filters capture expert strategies for searching that are known to improve the precision of searches. For example, Medline offers a small set of "clinical queries," which are pre-defined and validated search filters optimized to retrieve documents that are most likely to be clinically relevant, emphasizing disease etiology, diagnosis, therapy, or prognosis [10,11]. Such search filters are necessarily highly customized to the capabilities of individual information sources and their native search engines. For a federated search system to consistently use search filters, it would need to develop a generalized approach to search filters, or *meta-search filters*. Quick Clinical (QC) [4] is a federated evidence retrieval system designed to meet the specific needs of clinicians. Its design incorporates the novel use of meta-search filters to optimize search strategies, and it is based upon a wrapper-mediator architecture built around a universal query language. This paper describes the system architecture of QC and the technical challenges to the design of online evidence retrieval systems, and it reports on the technical performance of the system from a clinical trial with primary care physicians.

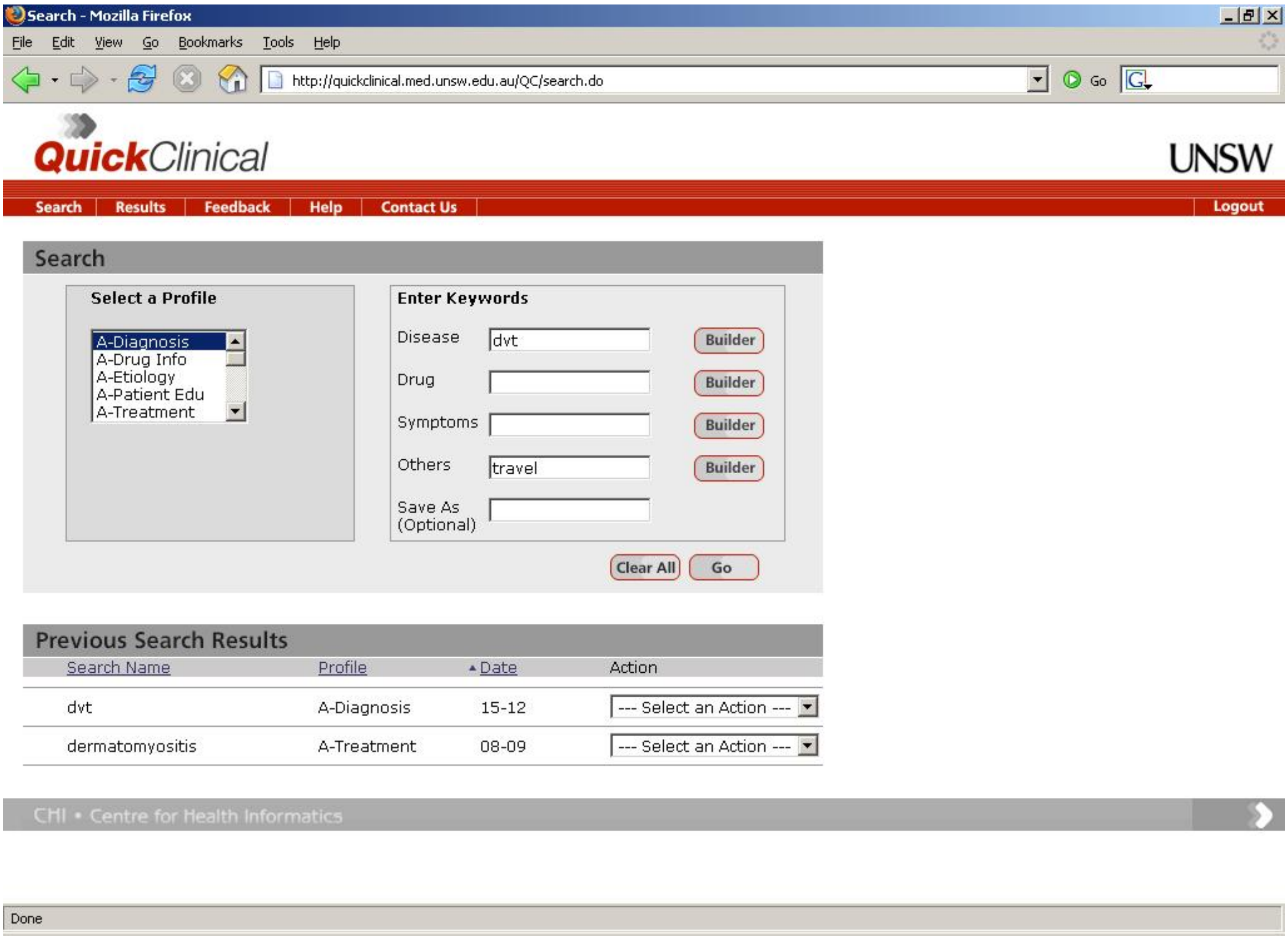
Methods

The Quick Clinical System

User Interface

In the QC user model, a user is presented with a single query interface, which connects to an arbitrarily large number of federated knowledge sources and incorporates query specific meta-search filters called "profiles." QC guides users to first consider the purpose of their search through selection of a profile, and it then asks them to provide specific keywords related to that search task. As a consequence, users are guided through a process that structures their query for them and improves the chances that they will ask a well-formed query and receive an appropriate answer. Figure 1 depicts the QC search interface. On the left hand is a list of search filters that describe typical search tasks and that are customized to the specific information needs of different user groups. Figure 1 shows filters specifically designed for use by primary care physicians.

Figure 1. The QC search query user interface



Search

Select a Profile

- A-Diagnosis
- A-Drug Info
- A-Etiology
- A-Patient Edu
- A-Treatment

Enter Keywords

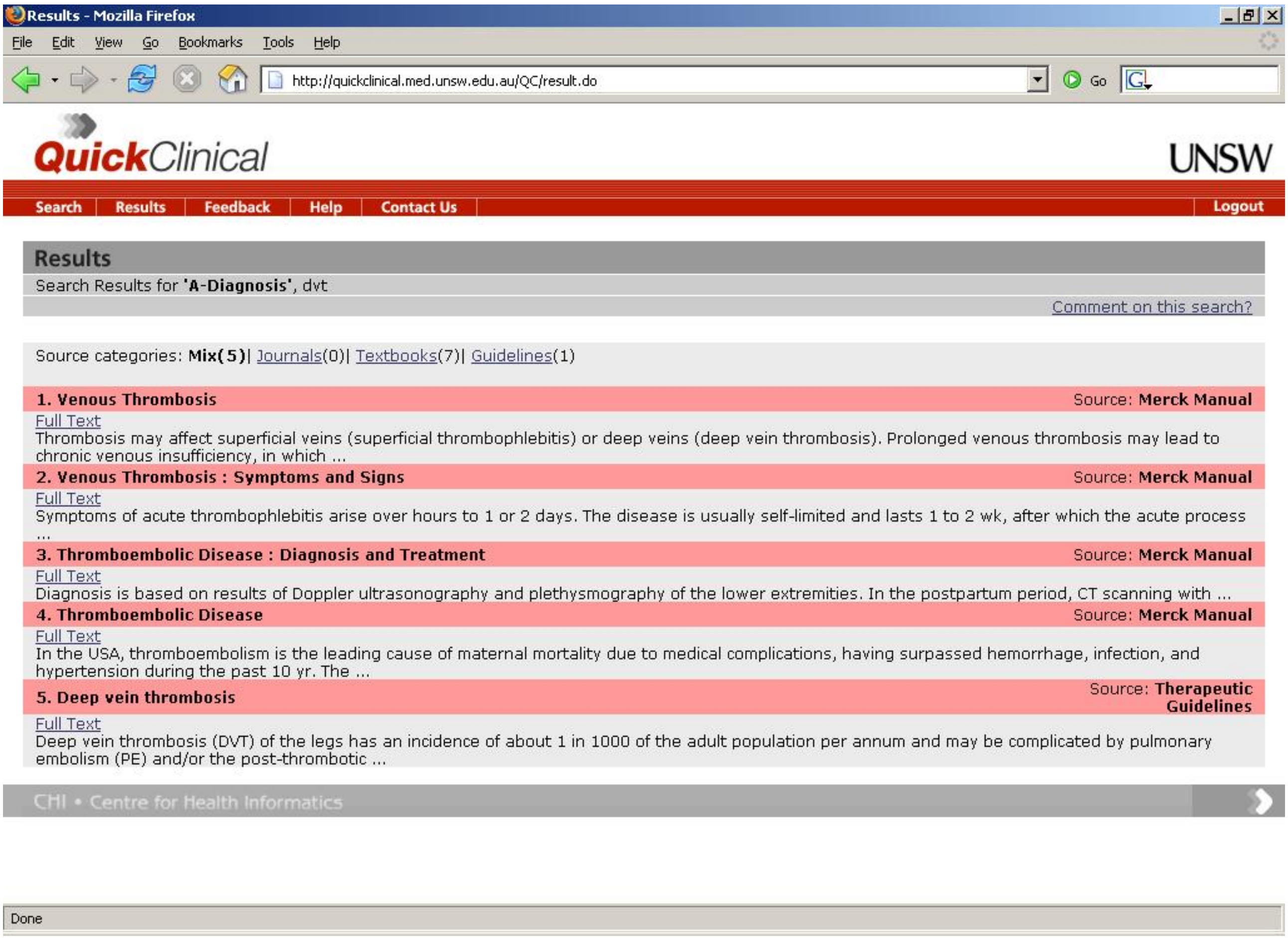
Disease: dvt
Drug:
Symptoms:
Others: travel
Save As (Optional):

Previous Search Results

Search Name	Profile	▴Date	Action
dvt	A-Diagnosis	15-12	--- Select an Action ---
dermatomyositis	A-Treatment	08-09	--- Select an Action ---

In QC, individual profiles are able to define different keyword types, such as “disease,” which describe the keyword classes typically associated with that profile. Thus, on the right of the interface are four fields where users can provide keywords describing the specific attributes of their search. Selection of a different profile may thus alter the keyword types requested from the user for a given search. QC then translates and submits search queries to the sources specified in the chosen profile, collects and processes the results, and presents them to the user as a list of documents (Figure 2). The title of a document is followed by the link and a short abstract of the content. A user can drill-down into a specific group of results by source type (eg, journal articles or guidelines).

Figure 2. Screenshot of a QC results page



Results

Search Results for **'A-Diagnosis'**, dvt

Comment on this search?

Source categories: **Mix(5)**| Journals(0)| Textbooks(7)| Guidelines(1)

1. Venous Thrombosis Source: **Merck Manual**

Full Text

Thrombosis may affect superficial veins (superficial thrombophlebitis) or deep veins (deep vein thrombosis). Prolonged venous thrombosis may lead to chronic venous insufficiency, in which ...

2. Venous Thrombosis : Symptoms and Signs Source: **Merck Manual**

Full Text

Symptoms of acute thrombophlebitis arise over hours to 1 or 2 days. The disease is usually self-limited and lasts 1 to 2 wk, after which the acute process ...

3. Thromboembolic Disease : Diagnosis and Treatment Source: **Merck Manual**

Full Text

Diagnosis is based on results of Doppler ultrasonography and plethysmography of the lower extremities. In the postpartum period, CT scanning with ...

4. Thromboembolic Disease Source: **Merck Manual**

Full Text

In the USA, thromboembolism is the leading cause of maternal mortality due to medical complications, having surpassed hemorrhage, infection, and hypertension during the past 10 yr. The ...

