Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 3 (2001), Issue 3 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Editorial

Original Papers

Editorial

Journal of Medical Internet Research is now indexed in Medline

G Eysenbach, MD

(J Med Internet Res 2001;3(3):e25) doi:10.2196/jmir.3.3.e25

We are very excited and proud to announce that in July 2001 the Journal of Medical Internet Research (JMIR) has been selected by the National Library of Medicine (NLM) to be indexed in Medline/Index Medicus, the world's most important biomedical bibliographical database. NLM uses an advisory committee, the Literature Selection Technical Review Committee, composed of authorities knowledgeable in the field of biomedicine, such as physicians, researchers, educators, editors, health science librarians, and historians, to review and recommend the journal titles NLM should index. Only 15-20% of medical journals are selected at each session for inclusion in Medline, and only very few pure electronic journals have been selected for indexing so far. The Committee recently completed a review of journals for possible inclusion in the National Library of Medicine's database systems and decided to select the "Journal of Medical Internet Research" for indexing and inclusion in Index Medicus and Medline. This will include not only articles which will be published in the future, but also retrospectively all articles which have been published in the Journal in the past 3 years, back to the first issue.

An electronic journal with no print equivalent must meet certain additional criteria to be selected for indexing. First, no restrictions on viewing the journal may exist (other than it might be fee-based). Secondly, NLM should be allowed to make copies of the content to preserve articles for posterity. Thirdly, publishers must provide XML/SGML-tagged data.

The Journal of Medical Internet Research is currently working on the required XML files to make these citations available to NLM. As soon as this process is finished, all articles published in the Journal of Medical Internet Research will be searchable via PubMed and other MEDLINE systems. We will then join the ranks of more than 150 publishers of nearly 1,000 MEDLINE journals (40% of MEDLINE's current citations) who are sending electronic XML data for citations and abstracts for quick inclusion in PREMEDLINE. PubMed will also offer links back to the respective article on the JMIR Web site. In addition, the Journal is working towards inclusion of its full-text articles into PubMed Central.

To date, only 20 pure electronic journals have been chosen for indexing by the National Library of Medicine, plus the BiomedCentral (BMC) specialty journals and the Cochrane database of systematic reviews. Table 1 lists all electronic journals chosen by NLM as of August 2001. Of the 20 electronic journals (excluding BMC and the Cochrane database), only 10, including JMIR, appear to be active. At least 3 electronic journals have been discontinued after a few years, and the remaining 7 journals have not provided any XML-tagged data or have for some other reasons no articles indexed to date.

The editorial board of the Journal of Medical Internet research is proud to be among the pioneers of successful and sustainable online journals and thanks everybody who has contributed to this success, most of all our authors, our peer-reviewers and our readers, who continue to give us useful feedback.

Gunther Eysenbach

Editor,

Journal of Medical Internet Research

Table 1. Electronic health journals selected for indexing in MEDLINE by the National Library of Medicine to date (as of Aug 28, 2001), in alphabetical order. Electronic journals were searched using Pubmed's journal browser, by entering the keyword *computer file*. Active journal titles (having published articles in 2001) are printed in bold