5. Deep vein thrombosis Source: **Therapeutic Guidelines**

Full Text

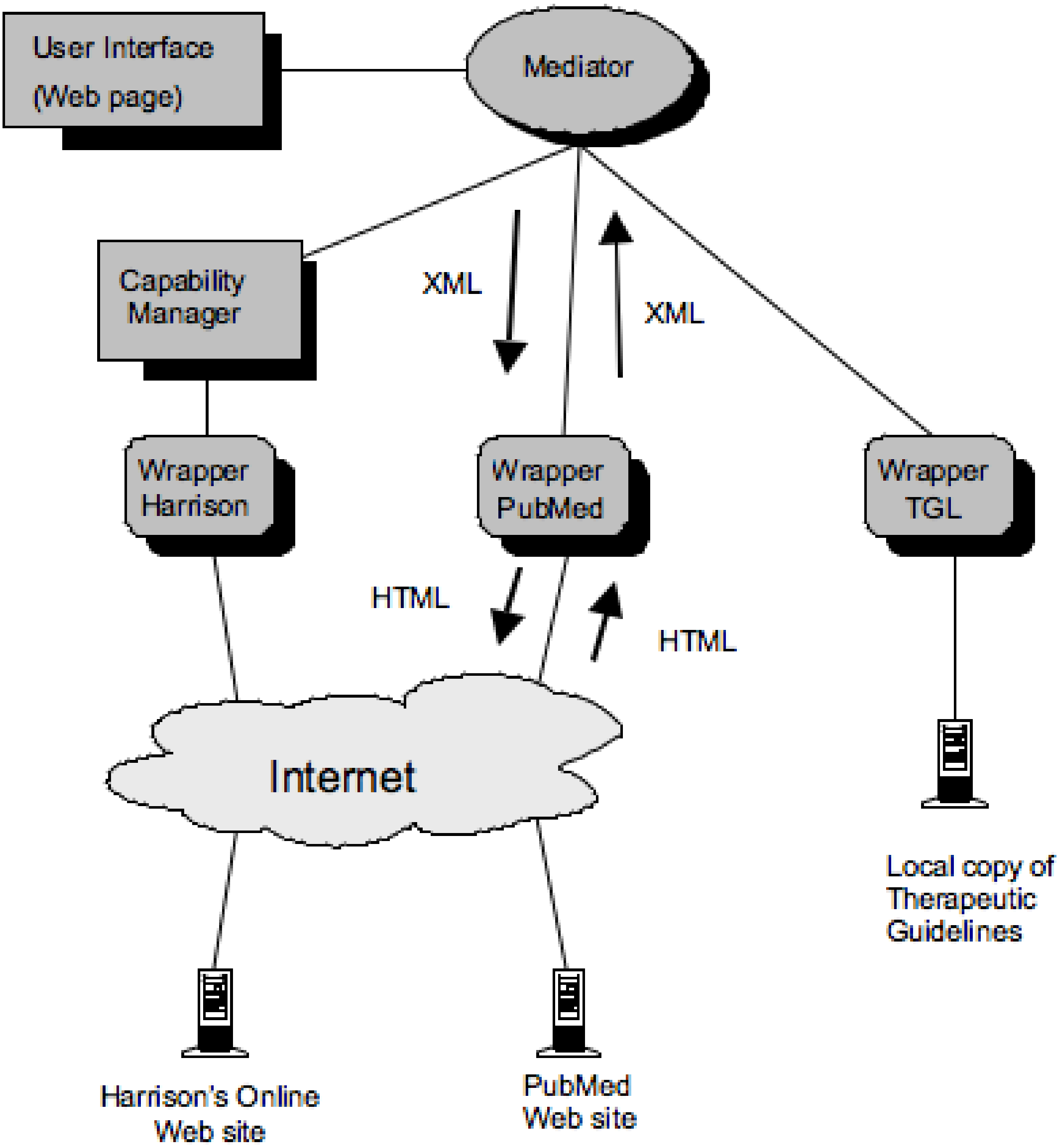
Deep vein thrombosis (DVT) of the legs has an incidence of about 1 in 1000 of the adult population per annum and may be complicated by pulmonary embolism (PE) and/or the post-thrombotic ...

Quick Clinical System Architecture

Overview

Most information sources such as websites, online texts, and databases have their own proprietary search interface, including query language and format for the display of results. Therefore, a federated meta-search engine that wishes to query a number of different information sources needs to first represent a user query using some internal query language [13] and then translate that internal query into the specific query languages of the relevant data sources. A well-documented [12] approach to this problem is to use a “wrapper” (Figure 3), which acts as an adapter between the proprietary language of individual information sources and the internal language used within a meta-search system. In QC, the internal query language is called the unified query language (UQL). Each information source known to QC has its own wrapper that translates queries from UQL into the native language of the source. As a result, internal components of QC only need to know UQL and not the individual query languages of the data sources. System maintenance is also simplified since the introduction of a new data source to the system only requires one new wrapper component to be generated. Once the results of a search are returned by an information source, the information must again be translated into a standard output format for presentation to the user, which, in QC, is called the unified response language (UReL). UReL also allows other components in the system to modify the presentation of search results without needing to understand the presentation format of individual sources (eg, to remove duplicate documents). In Figure 3, a search is initiated from the user interface, which forwards a query (in XML) to the mediator. The mediator splits the query into several subqueries and sends these to the appropriate wrapper (via a capability manager if required). Finally, the wrapper translates the query into the native query language of the data source (eg, in HTML for Web data sources). Similarly, the result from the data source gets translated back into the system’s XML representation and sent back to the user interface.

Figure 3. Architecture overview of Quick Clinical



Unified Query Language

UQL is used to represent queries obtained from users in a consistent internal way, and UQL statements identify query elements such as the external information sources to be searched and a set of search attributes used to delimit the search. For example, UQL expressions can store date range delimiters for a search. UQL also contains statements that indicate whether or not QC needs to process the query further. For example, we may wish to remove duplicate items obtained from different sources. In our current implementation, UQL is implemented using XML. To define the structure of the data within the XML document we use a data type definition (DTD), which allows various internal components of QC to validate the XML data received in the UQL query. The following example illustrates how a UQL query might look in XML.

```
<QUERY keyword = "iron AND deficiency"
  profile = "treatment"
  duplicateRemoval = "yes"
  sortBy = "rank"
  useLexicalVariants = "yes"
  timeout = "20"
  dateRangeBeginDay = "1"
  dateRangeBeginMonth = "1"
  dateRangeBeginYear = "1999" >
 <SOURCE name = "PubMed" />
 <SOURCE name = "Harrison’s online" />
 <SOURCE name = "Merck" />
 <SOURCE name = "MIMS" />
</QUERY>
```

Unified Response Language

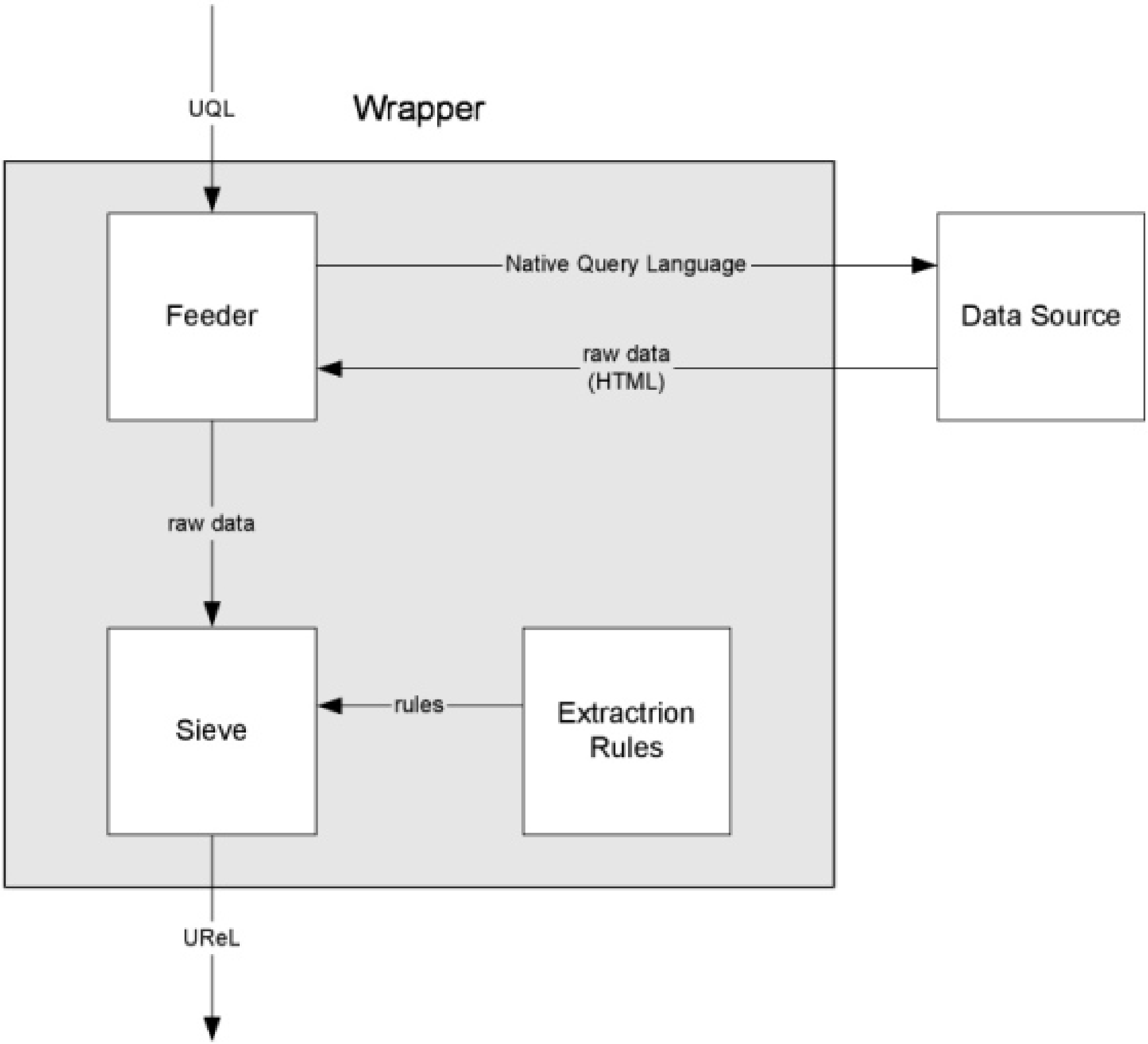
Similarly to the UQL, the unified response language (UReL) is used internally to guide display of information to users, also represented using XML. Each separate result, or “article,” from a source can be broken up into smaller chunks and given meta-data labels to represent the different sections of the data (eg, abstracts from journal articles). Since the majority of sources accessed by QC are journals, the data that are retrieved typically contain document elements such as Title, Author(s), Journal Name, Date of Publication, and the URL where the electronic version of the paper is accessed. Other sources, such as drug descriptions from pharmaceutical compendia, have sections such as Drug Name and Manufacturer. These different document elements, based upon the typical sources QC expects to find, are defined as specific fields in the UReL definition. The following example illustrates how a set of documents retrieved by QC might be represented in UReL.

```
<RESULT>
 <ARTICLE>
  <LINK>
   <HREF>
     http://www.ncbi.nlm.nih.gov:80
     /entrez/query.fcgi?cmd=Retrieve&db=PubMed
     &list_uids=12198020&dopt=Abstract
   </HREF>
   <LINKNAME>Abstract</LINKNAME>
  </LINK>
  <AUTHORLIST>Heath AL,
    Skeaff CM,
    Gibson RS.
  </AUTHORLIST>
  <TITLE>
    Dietary treatment of iron deficiency
  </TITLE>
  <DATE>
   <YEAR>2002</YEAR>
   <MONTH>9</MONTH>
  </DATE>
  <SOURCE>PubMed</SOURCE>
 </ARTICLE>
 <ARTICLE>
  <LINK>
   <HREF>
     http://mims.hcn.net.au
     /ifmx-nsapi/mims-data/?MIval=2MIMS_abbr_pi
     &product_code=288
     &product_name=Ferrum+H+Injection
   </HREF>
   <LINKNAME>More Information</LINKNAME>
  </LINK>
  <AUTHORLIST>
    Sigma Pharmaceuticals Pty Ltd.
  </AUTHORLIST>
  <TITLE>Ferrum H Injection</TITLE>
  <SOURCE>MIMS</SOURCE>
 </ARTICLE>
</RESULT>
```

Wrappers

For every information source known to QC, there is a specific wrapper that translates a UQL query into the native query language and format of the source. The wrapper also extracts the relevant information from the HTML result pages returned by the search engine and re-expresses it in UReL. Figure 4 shows the basic architecture of wrappers in our current system. Each wrapper has three main components: a feeder, extraction rules, and a sieve. The feeder converts the user query into the native query language of the data source. The data source responds to the query and returns HTML raw data. The feeder passes the raw data to the sieve, which converts it to UReL in XML format by using the extraction rules for the data source. The UReL is then sent back via other components to the user interface, which can interpret the XML and display the results.

Figure 4. Wrapper components



Mediator

A key requirement of a multisource information retrieval system is the ability to perform concurrent searches on multiple sources with a single query [6,7]. The mediator addresses this requirement. The mediator first analyzes a query and determines how many sources are to be searched. It then creates a separate search job for each of these sources and forwards the search job to other system components. Additionally, the mediator collects individual results as they arrive and amalgamates them for the user into a single result. By introducing parallelism, the time to perform a search across a number of resources should be reduced to the duration of the slowest source. However, the potential drawback of parallel processing is the increased administration overhead of running multiple parallel processes within a system. As a rule of thumb, we would expect the benefits of parallel execution should increase with the number of sources queried, as response times for Web resources can be many seconds long, and computational execution of processes to manage parallel search are typically much less than one second.

Connection speed and latency of response time from sources are, for practical purposes, nondeterministic in an Internet environment, and a meta-search engine can therefore experience large fluctuations in responses from the same source under different circumstances. Latency is subject to network traffic conditions, making it impossible to guarantee that all resources that are queried at a particular time will respond predictably and equally. To counter this, the mediator has a time-out feature. If a response is not received within the time-out specified by a profile, the mediator will cancel a subsearch and forward all the results currently available from other sources to the user interface. This effectively guarantees a defined response time irrespective of the state of the individual data sources and provides some control over the speed/accuracy trade-off.

Capability Manager

Search capabilities vary considerably between the search engines that QC might wish to interrogate, and some sources will have limitations in their ability to process search queries. One approach to this problem is to try to raise all sources to as high a level of common performance as possible by emulating missing capabilities locally, usually by modifying the query and/or search result [13]. A trivial example is mimicking the ability to perform a Boolean search when a data source does not have this capability. To emulate a Boolean AND, a meta-search engine would perform two parallel individual searches on the source and then itself perform the Boolean operation on the two results.

In QC, a capability manager (CM) is responsible for mimicking a range of search capabilities and is located between the mediator and wrapper. The CM may modify a query and/or the result depending on the capabilities of the sources about to be queried. Capabilities of the CM within the QC system included the following:

- Date-CM: search within a date range
- Duplicate-CM: remove document duplicates
- Sort-CM: sort results by title, author, document rank, or date
- Lexical-CM: expand a search term with lexical variants of the term. A lexical variant is a synonym, pluralization, hyphenation, or other modification that changes the text but not its meaning. Lexical variants are particularly important in the medical domain [14] because many concepts can be expressed in Latin or English (eg, cardio vs heart). Moreover, there is a common confusion between terms in American English versus British English (eg, hemoglobin vs haemoglobin, epinephrine vs adrenaline).

QC uses a stacking mechanism to insert individual CMs into the processing of queries for wrappers and the processing of results from a source. A component called the search planner, containing simple rules, is responsible for stacking the CMs. This means that the sequence of CMs can be ordered to ensure the correct outcome of query or result translations. Theoretically, this corresponds to a composition of operations. A lexical variant CM, for example, has to replace the search terms in the query before the wrapper executes the search. The Date-CM, on the other hand, can only perform its job after the successful execution of the wrapper.

Search Filters

Expert searchers typically will use search strategies that are more likely to accurately locate information, based upon an understanding of the specific capabilities of an evidence source. There is an increasing interest in the writing of search filters which capture such strategies, usually focusing on the major evidence repositories like Medline [10,11]. Search filters are designed for typical clinical queries such as “diagnosis” or “prescribing,” and they are crafted to find evidence most likely to satisfy the query by first selectively searching resources identified to be of high quality and, second, by automatically adding specialist keywords to the general question posed by a user. Within QC, search filters are stored in the profiles function. For example, if a clinician selects the “diagnosis” filter and enters the search term “asthma,” QC can add in the additional terms when it queries Medline [10]:

sensitivity and specificity [MESH] OR sensitivity [WORD] OR diagnosis [SH] OR diagnostic use [SH] OR specificity [WORD]

These terms have been shown to significantly enhance the quality of Medline results, but they are unlikely to be known to a typical clinical user.

Unlike standard search filters, QC profiles are meta-search filters because they encode search filters for multiple different sources. Profiles thus encode expert search strategies that are most likely to answer a certain class of query, and they encode, among other things, the most appropriate content sources to search (Table 1). For a primary care physician, these search profiles might be for diagnosis, prescription, review, and treatment [4], but any set of profiles can be created within QC to meet the specific query types and search contexts of different users. In Table 1, the Treatment profile describes a set of nine separate source-specific search filters, which collectively describe the search strategy believed most likely to retrieve an accurate search result from each resource. The # symbol delimits keyword variables that are to be instantiated with user keywords. For example, #1# represents the keyword type “disease,” and QC’s mediator component will substitute the user-provided keywords for “disease” throughout the profile, prior to sending the query to the individual wrappers for the different sources. More than one search string can be created for an individual source (eg, TGL 1 and TGL 2) as a single strategy may not always retrieve all the relevant documents.

Table 1. Quick Clinical meta-search filters

Source	Search String
TGL1	(#1# AND #2# AND #3# AND #4#) AND+ ("treatment" OR "therapy" OR "therapeutic use")
TGL2	(#1# AND #3#) AND+ ("treatment" OR "therapy" OR "therapeutic use")
HealthInsite3	#1# AND #2# AND #3# AND #4#
HealthInsite4	#1# AND #3#
PubMed5	(#1# ATTR+ [Title] AND #3# ATTR+ [Title] AND #4# ATTR+ [Title] ATTR+ /ther) English 10 years Human
PubMed6	#1# ATTR+ [Title] AND (#3# ATTR+ [Title] OR #4# ATTR+ [Title]) ATTR+ /drug ATTR+ ther English 10 years Human
Merck7	((#1# AND #3# AND #4#) OR (#1# AND #3#)) AND+ ("treatment" OR "therapy")
Harrison's8	Disconnected
Harrison's9	Disconnected

System Platform

The system was constructed using Java, the Struts Web application framework, and a MySQL database and is deployed on a RedHat Linux platform. The user interface (JSP, servlet, and HTML pages) is deployed through an Apache Web server connected to a Tomcat servlet engine. The Apache-Tomcat platform incorporates load balancing and fail-over and is suitable for scalability and large-scale deployment.

Technical Evaluation

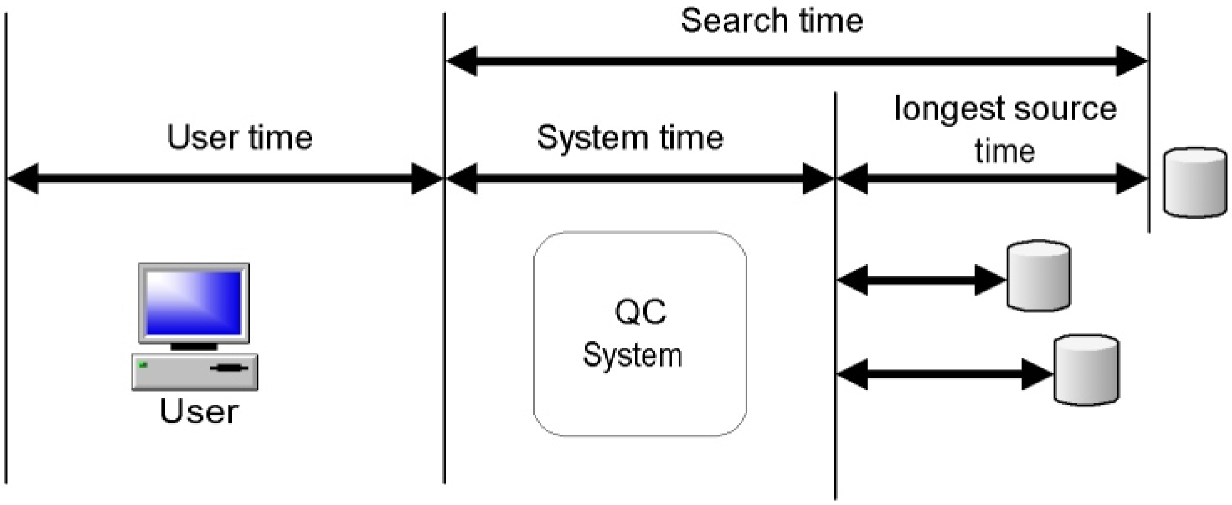
QC has undergone a series of clinical evaluations, which have been reported separately [4,15,16].

In total, 227 family physicians from across Australia participated in a trial of QC. Clinicians who had a computer with Internet access in their consulting rooms were recruited and asked to use QC for 4 weeks in routine care. Each participant was given a personal username and password to access the system. All clinicians completed an online pre-trial survey. QC was configured to search a set of eight sources, including remote sites such as PubMed, online journals such as *BMJ* and the *Medical Journal of Australia (MJA)*, and locally cached sources such as *The Merck Manual* and Therapeutic Guidelines Australia.

For every search, the time from the request arriving at the system to the time when the results were sent back to the user was recorded (Figure 5; search time = system time + slowest source time). Note that there is a cap on search time when the time-out cuts in. Time-outs are search-profile dependent and were set at either 15 or 30 s. The time it took to conduct the search on the individual sources was also recorded. The time taken to send data between QC and the user's computer (user time) is not incorporated in these measurements.

In the following section we report on the technical performance of the architecture and then reflect on its suitability for supporting evidence retrieval in clinical practice.