Title	No of articles indexed	Active	Issues indexed	MEDLINE Abbr. (and link to Pubmed)
AAPS pharmSci [computer file	---			AAPS PharmSci
BMC journals (consists of currently 58 specialty journals)	222	**yes**	2000 - today	**BMC Anesthesiol etc.**
CDR weekly [computer file] : communicable disease report	---			CDR Wkly (Online)
Cochrane database of systematic reviews (Online : Update Software)	1253	**yes**	2000 - today	**Cochrane Database Syst Rev**
Dermatology online journal [computer file].	86	**yes**	1997 - today	**Dermatol Online J**
EJIFCC [computer file] / IFFC	---			EJIFCC
Eurosurveillance weekly [computer file				--
Frontiers in bioscience [computer file] : a journal and virtual library.	565	**yes**	1996 - today	**Front Biosci**
Health information systems and telemedicine [computer file	5	no	1995-1996	Health Inf Syst Telemed
Journal of applied clinical medical physics [computer file] / American College of Medical Physics	---			J Appl Clin Med Phys
Journal of medical Internet research [computer file].	58	**yes**	1999 - today	**J Med Internet Res**
Journal of pharmacy & pharmaceutical sciences [computer file] : a publication of the Canadian Society for Pharmaceutical Sciences, Societe canadienne des sciences pharmaceutiques.	61	**yes**	1998 - today	**J Pharm Pharm Sci**
MedGenMed [computer file] : medscape general medicine.	123	**yes**	1999 - today	**MedGenMed**
Medscape women's health [computer file].	151	**yes**	1996 - today	**Medscape Womens Health**
Molecular vision [computer file].	172	**yes**	1995 - today	**Mol Vis**
Neurology & clinical neurophysiology [computer file] : NCN	---			Neurol Clin Neurophysiol
Neurosurgical focus [computer file	---			Neurosurg Focus
Online journal of issues in nursing [computer file].	26	**yes**	2001 - today	**Online J Issues Nurs**
transduction knowledge environment	---			Sci STKE
Sleep research online [computer file] : SRO.	67	**yes**	1998 - today	**Sleep Res Online**
The Online journal of current clinical trials [computer file	52	no	1992 - 1996	Online J Curr Clin Trials
The hospitalist [computer file] : the newsletter of the National Association of Inpatient Physicians	18	no	1997 - 1999	Hospitalist

Conflicts of Interest

None declared.

Submitted 28.08.01; ###Reviewer names will be inserted here### published 28.08.01.

Please cite as:
Eysenbach G
Journal of Medical Internet Research is now indexed in Medline
J Med Internet Res 2001;3(3):e25
URL: http://www.jmir.org/2001/3/e25/
doi:10.2196/jmir.3.3.e25
PMID:11720967

Original Paper

Evaluation of Controlled Vocabulary Resources for Development of a Consumer Entry Vocabulary for Diabetes

Timothy B Patrick[1], PhD; Harpreet K Monga[2], MD; MaryEllen C Sievert[3], PhD; Joan Houston Hall[4], PhD; Daniel R Longo[5], ScD

[1]Department of Health Management and Informatics, Center for Family Medicine Science, University of Missouri-Columbia, USA

[2]Department of Health Management and Informatics, University of Missouri-Columbia, USA

[3]School of Information Science and Learning Technology, Department of Health Management and Informatics, University of Missouri-Columbia, USA

[4]Dictionary of American Regional English Project, University of Wisconsin-Madison, USA

[5]Department of Family and Community Medicine, Center for Family Medicine Science, University of Missouri-Columbia, USA

Corresponding Author:
Timothy B Patrick, PhD
Department of Health Management and Informatics
University of Missouri-Columbia
324 Clark Hall
USA
Phone: +1 573 884 8155
Fax: +1 573 884 8155
Email: patrickt@health.missouri.edu

Abstract

Background: Digital information technology can facilitate informed decision making by individuals regarding their personal health care. The digital divide separates those who do and those who do not have access to or otherwise make use of digital information. To close the digital divide, health care communications research must address a fundamental issue, the consumer vocabulary problem: consumers of health care, at least those who are laypersons, are not always familiar with the professional vocabulary and concepts used by providers of health care and by providers of health care information, and, conversely, health care and health care information providers are not always familiar with the vocabulary and concepts used by consumers. One way to address this problem is to develop a consumer entry vocabulary for health care communications.

Objectives: To evaluate the potential of controlled vocabulary resources for supporting the development of consumer entry vocabulary for diabetes.

Methods: We used folk medical terms from the Dictionary of American Regional English project to create exended versions of 3 controlled vocabulary resources: the Unified Medical Language System Metathesaurus, the Eurodicautom of the European Commission's Translation Service, and the European Commission Glossary of popular and technical medical terms. We extracted consumer terms from consumer-authored materials, and physician terms from physician-authored materials. We used our extended versions of the vocabulary resources to link diabetes-related terms used by health care consumers to synonymous, nearly-synonymous, or closely-related terms used by family physicians. We also examined whether retrieval of diabetes-related World Wide Web information sites maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations can be improved by substituting a physician term for its related consumer term in the query.

Results: The Dictionary of American Regional English extension of the Metathesaurus provided coverage, either direct or indirect, of approximately 23% of the natural language consumer-term-physician-term pairs. The Dictionary of American Regional English extension of the Eurodicautom provided coverage for 16% of the term pairs. Both the Metathesaurus and the Eurodicautom indirectly related more terms than they directly related. A high percentage of covered term pairs, with more indirectly covered pairs than directly covered pairs, might be one way to make the most out of expensive controlled vocabulary resources. We compared retrieval of diabetes-related Web information sites using the physician terms to retrieval using related consumer terms We based the comparison on retrieval of sites maintained by non-profit healthcare professional organizations, academic organizations, or governmental organizations. The number of such sites in the first 20 results from a search was increased by substituting a physician term for its related consumer term in the query. This suggests that the Dictionary of American Regional English extensions of the Metathesaurus and Eurodicautom may be used to provide useful links from natural language consumer terms to natural language physician terms.

Conclusions: The Dictionary of American Regional English extensions of the Metathesaurus and Eurodicautom should be investigated further for support of consumer entry vocabulary for diabetes.

(J Med Internet Res 2001;3(3):e24) doi:10.2196/jmir.3.3.e24

KEYWORDS

Communication barriers; vocabulary, controlled; public health

Introduction

The consumer vocabulary problem

Digital information technology can facilitate informed decision making by individuals and can contribute to the realization of 2 important goals of the United States (US) health care establishment. The first of these goals is the US Healthy People 2010 goal of increasing "life expectancy and quality of life . . . by helping individuals gain the knowledge, motivation, and opportunities they need to make informed decisions about their health" [1]. The second goal is the US National Cancer Institute (NCI) goal of improving the "use of high quality, evidence-based cancer communications regardless of race, ethnicity, health status, education, income, age, gender, culture, or geographic region" [2]. It is certainly the case that " [a]t no other time in history has it been so easy for so many people to access such a vast wealth of information" [3], and it is also the case that consumers of health care, particularly laypersons, increasingly use the Web as a source of health care information [4]. The digital divide [5] has been described as the separation of populations into those who do and those who do not have access to or otherwise make use of digital information. The digital divide is problematic for the goals of Healthy People 2010 and NCI.

A promise of health care-communications research is that it can offer solutions to the digital divide. Ensuring universal access to information technology will not, however, be a sufficient solution. To close the digital divide, health care-communications research must address an issue that is more fundamental than the digital divide: the issue that we call the consumer vocabulary problem.

The consumer vocabulary problem is that consumers of health care, at least those who are laypersons, are not always familiar with the professional vocabulary and concepts used by providers of health care and by providers of health care information, and, conversely, health care and information providers are not always familiar with the vocabulary and concepts used by consumers. This bidirectional communication problem is more fundamental than the digital divide because it affects the use of information in all forms, both digital and non-digital. Access to a high speed network and a World Wide Web (Web) browser will serve little purpose for a health care consumer who lacks the vocabulary necessary to: ask questions about his or her health care, search for information to support his or her decisions about it, or understand such information when he or she does manage to access it.

Vocabulary differences in health are Communications: misunderstandings, outcomes, and information access

Potential misunderstandings attendant upon vocabulary differences in health care communications may reduce the quality of patient-physician interaction, result in poor health outcomes and patient satisfaction, impact consumer access to health care information, and have implications for informed consent.

The effect of vocabulary differences on patient-physician communication has long been recognized and studied by medical anthropologists and practitioners [6-8]. The importance of patient-physician communication for patient satisfaction and health outcomes is increasingly recognized [9] and communication skills are increasingly promoted as a necessary part of medical education [10-12].

Not all problems in patient-physician communication are the result of vocabulary differences. Some communication problems result from differences in values between the patient and the physician [13]. However, the quality of patient-physician communication may be compromised by differences in vocabulary that result in poor health outcomes. The patient's misunderstanding of professional language (or misunderstandings by the patient's caretaker), whether or not due to illiteracy [14], may lead to misunderstandings of crucial diagnostic or treatment information, and this may lead to a lack of proper compliance [15]. Previous studies, for example, have demonstrated a need for close attention to vocabulary differences in the education and treatment of patients with chronic conditions such as diabetes [16,17]and asthma [18-22].