Figure 5. Search time metrics



Results

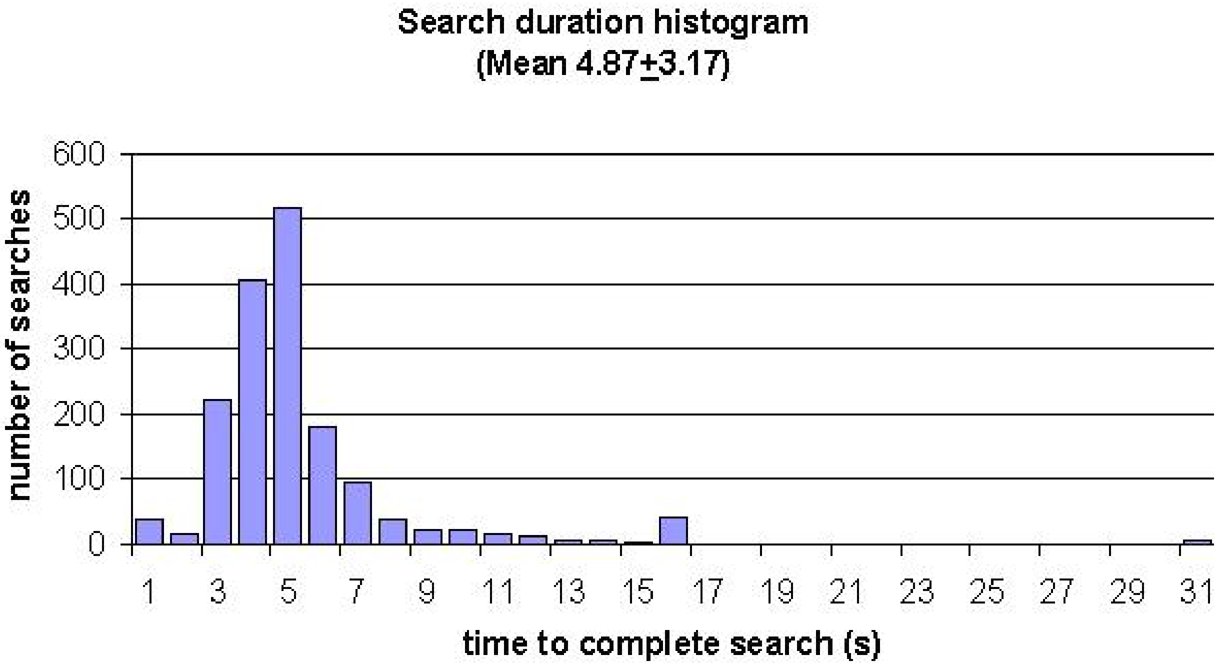
In the pre-trial questionnaire, 40% of the clinicians reported having a broadband (ADSL, cable, satellite) connection, while 43% used a 56k or 64k modem connection. The remaining 17% either did not know the type of connection used or had a slower connection. A total of 1662 searches were performed over the trial.

Search Speed

Under local network conditions (LAN, 100MBit), the user time (from starting the search on a client computer to displaying the results) was approximately 1.5 s. However, since most users accessed the system through the Internet, latency was significantly longer and slowed down the overall search speed.

The average search time was 4.9 s, with a standard deviation of 3.2 s (N = 1662 searches). Figure 6 shows the distribution of all search times over the trial. There are four distinctive features in this chart. The first is a small peak at 1 s (ie, searches that took up to 1 s to complete). The second feature is a peak around the mean value. Third, there is a small peak at 15 s, and, fourth, there is a small peak at 30 s.

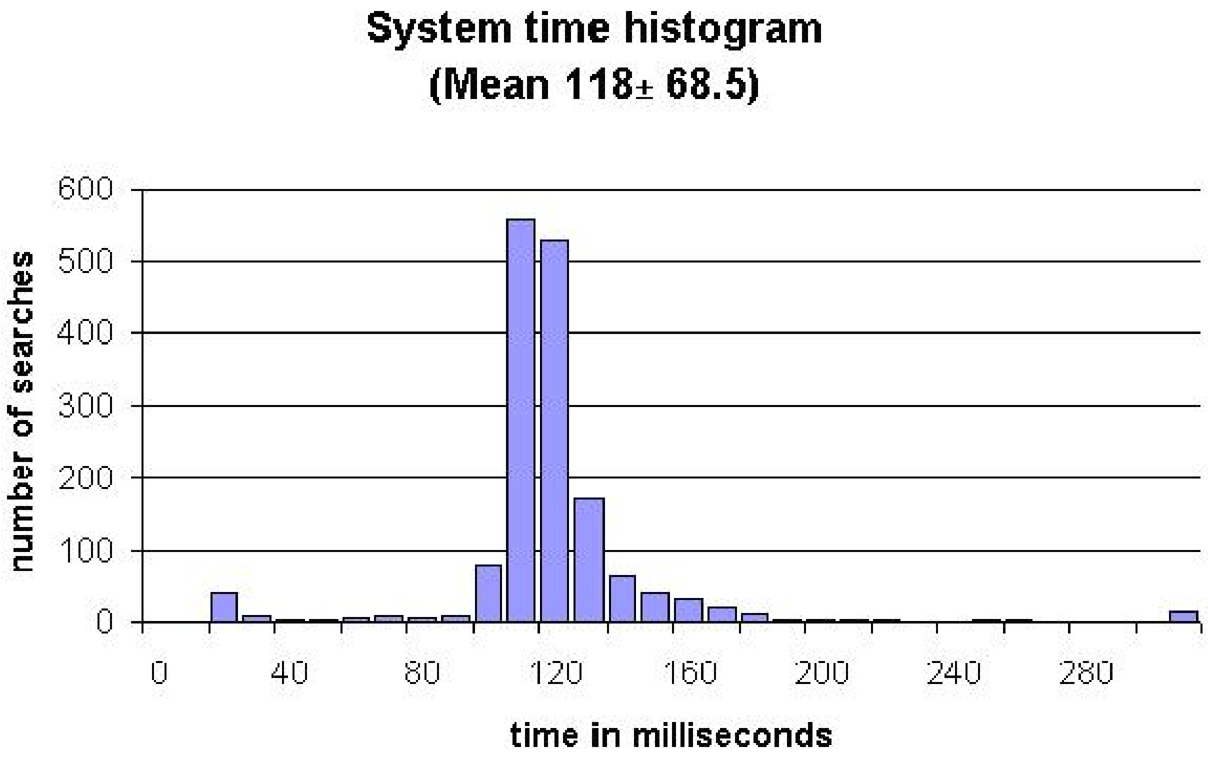
Figure 6. Distribution of search time for all 1662 searches



System Time

System time for a search was computed by subtracting the duration of the slowest source in every search from the search time (see Figure 5). From the system time histogram in Figure 7, it can be seen that for the majority of the searches the system takes between 100 ms and 130 ms (mean = 117.9 ms; SD = 68.4 ms; N = 1614 [48 searches had missing data, hence 1614 searches]).

Figure 7. Distribution of the system time for 1614 searches



System Time Versus Number of Individual Sources Involved

Depending on the search profile selected, the system will query a certain number of information sources and combine the results. To illustrate the dependency between system time and the number of sources queried, Table 2 shows average system time versus the number of sources queried in a search. The number of sources queried is predefined by the search profile, and none of the search profiles tested queried five, six, or eight sources.

Table 2. System time vs number of sources queried

Number of Sources Queried	N	Average System Time (ms)
1	48	18.1
2	9	31.8
3	15	73.3
4	7	59.7
5	0	-
6	0	-
7	1373	122.2
8	0	-
9	162	122.6

Speed and Reliability of Individual Data Sources

In addition to the performance measurements of the whole searches, the speed and reliability of the individual data sources was measured. Reliability was measured as the number of error cases (ie, queries that were not answered due to an error condition, such as a network error, an HTTP error, or queries that timed out). Reliability and speed figures are summarized in Table 3.

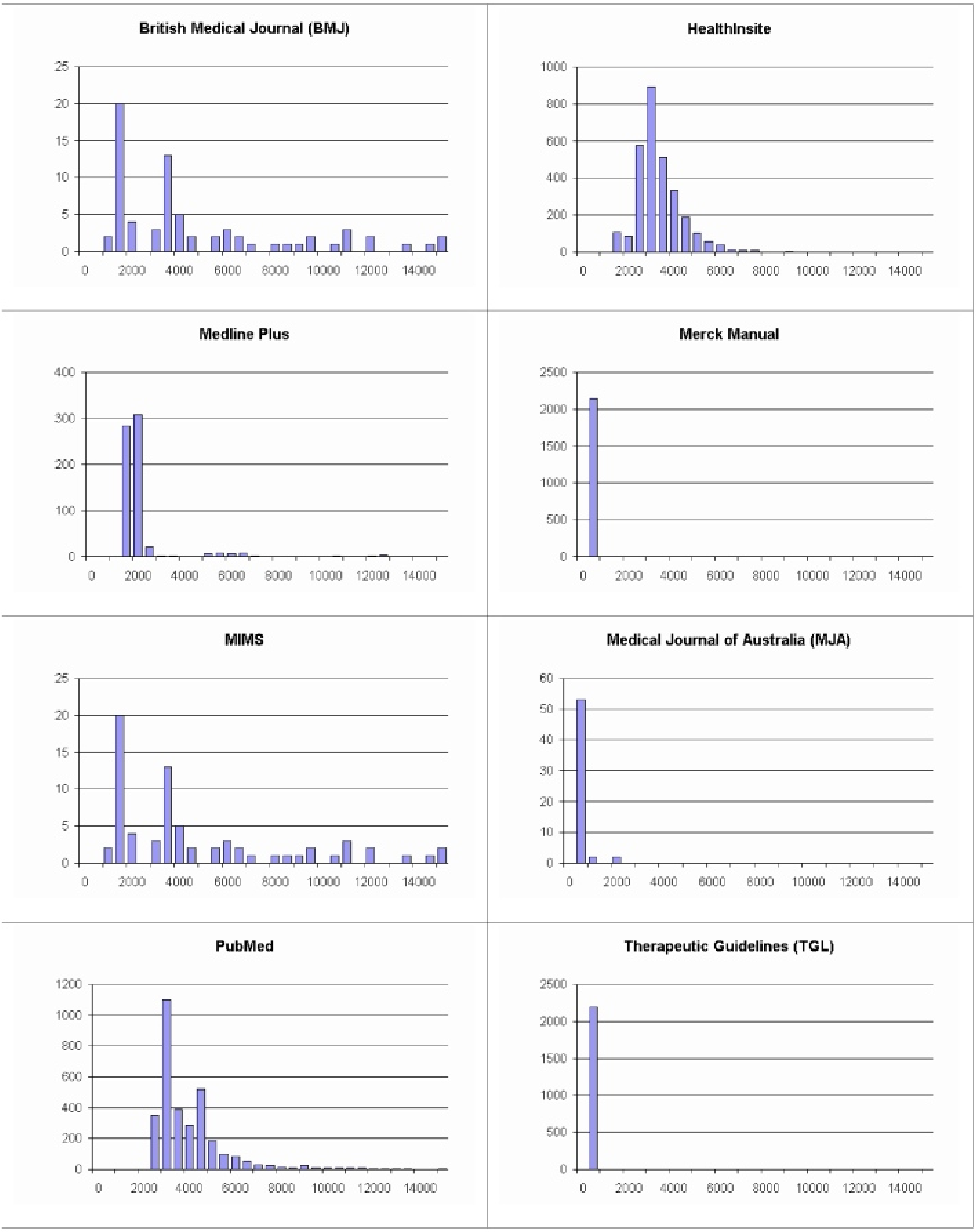
Table 3. Reliability and speed of data sources

Source	Type	Number of Searches	Number of Errors	Error (%)	Mean Speed (s)	SD (s)	Min (s)	Max (s)
Merck	Local	2144	0	0.0	0.06	0.11	0.01	2.89
TGL	Local	2193	0	0.0	0.05	0.12	0.01	2.85
BMJ	Remote	73	1	1.4	4.55	3.92	0.99	17.5
HealthInsite	Remote	2993	55	1.8	3.09	1.08	1.08	22.3
MedlinePlus	Remote	653	0	0.0	1.87	1.36	1.09	12.5
MIMS	Remote	650	3	0.5	0.98	1.14	0.28	8.30
MJA	Remote	58	1	1.7	0.25	0.31	0.10	1.73
PubMed	Remote	3288	39	1.2	3.76	1.69	1.87	15.0
Total		12052	99					
Mean				0.8	1.83	0.63*	0.68	10.4

* standard error of the mean

The most reliable sources were the locally indexed sources Merck (*The Merck Manual*) and TGL (Therapeutic Guidelines Australia), both which did not have any error cases. On the other end of the scale are HealthInsite (a national consumer site for health information) and *MJA*. The slowest source in the trial was *BMJ*, with an average of 4.55 s to process a query (SD = 3.92 s; N = 73). This was followed by PubMed, which returned results at an average of 3.76 s (SD = 1.69 s; N = 3288). The two locally indexed sources (Merck and TGL) returned search results within an average of 0.061 s and 0.047 s, respectively. However, the two local sources do have a relatively large standard deviation. Figure 8 shows the distribution of query times to the eight individual data sources.