Vocabulary and other language differences may be particularly significant in cross-cultural contexts [23,24]. For example, a physician not familiar with the term *the blood disease* as used by some African Americans to refer to cancer [25], might very easily mistake it for a synonym of *low blood* as used by some African Americans to refer to anemia [8]. For another example, consider *sugar*, *sugar diabetes*, and *diabetes*. A common belief among clinicians in the Midwestern and Southern US is that *sugar* is a term used by rural African Americans to refer to diabetes. However, a recent study [26] suggests that in some cases persons who say they have *sugar* have very different health beliefs regarding their disease than do persons who say they have *sugar diabetes* or *diabetes*. According to [26], persons who said they had *sugar* were "more likely to say that their condition was not serious and was curable." These different health beliefs have clear implications for compliance. If persons who believe they have *sugar* also believe their condition is not serious, they may be less diligent in matters of controlling their blood glucose levels.

With regard to health care-information access, in a previous study [27] we showed that the vocabulary used for the retrieval of health care information on the Web is problematic. In that study, we compared retrieval results for several commonly-used lexical variants, ie, different ways of writing the same term commonly used by laypersons. We found that the variant used clearly influenced the number of items retrieved. But, lexical variants are not the only sort of vocabulary barrier to health care information. In a recent Associated Press (AP) story [28], it was reported that former US Rep Geraldine Ferraro has *blood cancer*. Regardless of whether it was used appropriately in this case, *blood cancer* is a clear example of consumer vocabulary. If recent reports of the increasing use of the Web by consumers to get health care information [4] are accurate, it is likely that the AP story provided an impetus for some consumers to search for related information on the Web. Accordingly, we conducted a search (on June 25, 2001) on www.altavista.com for *blood cancer*. The search returned 2,616 results while a search with *multiple myeloma*(which the AP story used, apparently as a more specific term than *blood cancer)* returned 33,279 results. On www.google.com , the difference was 5,450 versus 46,700.The number of items retrieved is not at issue as much as the quality and relevance of the items retrieved [29-32]. Nevertheless, these results do suggest that there may be significant practical differences between consumer vocabulary and professional vocabulary for Web information retrieval.

Vocabulary differences affecting information retrieval and conceptual frameworks for reconciling those differences have long been studied in the fields of information science and informatics [33-36]. The problem of vocabulary differences is complicated for health care consumer informatics due to variations in the precision of the meanings of health care vocabulary as used by laypersons, and the lack of consistent semantic overlap of layperson vocabulary with professional vocabulary [6,35].

The lack of semantic overlap may have implications for informed consent. Perhaps the best-known US example of this is the infamous Tuskegee syphilis study [37]. The lack of informed consent by participants in the Tuskegee study may have been the result of mutual misunderstandings of the term *bad blood* as used by the rural African American subjects and as used by the US Public Health Physicians running the study. According to [37], some subjects of the study were told they had *bad blood*. The US Center for Disease Control (CDC) was in charge of the study in its later stages and an official of the CDC "stated that he understood the term 'bad blood' was a synonym for syphilis in the black community" [37]. However, also according to [37], a surviving subject of the study stated, ""That could be true. But I never heard no such thing. All I knew was that they just kept saying I had the bad blood -- they never mentioned syphilis to me, not even once." [37] . Indeed, such misunderstandings concerning *bad blood* may yet persist. In 1985, after the Tuskegee study had been widely reported in the popular press, *bad blood* was still reported to be a slang or nonstandard dialect synonym for syphilis [8].

The need for a consumer entry vocabulary for health care communications

The types of problems just described will differ in degree from case to case; some health care consumers are more familiar than others with professional vocabulary, and likewise some health care providers are more familiar than others with consumer vocabulary. In addition, there may be differences between countries, eg, differences between US physicians and British physicians in their familiarity with consumer language. Nevertheless, in general the kind of bidirectional communication problems we are discussing can make it difficult for individuals to get the information they need to support their decisions about their personal health care, and so these problems run counter to the goals of Healthy People 2010 and the US National Cancer Institute (NCI).