Figure 8. Histogram of search times for each of the eight data sources (x-axis is time taken for a search, in ms; y-axis is number of searches)



Discussion

System Time

From the results of the system time versus source time, we can observe that system-processing time is only a fraction of the total search time. However, there are exceptions, namely when local data sources are used exclusively. From a user's perspective this still would not be an issue as the overall user time is greater by at least a factor of ten. It could, however, become a problem in a situation where many searches are dependant on the result of a previous search and have to be executed in series. System time has thus been kept relatively short, removing the initial reservation that too much parallelism could slow down the system excessively. From Table 3 it can be seen that the system time generally increases in line with the number of sources queried (with the exception of four sources queried). However, the order of this increase does not appear to be squared or even exponential, but rather linear.

Search Times

The four distinct features in the histogram of search times described in Figure 6 are due to the nature of the data sources and the value of the time-outs. The first small peak at 1 s is from search profiles that use exclusively local data sources. The second feature is a peak around the mean value and is caused by the six Internet resources. The small peak at 15 s is due to the large number of search profiles that have this value as a time-out. And finally, the tiny peak at 30 s is where the remaining searches time out.

It was to be expected that local sources would be more reliable and have a shorter latency in response time. This is due to the controlled environment, compared to the uncontrolled Internet environment of the external sources. It is interesting to note the difference between the six external data sources. While some sources are very popular (eg, PubMed) and therefore are expected to be busy, others might lack the resources to keep up with demand. The time-out value of individual data sources is a trade-off between speed and quality of results and is determined by the intended usage of the system. However, under certain circumstances there are optimizations that can be carried out without affecting quality of results. For example, the search duration histogram for HealthInsite (Figure 8; top right) reveals that if a search has not completed within 10 s it is highly unlikely it will complete within 15 s. Therefore, a time-out value of 15 s can safely be reduced to 10 s without significantly compromising search quality.

Future Work

The current QC architecture has demonstrated in trials that it meets the technical design goals set for it, and it provides good evidence that our general approach to federated searching is sustainable and maintainable. We intend to pursue research and development in areas of current interest to meta-search engines, information retrieval systems, and artificial intelligence. These include automatic wrapper generation [17,18] so that new data sources can be easily integrated into QC. Using this approach, a component could automatically generate a wrapper from knowledge of the data source query inputs and results. Another area of continued research will be automated data source consistency checking. Data sources often change in their formats, and this needs to be monitored with either automated or human intervention in order to modify wrappers accordingly. A third area will be intelligent search agents [19]. We envisage incorporating an intelligent agent that will guide users through the search process, using domain knowledge to help frame clinical questions and choose search parameters. This agent could learn to work with its user. An area of continued development will be semantic understanding of result sets. We would like QC to combine search results into a meaningful coherent story that presents a concise, relevant, and digestible response to the user [20]. These approaches, coupled with user support, will allow us to develop and improve the system with a view to it becoming an integral part of a clinician's daily practice. Even without these enhancements, we have demonstrated that the QC framework is a functional and useful approach for the delivery of online, just-in-time clinical evidence.

Acknowledgments

This work was supported by a National Health and Medical Research Council Development Grant and an Australian Research Council SPIRT grant, in partnership with Merck Sharp and Dohme (Australasia) Pty Ltd.

Conflicts of Interest

Quick Clinical was developed by researchers at the Centre for Health Informatics at the University of New South Wales, and the university and some of the authors could benefit from commercial exploitation of QC or its technologies.

References

1. Coiera E. Information economics and the Internet. J Am Med Inform Assoc 2000;7(3):215-221. [PMC: 10833157] [Medline: 20290754]
2. Lawrence S, Giles CL. Searching the world wide Web . Science 1998 Apr 3;280(5360):98-100. [Medline: 0009525866] [doi: 10.1126/science.280.5360.98]
3. Hersh WR. Information retrieval: a health and biomedical perspective. 2nd edition. New York: Springer-Verlag; 2003.
4. Magrabi F, Coiera EW, Westbrook JI, Gosling AS, Vickland V. General practitioners' use of online evidence during consultations. Int J Med Inform 2005 Jan;74(1):1-12. [Medline: 15626631] [doi: 10.1016/j.ijmedinf.2004.10.003]
5. Open Archives Initiative. URL: http://www.openarchives.org/ [accessed 2005 Sep 19]
6. Arens Y, Chee C, Hsu C, Knoblock C. Retrieving and integrating data from multiple information sources. Int J Intell Coop Inform 1993;2(2):127-158. [doi: 10.1142/S0218215793000071]
7. Bergamaschi S, Deneventano D. Integration of information from multiple sources of textual data. In: Intelligent Information Agents. New York: Springer; 1999:53-77.

8. Schmitt B, Schmidt A. METALICA: an enhanced meta search engine for literature catalogs. Presented at: Second Asian Conference on Digital Libraries; 1999; Taipei, Taiwan.
9. Coiera E. Guide to Health Informatics. London, UK: Arnold; 2003.
10. Haynes RB, Wilczynski N, Mckibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994;1(6):447-458. [PMC: 7850570] [Medline: 95153434]
11. Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. J Am Med Inform Assoc 2001;8(4):391-397. [PMC: 11418546] [Medline: 21311701]
12. Huang L, Hemmje M, Neuhold EJ. ADMIRE: an adaptive data model for meta search engines. Computer Networks (The International Journal of Computer and Telecommunications Networking) 2000;33(1):431-448.
13. Levy AY, Rajamaran A, Ordille JJ. Querying heterogeneous information sources using source descriptions. Presented at: Twenty-Second International Conference on Very Large Databases; September, 1996; Mumbai, India p. 251-262.
14. Divita G, Browne A, Rindflesch TC. Evaluating lexical variant generation to improve information retrieval. Presented at: AMIA Annual Symposium; 1998; Philadelphia, Penn p. 775-779.
15. Westbrook JI, Coiera EW, Gosling AS. Do online information retrieval systems help experienced clinicians answer clinical questions? J Am Med Inform Assoc 2005;12(3):315-321. [Medline: 15684126] [doi: 10.1197/jamia.M1717]
16. Westbrook JI, Gosling AS, Coiera EW. The impact of an online evidence system on confidence in decision making in a controlled setting. Med Decis Making 2005;25(2):178-185. [Medline: 15800302] [doi: 10.1177/0272989X05275155]
17. Chidlovskii B, Ragetli J, de Rijke M. Automatic wrapper generation for web search engines. Presented at: Web-Age Information Management, First International Conference; June 21-23, 2000; Shanghai, China p. 399-410.
18. Kushmerick N. Wrapper induction: efficiency and expressiveness. Artif Intell 2000;118(1-2):15-68. [doi: 10.1016/S0004-3702(99)00100-9]
19. Zick L. The work of information mediators: a comparison of librarians and intelligent software agents. First Monday 2000 May;5(5).
20. Levy AY. Combining artificial intelligence and databases for data integration. In: Artificial Intelligence Today: Recent Trends and Developments. Lecture Notes in Comput Sci 1999:249-268.

Abbreviations

CM: capability manager
QC: Quick Clinical
UQL: unified query language
UReL: unified response language

---

submitted 15.12.04; peer-reviewed by W Hersh, S Isaacksz; comments to author 01.03.05; revised version received 05.10.05; accepted 13.09.05; published 24.10.05.

Please cite as:
Coiera E, Walther M, Nguyen K, Lovell NH
Architecture for Knowledge-Based and Federated Search of Online Clinical Evidence
J Med Internet Res 2005;7(5):e52
URL: http://www.jmir.org/2005/5/e52/
doi:10.2196/jmir.7.5.e52
PMID:16403716

Original Paper

What Is eHealth (5): A Research Agenda for eHealth Through Stakeholder Consultation and Policy Context Review

Ray Jones[1], PhD; Ray Rogers[2], BSc; Jean Roberts[2], BSc; Lynne Callaghan[1], PhD; Laura Lindsey[1], BSc; John Campbell[3], MD; Margaret Thorogood[4], PhD; Graham Wright[2], MPhil; Nick Gaunt[5], PhD; Chris Hanks[1], MSc; Graham R Williamson[1], PhD

[1]School of Nursing and Community Studies, University of Plymouth, Plymouth, UK
[2]Centre for Health Informatics Research and Development, University College, Winchester, UK
[3]General Practice and Primary Care, Peninsula Medical School, Exeter and Plymouth, UK
[4]Epidemiology, Warwick Medical School, University of Warwick, Coventry and Leicester, UK
[5]NHS Institute for Innovation and Improvement, Coventry, UK

Corresponding Author:
Ray Jones, PhD
School of Nursing and Community Studies
University of Plymouth
Faculty of Health and Social Work
Plymouth PL3 4SP
UK
Phone: +44 01752 233886
Email: ray.jones@plymouth.ac.uk

Abstract

Background: In 2003, the National Health Service in England and Wales, despite its large investment in information and communication technology, had not set a national research agenda. The National Health Service has three main research and development programs: one is the Service Delivery and Organisation program, commissioned in 2003, and the others are two parallel “scoping exercises” to help set a research agenda. This paper reports on one of those projects. A parallel literature review was carried out by others and has been reported elsewhere.

Objective: The objective was to explore the concerns of stakeholders and to review relevant policy in order to produce recommendations and a conceptual map of eHealth research.

Methods: There were two parallel strands. For the stakeholder consultation, 37 professionals representing 12 “stakeholder” groups participated in focus groups or interviews. Discussion was prompted by eHealth “scenarios” and analyzed using thematic content analysis. Subsequently, 17 lay participants, in three focus groups, discussed and prioritized these themes. For the policy review, 26 policy makers were interviewed, and 95 policy documents were reviewed. Recommendations were subsequently reviewed in a conference workshop. Recommendations for research from both strands were combined into a conceptual map.

Results: Themes from stakeholder consultation and policy review were combined as 43 recommendations under six headings. Four of these headings (using, processing, sharing, and controlling information) describe the scope of eHealth research. The other two relate to how research should be carried out (ensuring best practice is first identified and disseminated) and to the values considered important by stakeholders (in particular, measuring improvement in health).

Conclusions: The scope of eHealth research (using, processing, sharing, controlling information) derived empirically from this study corresponds with “textbook” descriptions of informatics. Stakeholders would like eHealth research to include outcomes such as improved health or quality of life, but such research may be long term while changes in information technology are rapid. Longer-term research questions need to be concerned with human behavior and our use of information, rather than particular technologies. In some cases, “modelling” longer-term costs and benefits (in terms of health) may be desirable.