One way to address these bidirectional communication problems is to develop a *consumer-entry vocabulary* for health care communications. An entry vocabulary links commonly used terms to terms in some specialized vocabulary [36]. A consumer entry vocabulary for health care communications will link terms familiar to the consumer to possibly-unfamiliar professional terms used to index, describe, and communicate health care information. It will also function as an aid to communication between providers and patients (or their caretakers) [15]. From the point of view of the health care provider, the entry vocabulary will link terms familiar to the provider with commonly used, but perhaps unfamiliar, consumer terms. This bidirectional linking of consumer and professional vocabulary has elsewhere been described as the problem of reconciling groups of sub-languages [34].

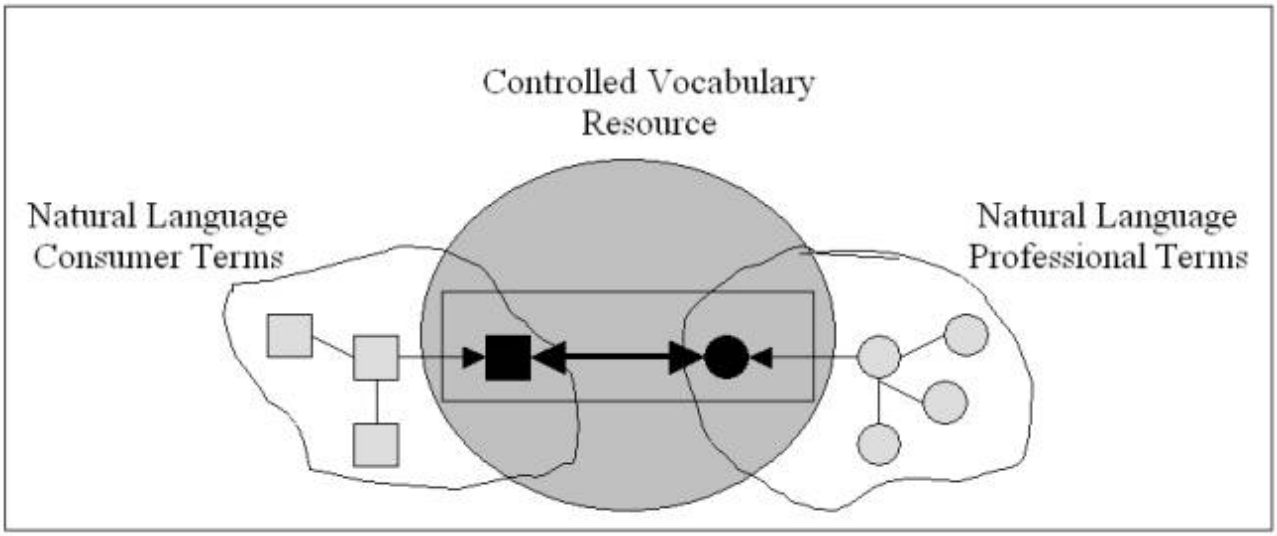
Overview of the study

A controlled vocabulary may be defined minimally as a concept-based vocabulary in which both the concepts and the terms used to express them are subjected to some level of control; only some subset of concepts is expressed, and only some terms and term variations are allowed as expressions of those concepts. Typically, an entry vocabulary will link natural language terms to terms in a controlled vocabulary. Since, however, health communication begins and ends with the natural language of professional practice and that of common sense, we were interested in using a controlled-vocabulary resource to link 2 natural language domains of vocabulary. (We use *vocabulary resource* here to allow for an organized collection of vocabularies, such as the Unified Medical Language System [UMLS] Metathesaurus.) We were *not* interested in questions of how much of a given natural language domain, whether professional or layperson, was contained in the controlled vocabulary resource. No matter how much is invested in controlled vocabulary resources, and no matter how comprehensive they appear, we assume that there will always be some natural residue of current vocabulary, whether professional or layperson, that is not contained in any controlled vocabulary resource, and that it may be associated with bidirectional communication problems in health care. One way to reap the benefits of the investment that has been made in building vocabulary resources may be to use them to provide

indirect, though perhaps only approximate, links between natural language domains of vocabulary.

By using a controlled vocabulary resource to link 2 natural language domains of vocabulary, we may provide entry vocabulary from one domain to the other. Specifically, in this pilot study we evaluated controlled vocabulary resources for their capacity to support an entry vocabulary from natural language health care consumer vocabulary to natural language health care professional vocabulary. The conceptual model we used is depicted in Figure 1.

Figure 1. A controlled vocabulary resource providing a bridge between consumer terms and professional terms



As shown in Figure 1, a controlled vocabulary resource (such as the Metathesaurus) may function as a bridge between natural language consumer terms and natural language professional terms. The point of the model is to address the need for linking natural language domains, and depends on there being some overlap between the natural language consumer vocabulary, the controlled vocabulary resource, and the natural language professional vocabulary. This overlap allows us to make use of the controlled vocabulary resource to link the natural language consumer terms to the natural language professional terms. Figure 1 shows 3 natural language consumer terms linked to 4 natural language professional terms.

In Figure 1, assume that the 3 consumer terms are located in the same semantic neighborhood [38], that is, are synonymous, nearly synonymous, or otherwise closely related. Also assume that the 4 professional terms are themselves colocated in their own semantic neighborhood. Assume that the controlled vocabulary resource contains 1 of the consumer terms and 1 of the professional terms. Also assume that the controlled vocabulary resource, according to its internally specified rules and structure, locates the contained consumer term and the contained professional term in the same semantic neighborhood. Finally, assume that the relationship between semantic neighborhoods is transitive, that is, if the consumer neighborhood may be linked to the controlled neighborhood, and the controlled neighborhood may be linked to the professional neighborhood, then the consumer neighborhood may be linked to the professional neighborhood. In that case the controlled vocabulary resource may be used to indirectly relate the remaining consumer and professional terms, thus providing a consumer entry vocabulary to the natural language professional vocabulary. The character of that relation will vary with the properties that define the respective semantic neighborhoods, as well as with the degree of control that characterizes the vocabulary resource.

In order to link natural language consumer terms to natural language professional terms we must, according to our model, group the consumer terms and professional terms into semantic neighborhoods (**Methods, Step 3**). In addition, the links between consumer terms and professional terms provided by a vocabulary resource are appropriate only if those terms should be linked. Thus, in order for us to evaluate whether a given vocabulary resource may be useful in linking natural language consumer terms to natural language professional terms, we need to independently determine whether the consumer terms should be colocated with the professional terms in the same semantic neighborhood (**Methods, Step 3**). Finally, while the conceptual appropriateness of the links provided by a vocabulary resource is important, we are most interested in the practical value of the resulting consumer entry vocabulary with respect to the previously mentioned goals of Healthy People 2010 and NCI. Thus, we need to evaluate the consumer entry vocabulary with respect to some context of information use or health care communication that can provide support for an individual's decision making about his or her personal health care. In this pilot study, we evaluated the consumer entry vocabulary for its support for Web retrieval of health care information (**Methods, Step 8**). Certainly, Web retrieval should be evaluated not only for the *amount* of information retrieved but also its *quality*. For this pilot study, we examined Web retrieval, using the entry vocabulary, for sites maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations. We recognize that maintenance of a site by a professional, academic, or governmental organization does not guarantee that the site is of high quality, but we do think it does increase the likelihood that the site is of high quality.

To keep this initial exploratory study to a manageable size, we have restricted ourselves to consumer vocabulary related to diabetes, and have attempted to relate this consumer vocabulary to diabetes vocabulary used by US family physicians. As such, this study complements a previous study concerning consumer and professional vocabulary for diabetes that focused on the Read Thesaurus [35].

Research Questions

We addressed 2 research questions:

1. To what extent can the Metathesaurus [39], the Eurodicautom of the European Commission's Translation Service [40], and the European Commission Glossary of popular and technical medical terms (or European Glossary for short) [41], as extended by the Dictionary of American Regional English (DARE) project [42-45], be used to derive pairs of consumer and physician terms concerning diabetes that are in the same semantic neighborhood?
2. Can retrieval of diabetes-related Web information sites maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations be improved by substituting a physician term for its paired consumer term in a Web query?