(J Med Internet Res 2005;7(5):e54) doi:10.2196/jmir.7.5.e54

KEYWORDS

medical informatics; research priorities

Introduction

In 2002, the National Health Service (NHS) in England and Wales planned to invest over £2 billion in information and communication technology (ICT) [1]. This includes initiatives such as electronic patient records, electronic prescribing, the NHS Direct Telephone and Internet Service, and the National Electronic Library for Health [1,2]. Researchers from multiple disciplines in the UK and elsewhere had been investigating health informatics, but the NHS, despite its large investment in ICT, had not set a national research agenda for ICT.

A one-day conference in 2002 on Health Informatics Research and Development, sponsored by the research councils, Department of Health, and Department of Trade and Industry, concluded that the "lack of national strategy, capacity and career paths in health informatics have been weaknesses and remain threats to realising the informatics potential of the National Health Service…. [L]arge investment in the National Health Service and e-Science is unlikely to achieve its objectives without radical improvement in support for academic health informatics…. [This emphasizes the] importance of...clarifying the academic agenda for health informatics" [3]. The short- and medium-term challenges were seen as the following: (1) establishing the foundations of a knowledge infrastructure, (2) innovations in the clinician computer interface, (3) workable privacy protection, (4) more creation of knowledge from routinely collected data, and (5) finding the metrics of success for health informatics.

The NHS has three main national NHS research and development programs: Health Technology Assessment, New and Emerging Applications of Technology, and Service Delivery and Organisation (SDO). The SDO was launched on March 30, 2000, to consolidate and develop the evidence base on the organization, management, and delivery of health care services [4]. To respond to the needs of the "stakeholders" (service users, health professionals, and policy makers), the NHS, through its SDO research program, undertook an initial "listening exercise"[5] to produce a document outlining its overall priorities for research. It has continued to use this approach to develop and commission research [6]. It commissions a "scoping exercise" (normally a literature review and a stakeholder consultation) and then uses that in subsequent calls for proposals.

This study explored the concerns of professional and lay stakeholders regarding future developments of eHealth and reviewed relevant policy to produce recommendations for eHealth research. A parallel literature review was carried out by others and has been reported elsewhere [7-9].

Methods

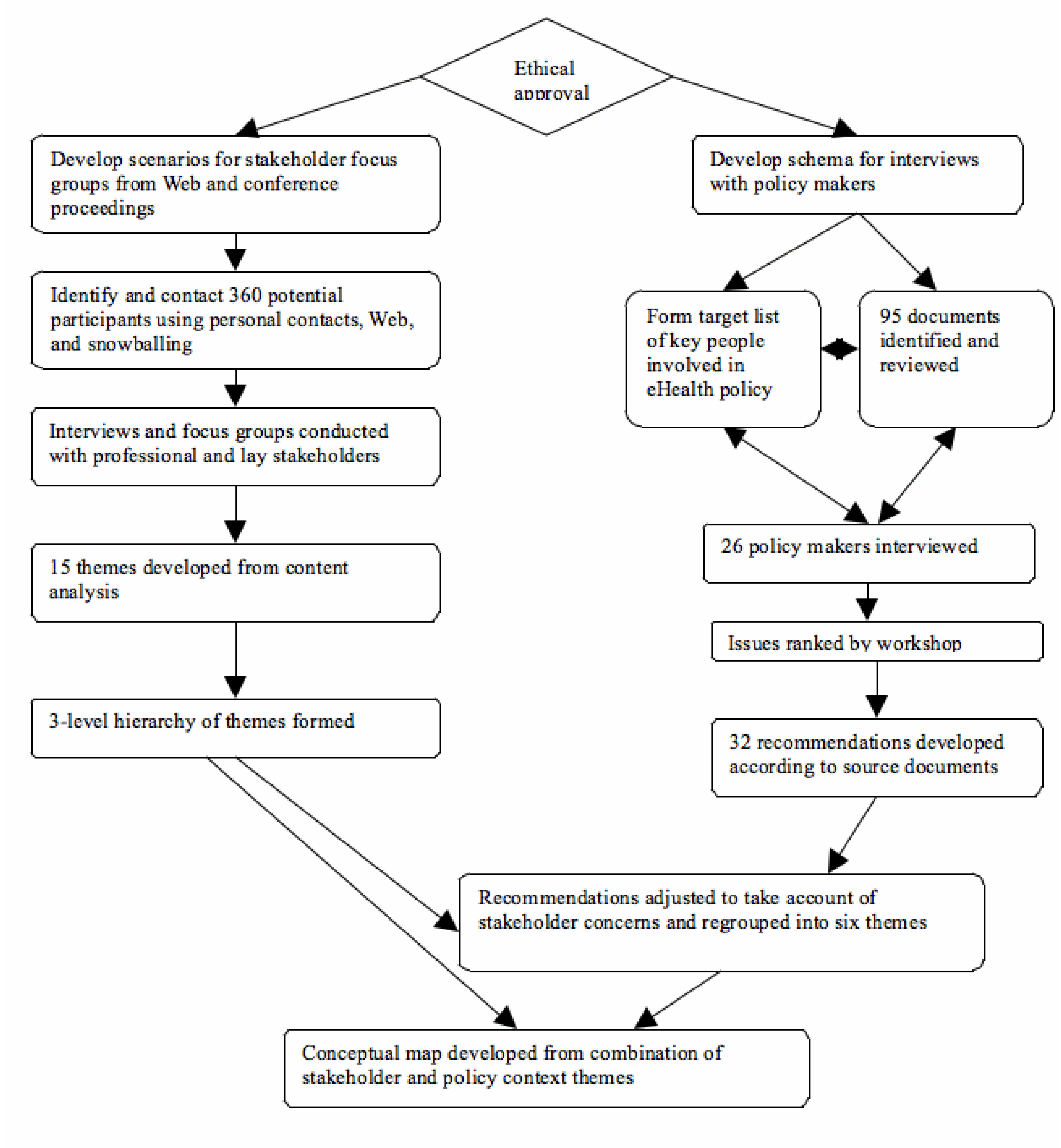
Study Design

The study was reviewed and approved by the SouthWest Multi-Centre Research Ethics Committee and the University of Plymouth Faculty of Health and Social Work Ethics Committee. Data collection was carried out between November 2003 and June 2004. There were two parallel strands (Figure 1):

1. Stakeholder consultation: Focus groups and interviews with "professional" stakeholders generated themes that were subsequently prioritized by lay participants.
2. Policy context review: Policy makers were interviewed and policy documents were reviewed in order to produce recommendations that were subsequently reviewed in a conference workshop.

The themes and research questions arising from the stakeholder consultation and policy context review were compared, and recommendations from policy context were adapted to take account of stakeholder concerns. Diagrams were developed to bring together stakeholder and policy maker views of the scope of eHealth research.

Figure 1. Parallel methods of stakeholder consultation and policy review



Stakeholder Consultation

In all, 12 groups (30 in each group) of professional stakeholders were contacted via email. Potential participants were identified by Web searches and by "snowballing" from existing contacts, trying to get geographical coverage within England. Stakeholders were defined as the following:

- NHS eHealth innovators and implementers
- University researchers in health informatics
- NHS staff in primary care
- NHS staff in secondary care
- NHS primary care trust managerial staff
- NHS acute trust managerial staff
- Suppliers of ICT to the NHS
- Professional organizations and royal colleges
- Informatics trainers
- Governance and other regulators
- Charities and other information providers
- Other NHS managers

These 360 people were sent an email inviting them to take part in the study, with a consent form to return to the researcher via email or post if they agreed to take part in the study. We asked professional participants to rank themselves on a four-point scale (from "I am pretty sceptical that eHealth will have any benefits at all" through to "I am very positive that eHealth can help improve the NHS if used appropriately”) as a rough guide to attitudes toward eHealth (Table 1).

Table 1. Professional participants: locations and eHealth attitudes

Stakeholder Group (number of participants)	Location in England	Self-Rating on eHealth*
Innovators and implementers (5)	South West	3
	South East	4
	South East	3
	South East	4
	South West	3
Academic researchers in eHealth (4)	Midlands	3
	Midlands	3
	South West	3
	South East	3
NHS staff in primary care (3)	South West	3
	South West	4
	South West	4
NHS staff in secondary care (4)	South East	3
	South East	3
	South West	4
	South East	3
Primary care managerial staff (1)	South West	3
Acute care managerial staff (1)	South East	3
Suppliers (3)	South East	3
	South West	4
	North	3
Royal colleges (3)	South East	3
	South East	4
	South East	4
Informatics trainers (5)	South East	3
	Midlands	3
	Midlands	3
	Midlands	3
	South East	3
Governance (3)	South East	3
	South East	4
	South East	3
Charities and other providers (2)	Midlands	3
	South East	3
Other NHS managers (3)	South East	3
	South West	4
	North	3

* Self-ratings: (1) I am pretty sceptical that eHealth will have any benefits at all. (2) I think that there could be some possible benefits to eHealth methods but on balance think that it is unlikely that the benefits will outweigh the costs. (3) I think that there are definitely benefits to eHealth but that we need to choose and develop methods carefully. (4) I am very positive that eHealth can help improve the NHS if used appropriately.

A convenience sample of lay participants was recruited via snowballing from contacts in a local children's nursery and from older friends of the research team. Potential participants were sent study information sheets and consent forms either via email or post. Two groups of older people and one group of parents took part in the study.

Scenarios depicting the current or future use of eHealth technologies were constructed to prompt discussion among the professional groups of the relevant themes regarding the use of eHealth technologies. Subject matter of the scenarios was developed from the content of news reports, informatics conference proceedings, and general Web searches. Both patient- and professional-centered scenarios were developed in order to achieve a balance of perspectives. The research team developed and discussed 32 scenarios: 15 were omitted, 7 were added, and a number were reworded to ensure neutrality in presentation.

In total, 24 scenarios (see Multimedia Appendix 1) were allocated to the 12 professional groups using a balanced incomplete block design [10]. Each group had four scenarios, and each scenario was used twice. A semistructured schedule based on the scenarios was constructed for use either as focus group or interview prompts. Some scenarios described patients being monitored by an implanted device that sent physiological information to hospital, a family doctor booking a hospital appointment during consultation in primary care, and a woman having an antenatal ultrasound in the community with expert diagnosis from abroad. Textbox 1 provides four examples of scenarios used as prompts in telephone focus groups. The topics covered in the scenarios included patients' use of the Internet to order prescriptions, arrange doctor's appointments, or join patient discussion forums. Other topics were about patients accessing their own health record, assessing the quality of a website, using a digital interactive television for a program on multiple sclerosis, or using a public access touch screen health information point.

Textbox 1. Four examples of scenarios used as prompts in telephone focus groups

1. Ordering prescriptions: Sam, 45, drops off a repeat prescription for his high cholesterol medication every month and has a check-up routinely every three months. Recently, the local pharmacy and Sam's family doctor have started a scheme whereby patients can order their repeat prescriptions online, thereby relieving the burden on administrative staff at the surgery (primary care health center). Following an order being made by a patient, the pharmacy provides the doctor with a list of repeat prescriptions, which the doctor approves or not. The pharmacy then sends an email to Sam when his medication is ready for collection.
2. Use of implanted device: James, a diabetic, has an implanted device that measures his blood glucose level and transmits this reading to the hospital. If the reading is below a certain level, James is contacted on his mobile phone by an automated system. Recently, the hospital received a signal that James's blood glucose had dropped to 1.5. The doctor was alerted and visited the patient to review his medication.
3. e-booking: Peter, aged 45, attends an appointment with his doctor about his recent weight loss. His doctor decides that Peter should be referred for an appointment at the hospital and uses the new e-booking system. Upon inputting Peter's details into the system, an appointment was set up immediately, and Peter was able to leave the surgery with his hospital appointment arranged.
4. Wireless technology: Ann Young is a district nurse who uses a palmtop with wireless access to the internet and PCT intranet. Ann regularly uses her palmtop in order to ask advice of her colleagues or to obtain test results, and she now views her palmtop as an invaluable resource. At the next practice meeting, Ann intends to present the benefits of using a palmtop to her colleagues.

Policy Context

A parallel, two-stage process was used to review the policy context. First, policy makers were interviewed and policy documents were identified and reviewed. The contents of documents and interview notes were categorized under English policies on ICT specific to health, English health policies influencing eHealth, nonhealth policies influencing eHealth, and European Union policies influencing eHealth. The reviewers focused on seven specific topics: (1) birth to death records, (2) country-wide access to quality health advice, (3) application of ICT to pharmacy, (4) telemedicine, (5) reduction of adverse incidents, (6) confidentiality, and (7) health data cards. In a second stage, 28 recommendations from the policy makers and documents were reviewed in a health informatics conference workshop by 60 participants using discussion and an interactive voting system. Participants scored each recommendation for relevance to the needs of the NHS (using a nine-point scale from "not relevant" to "highly relevant"). Recommendations with "middle scores" were discussed with the audience in more detail to obtain their views and decide if the recommendation needed to be worded more clearly.

Synthesis and Conceptual Mapping

Two members of the research team (RJ and LC) independently ranked the correspondence between the stakeholder concerns and the policy context recommendations on a scale of 1 (no correspondence) to 3 (strong correspondence). Analysis showed there to be strong agreement between the researchers. Stakeholder concerns, particularly of "technology meeting needs and improving health and quality of life," were not consistently addressed by the policy context recommendations. Recommendations were adapted to take stakeholder concerns into account and were regrouped from "source-oriented" to "research-oriented" groupings. The new policy context recommendations were agreed upon between team members by telephone conference discussion. The two lists, one from stakeholder consultation and the other from the policy context review, were then reviewed again and similar areas of research were grouped. Reference was also made to the work of the Scottish Consumer Health Informatics Network [11]. We concluded that the scope of eHealth research could be described by a simple block diagram with four elements with linked areas of "best practice." The recommendations were regrouped according to this "conceptual map."

Results

Subjects and Sources

In total, 37 (10%) professional stakeholders consented to take part in the study and were consulted either via telephone focus group (25), telephone interview (6), videoconference (4), or in-person interview (2). There were 17 lay people (12 older people and 5 parents of young children) that took part in in-person focus groups. We interviewed 26 policy makers and identified and reviewed 95 policy documents.

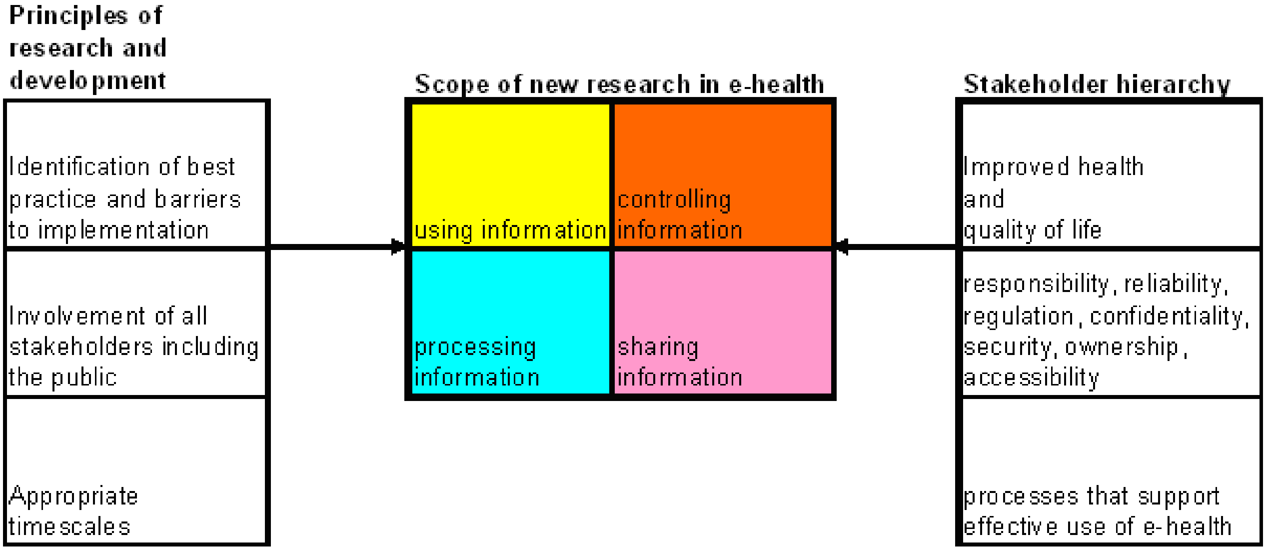
Validity of Methods

One of the limitations of stakeholder consultation is that participants have to be sufficiently interested in the topic to take part. None of the 37 professional participants were "sceptical that eHealth will have any benefits at all," and none thought "there could be some possible benefits to eHealth methods but on balance think that it is unlikely that the benefits will outweigh the costs." On the other hand, 27 thought "that there are definitely benefits to eHealth but that we need to choose and develop methods carefully," and 10 were "very positive that eHealth can help improve the NHS if used appropriately."

Scope of eHealth

The research questions identified by stakeholders and policy review fell into six groups. Four of these (using, controlling, processing, sharing information) were used to describe the "scope" of eHealth. The other two groups of research questions fall under principles of research and development and stakeholder hierarchy (Figure 2).

Figure 2. Scope of eHealth research



Principles of Research and Development

Both stakeholders and policy makers referred to many examples where, before innovative approaches are introduced, best practice procedures and barriers to implementation should be identified, and where professional and public stakeholders should be involved in research and development. A number of areas were suggested (Textbox 2).

Textbox 2. Seven research aims related to the identification and implementation of best practice in eHealth

1. Informatics training for health professionals, identification and exploration of examples of best practice to see how these can be disseminated to achieve improved health care, exploration of the attitudes of health professionals toward such training and use of the skills acquired in practice
2. Working practices in other sectors (eg, e-business) to identify best practice and barriers to similar uses of ICT in the health sector
3. Web-based services for citizens in other sectors to see what lessons can be learnt on when to implement eHealth solutions for patients
4. Research on telemedicine (eg, using coronary heart disease or cancer services) and barriers to its implementation
5. The costs and benefits (including improved patient safety) of hospital systems that combine e-prescribing, order entry, decision support, bar coding for medication management, and robotic dispensing
6. NHS procedures that aim to safeguard confidentiality of patient data and disseminate best practice
7. Assessment of the experience of UK citizens accessing health care in other countries (and vice versa) and identification of where health and other outcomes could be improved through the use of ICT

Scope of New Research in eHealth

Using Information

Information is used in decision support, in the organization of services, for reassurance of professionals and patients, and in information-based therapies. Four research aims (Textbox 3) from the policy context review concerned the use of information in decision support.

Textbox 3. Four research aims related to the way information is used

1. To assess clinicians' and patients' perceptions of the benefits and barriers to using decision support tools—in particular, to compare clinicians who use decision support tools with clinicians who do not
2. To assess the quality of information available from repositories of health data and to assess how it can be legally, ethically, and cost-effectively aggregated for public health policy and decision support
3. To explore the costs and potential benefits of birth to death records in relation to decision making and other aspects of health care and to identify policy changes required to achieve them
4. To review decision support and expert systems used in the NHS to ascertain their impact on patient services

Sharing Information

Both stakeholders and the policy context review identified a large number of research questions related to sharing information (Textbox 4). These included how information should be shared across sites (eg, between hospital and home), across sectors (eg, between social services and NHS), and between different professional (and patient) groups (eg, between doctors, nurses, dentists, patients).

Textbox 4. Thirteen research aims related to the way information is shared

1. To examine how the NHS can work with other information and education providers to facilitate patient involvement in eHealth
2. To explore patient attitudes toward initiatives of patient involvement in eHealth
3. To identify the extent to which implanted or wearable technology removes patients' control of their condition and to identify how ICT may best be used to encourage and facilitate patients to take responsibility for their health
4. To investigate the extent to which recently introduced information technology–based systems (such as e-booking) increase patient expectations and consequently decrease satisfaction if those expectations are not met
5. To investigate the efficacy of developing a code of collaboration under which organizations can explicitly share data and input to health records consistently, unambiguously, and sensitively
6. To determine how we can best deal with combining multiple sources of data, dealing with apparently conflicting information from different sources, with minimum patient risk, minimum cost, and patient consent and confidence
7. To examine the costs and benefits of cross-sectoral records and patient safety issues associated with cross-sectoral working
8. To investigate how ICT can best contribute to pharmacy clinic services sharing data between the NHS and patient
9. To investigate the potential of eHealth to enable effective interfaces, for example, between health and social care, local specialists and specialist services, care givers and professionals
10. To investigate the costs and benefits of using different technologies to support community-based staff (eg, notepad computers, electronic links to supporting organizations, teleconferencing in cancer services)
11. To explore the changes in work patterns, the potential for patient involvement, and legal issues in home care (eg, for older people)
12. To investigate ICT use in multisite work in relation to such issues as culture change, governance, health professional training, patient expectations, and changes to health outcomes
13. To determine the costs and benefits of the use of health data cards

Controlling Information

This group of concerns was ranked second most important by stakeholders. Ten research aims (Textbox 5) incorporated issues of control, accessibility, reliability, confidentiality, security, ownership, and regulation.

Textbox 5. Ten research aims related to the way information is controlled

1. To investigate how health professionals and patients discriminate between reliable and unreliable information
2. To examine the circumstances in which regulation of information provision and use is necessary and, further, when education and empowerment of professionals is a more effective option—additionally, what are health professional and patient attitudes toward the regulation of health information?
3. To investigate the extent to which health professionals advise patients of reliable sources of information on the Web, television, and other media, and further, to examine the level of preparation and support that health professionals require to provide such advice and to examine patients' expectations of this advice
4. To determine the site of responsibility if health care errors are made as a result of information transfer
5. To explore how social organization and different technologies can be used to help prevent inequity of access to information for both patients and professionals, and to identify initiatives whereby groups traditionally considered to have restricted access have successfully achieved training and access to new technologies
6. To explore health professional and patient attitudes toward ownership and sharing of data
7. To develop and test guidance on regulation and responsibility
8. To examine the costs and benefits of different ways of addressing equity to inform citizens
9. To investigate ways (quality marks, portals, patient and health professional training) to assist the public in obtaining quality information from the Web
10. To investigate patients' knowledge and views on confidentiality and their attitudes as to how their data should be used (eg, in research) in terms of potential benefits to health and quality of life

Processing Information

This covers a range of issues, including how best to present information (eg, should it be tailored for different users) and where it might be presented (eg, should it be sent to the user [push], or should it wait until the user seeks it [pull]). It also includes the coordinated integration of information derived from a variety of sources, as demonstrated in the electronic ordering and home delivery of medicines (Textbox 6).

Textbox 6. Seven research aims related to the way information is processed

1. To identify what extent health information should be tailored to the needs of certain groups of patients, professionals, or individuals
2. To examine the costs and benefits of providing information in different locations (eg, mobile versus static for professionals, NHS versus home for patients)
3. To investigate how information can be better integrated so that patients can, for example, access their own medical record on the Web, obtain relevant and validated information about it, and order a prescription
4. To identify instances or circumstances when patients want to enquire about health information through known professionals (eg, family doctor) and when they want to use an anonymous source
5. To investigate what services patients desire for electronic ordering and home delivery of medicine and how they can be delivered safely, equitably, and cost-effectively
6. To identify how eHealth technologies can enable or improve family support for seriously ill children and provide just-in-time information tailored to individuals
7. To examine the costs, benefits, attitudes toward, and the use of, ICT support in their homes for patients with severe chronic disease (eg, video links to NHS and voluntary services, smart cards with patient records)

Hierarchy of Stakeholder Concerns

The overriding concern of stakeholders was that spending money on eHealth should be worthwhile and should lead to improved health and quality of life. Particular research aims suggested by the data included "to review the costs and benefits of a range of recent eHealth applications, including the modelling of new forms of care made possible by ICT support," and "to present those examples of eHealth applications, shown to have a demonstrable effect on improved health and quality of life, to professional and public stakeholders to obtain their views as to the nature of the most appropriate investment in eHealth." In addition, the stakeholders identified themes concerned with controlling information (responsibility; reliability; regulation; accessibility; confidentiality, security, and ownership) as being particularly important, and so placed them on the second level of a hierarchy of concerns.

Discussion

In the context of the British health service, which is mostly free at the point of delivery, the overriding concern of stakeholders was that spending money on eHealth should be worthwhile and should lead to improved health and quality of life. At first this appears an unremarkable finding. However, although such an aim is part of the political rhetoric and may be an unstated assumption of policy documents in the United Kingdom, it is not often explicitly addressed in service development and use of information and communication technology. It is significant, for example, that the NHS recently funded a program originally called the National Programme for Information Technology, rather than a program for eHealth. (Subsequently, the program has been renamed Connecting for Health).

On the other hand, we know that doing research that can show a difference in health as a result of an eHealth intervention is difficult, partly because partial implementation of an e-booking system or a hospital information system is difficult. Gold standard randomized trials whose results can be generalized for widespread implementation are very difficult. In addition, to see changes in health or to measure cost benefit is slow and expensive, made more so now in the United Kingdom by the time needed for increased levels of ethical and research control and approval [12,13]. Thirty years ago, Blum noted that half the papers about computer applications concerned systems that were no longer operational [14]. We all know that ICT changes become ever more rapid. Research, therefore, has to be more about human behavior and how we use information and less about specific organizational or technological environments. It is essential to recognize the difficulties of addressing stakeholders' needs by measuring change in health outcomes.

In some cases, modelling the longer-term costs and benefits (in terms of health) may be desirable. As systems continue to evolve, the health benefits may be seen not in the immediate change, but in a future evolution made possible by the initial change [15-17].

Many of the recommendations derived from both the stakeholder consultation and policy review confirmed the need to identify best practice and the barriers to implementation of that best practice. There are many examples of medical informatics research with demonstrable benefit which decades later still wait to be implemented more widely. For example, computers have been used successfully for patient interviews for nearly 30 years. Slack et al first reported on a computer-based medical history system in the *New England Journal of Medicine* in 1966 [18]. Yet, despite numerous research reports [19], the method has not been routinely adopted.

Rogers' description of the diffusion of innovations [20] is well known, but stakeholders and policy makers want to see this process accelerated. The recommendations included, for example, identifying and exploring examples of effective informatics training for health professionals to see how these can be disseminated, or investigating working practices in other sectors (eg, e-business) to identify best practice and barriers to similar uses of ICT in the health sector.

In discussion, professional stakeholders often reverted to their role as patient or consumer rather than, for example, speaking as a supplier of ICT or from the point of view of primary care. Thus, although we sampled from all segments of the stakeholder population, there were no obvious differences between the different types of stakeholder. (Our sample was small, however, making our power to detect differences limited.)

The four categories which emerged as the grouped themes from the data are similar to classifications and descriptions found in textbooks of health informatics (eg, [21]). Blum, in a personal review of Medical Informatics in the United States, 1950-1975, presents a historical table of the Scope of Medical Computing from 1950 to 1980 using the three headings Data, Information, and Knowledge Applications [14]. He said, “Within a category, research begins only after the supporting technology is mature enough to support it beyond the conceptual level.” His table showed that he viewed data applications in the 1980s as refined, information applications as mature, and knowledge applications as prototype.

Four limitations of our study are the following: (1) We were not able to recruit people who were very sceptical about the potential of eHealth. However, we think that a more sceptical sample would be likely to have expressed similar concerns about improved health and value for money. (2) We did not have equal representation for the 12 predefined groups. However, we achieved coverage (although, in some cases, only one member; see Table 1) for all groups. Furthermore, as most participants often “reverted” to their role as patient in the discussion, our original idea that different professional groups might have certain biases and try to put forward ideas to their advantage seemed wrong (although numbers were small to detect any differences between groups). (3) Our lay sample was recruited only in the South East and the South West of England and did not include any people from ethnic minorities. We have no knowledge or hypotheses about how views may differ in groups not represented. (4) An eHealth agenda derived from policy review and stakeholder consultation in a country with a state-run health service may not transfer well to countries with private or insurance-based systems.

The purpose of our study was to produce a conceptual map and research agenda for eHealth based on stakeholder views and policy review. In a parallel SDO project, Pagliari et al [7-9] developed recommendations based on a review of existing academic and wider evidence sources, indicating the scope of the eHealth concept, the effectiveness of eHealth innovations, issues for implementation, and future directions for eHealth. Their research evidence is grouped by four broad technological categories:

1. Decision support tools for patients and clinicians
2. Networked digital technologies (Internet) used by patients (eg, for information, self-management, or peer support) and professionals (eg, for interprofessional communication, education, or communication with patients)
3. Computerized patient records, including issues relating to patient access and confidentiality and influences on health care delivery
4. Telemedicine and telecare

Their results are also interpreted with respect to the following broad content areas:

1. Specific research needs (evidence of effectiveness in specific areas)
2. Generic research needs (eg, methodological challenges to eHealth research, factors affecting implementation, effects on behavior and relationships, educational interventions, health inequalities, alternative delivery media, risks to the health service and society, role in self-care, and consumer empowerment)

3. Challenges for implementation (demonstrating impact, high-level support, strong project management, stakeholder engagement, the digital divide, ensuring credibility and quality, ethical, security and privacy issues, standards)
4. Emerging trends and future directions (eg. personalized and tailored systems, new technological advances for the information management and care facilitation, and delivery modes)

While differences in emphasis were expected due to the methodology of each study, the clear parallels between the results offer support for our recommendations. Our conceptual map, which has come from stakeholder discussion and policy review, also helps to put both our own and other detailed recommendations into a framework concerned with information and how we use it.

Acknowledgments

We thank all those who participated in this study. The study was funded by the NHS SDO Programme.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of 24 scenarios. [PDF file, 92 KB - jmir_v7i5e54_app1.pdf]

References

1. ; Department of Health. Delivering 21st Century IT Support for the NHS: National Strategic Programme. London, UK: Department of Health Information Policy Unit; 2002.
2. ; Department of Health. The NHS Plan. A Plan for Investment. A Plan for Reform. London, UK: The Stationery Office; 2000.
3. ; Medical Research Council. Proceedings of the Workshop on Health Informatics Research and Development; July 18, 2002. London, UK: Royal College of Obstetricians and Gynaecologists. URL: http://www.mrc.ac.uk/pdf-health_informatics_r_and_d_workshop.pdf [accessed 2005 Jul 23]
4. ; NHS Service Delivery and Organisation (SDO) Programme. Home page. URL: http://www.sdo.lshtm.ac.uk/index.html [accessed 2005 Jul 23]
5. Fulop N, Allen P. Service Delivery and Organisation National Listening Exercise: Report of the Findings. London, UK: National Health Service Executive; 2000.
6. Lomas J, Fulop N, Gagnon D, Allen P. On being a good listener: setting priorities for applied health services research. Milbank Q 2003;81(3):363-388. [Medline: 22823255]
7. Pagliari C, et al. EH1 E-Health Scoping Exercise. Review of the Traditional Research Literature. URL: http://www.sdo.lshtm.ac.uk/ehealth.htm [accessed 2005 Jul 23]
8. Pagliari C, et al. EH1 E-Health Scoping Exercise. Review of the Non-Traditional Research Literature. URL: http://www.sdo.lshtm.ac.uk/ehealth.htm [accessed 2005 Jul 23]
9. Pagliari C, Sloan D, Gregor P, Sullivan F, Detmer D, Kahan JP, et al. What is eHealth (4): a scoping exercise to map the field. J Med Internet Res 2005 Mar 31;7(1):e9 [FREE Full text] [Medline: 15829481] [doi: 10.2196/jmir.7.1.e9]
10. Armitage P, Berry G. Statistical Methods in Medical Research. Oxford: Blackwell Scientific Publications; 1987.
11. Marsden J, Jones R. Establishing and maintaining the Consumer Health Informatics Network for Scotland (CHINS). Health Care Computing 2004:169-177. URL: http://www.health-informatics.org/hc2004/P27_Marsden.pdf [accessed 2005 Sep 26]
12. Elwyn G, Seagrove A, Thorne K, Cheung WY. Ethics and research governance in a multicentre study: add 150 days to your study protocol. BMJ 2005 Apr 9;330(7495):847. [Medline: 15817562] [doi: 10.1136/bmj.330.7495.847]
13. Hearnshaw H. Comparison of requirements of research ethics committees in 11 European countries for a non-invasive interventional study. BMJ 2004 Jan 17;328(7432):140-141 [FREE Full text] [Medline: 14726341] [PMC: 14726341] [doi: 10.1136/bmj.328.7432.140]
14. Blum BI. Medical Informatics in the United States, 1950-1975. In: Blum BI, Duncan K, editors. A History of Medical Informatics. Proceedings of the ACM Conference on the History of Medical Informatics; Bethesda, Maryland; November 5-6 1987. New York: ACM Press; 1990:xvii-xxxx.
15. Friedman CP, Wyatt JC. Evaluation Methods in Medical Informatics. London, UK: Springer-Verlag; 1997.
16. Mumford E. A socio-technical approach to systems design. Requir Eng 2000;5(2):125-133.
17. Jones RB, Hedley AJ. Evaluation of a diabetes register and information system. In: Richards B, editor. Current Perspectives in Health Computing. Weybridge, UK: BJHC Books; 1986:80-87.
18. Slack WV, Hicks GP, Reed CE, Van Cura LJ. A computer-based medical-history system. N Engl J Med 1966 Jan 27;274(4):194-198. [Medline: 66074575]

19. Jones R, Knill-Jones RP. Electronic patient record project: direct patient input to the record. Report for the Strategy Division of the NHS Information Management Group. Glasgow: University of Glasgow; 1994.
20. Rogers E. Diffusion of Innovation. 4th ed. New York: The Free Press; 1995.
21. ; World Health Organization. Informatics and Telematics in Health. Present and Potential Uses. Geneva: WHO; 1988.

Abbreviations

ICT: information and communication technology
NHS: National Health Service
SDO: Service Delivery and Organisation

---

submitted 18.08.05; peer-reviewed by C Pagliari, S Olsson; comments to author 15.09.05; revised version received 16.09.05; accepted 21.09.05; published 10.11.05.

Please cite as:
Jones R, Rogers R, Roberts J, Callaghan L, Lindsey L, Campbell J, Thorogood M, Wright G, Gaunt N, Hanks C, Williamson GR
What Is eHealth (5): A Research Agenda for eHealth Through Stakeholder Consultation and Policy Context Review
J Med Internet Res 2005;7(5):e54
URL: http://www.jmir.org/2005/5/e54/
doi:10.2196/jmir.7.5.e54
PMID:16403718

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