Vocabulary Resources

The Metathesaurus is a large database developed by the National Library of Medicine that links together terms from over 50 health care vocabularies. The Metathesaurus links terms together when they express the same, or nearly the same, concepts. When terms express the same concept, they are assigned, in the language of the Metathesaurus, to the same *metaconcept*. For example, the terms *Unspecified diabetes mellitus* and *Diabetes mellitus* are assigned to the same metaconcept.

The Eurodicautom is the multilingual terminological database of the European Commission's Translation Service. The Eurodicautom contains 124,551 entries for medicine, with many of these in English. Altogether, the Eurodicautom contains 990,672 English term entries. The Eurodicautom is organized by terminology collections maintained by different terminology offices, and, like the Metathesaurus, links different terms together when they express the same, or nearly the same, concepts. For example, the Eurodicautom entry for *hypertension*, in the terminology collection EUT97, includes *hyper blood pressure* as a synonym for *hypertension*.

The European Glossary was developed by the European Commission and is separate from the Eurodicautom. The European Glossary contains pairs of synonymous or nearly-synonymous consumer terms and professional terms. For example, the European Glossary pairs the consumer term holding your breath with the professional term hypoventilation.

DARE is intended to "document the varieties of English that are not found everywhere in the United States--those words, pronunciations, and phrases that vary from one region to another, that we learn at home rather than at school, or that are part of our oral rather than our written culture" [45]. DARE is based on fieldwork involving interviews carried out in 1,002 US communities in all 50 US states between 1965 and 1970. DARE fieldworkers collected a variety of responses to this interview question concerning diabetes, for example: "When a person has too much sugar in his blood and may have to take insulin for it, you'd say he has_____?"

The responses included *diabetes*, *high blood*, *sugar blood*, *sugar diabetes*, and *the sugar*. Additional terms included in DARE are provided by a comprehensive collection of written materials (diaries, letters, novels, histories, biographies, newspapers, and government documents) that cover US history from the colonial period to the present [45].

Methods

Step 1. We selected consumer terms related to diabetes from a list of terms previously extracted from 2 consumer sources. The first source of terms was a corpus of 1,500 e-mail questions submitted by consumers to University of Missouri Healthcare's On Call Online Web health care advice site during the period 1997-1999 (University of Missouri-Columbia Human Subjects Research Institutional Review Board [IRB] exemption approval #6843.) The second source of terms was a log of 348,000 queries submitted to the consumer portal of a major for-profit Web health care information site (by agreement with the site owner, we will not mention the name of the site.) The diabetes-related terms were selected by searching a list of previously-extracted terms for those containing the substrings *diab*, *dib*, *gluco*, or *insul*. We corrected the spelling of the consumer terms and eliminated duplicates. (We recognize at the outset a limitation of this study-that while these terms are ostensibly *consumer* terms, their pedigree as *layperson* terms must remain somewhat in doubt, since there is no clear evidence, particularly for terms extracted from the query log, to settle the question whether a health professional or a layperson entered the query.)

Step 2. We next extracted physician terms related to diabetes from terms previously extracted from a corpus of 25,000 family-medicine progress notes authored by University of Missouri Healthcare family physicians during the period 1999-2000 (University of Missouri-Columbia IRB exemption approval #7054.) As with the consumer terms, physician terms related to diabetes were selected by searching a list of previously extracted terms for terms containing the substrings *diab*, *dib*, *gluco*, or *insul*. We corrected the spelling of the physician terms and eliminated duplicates.

Step 3. Two co-authors (H.K.M. and T.B.P.) used the following informal procedure to produce pairs of consumer and physician terms that were judged to be in the same semantic neighborhood: (1) each consumer term was matched to one or more synonymous, nearly-synonymous, or closely-related physician terms; (2) the consumer terms and physician terms were separately grouped into sets of synonymous, nearly-synonymous, or closely-related terms; (3) each group of consumer terms was matched to one or more groups of physician terms; and (4) the consumer-term-physician-term matches were adjusted based on (3). H.K.M. did the initial matching and groupings of terms and then T.B.P. reviewed the matches and groupings. Disagreements were discussed and reviewed using medical dictionaries until consensus was achieved. This informal procedure was not intended to produce pairs of

strictly-synonymous consumer terms and physician terms, but rather to produce pairs of terms we judged to be synonymous, nearly synonymous, or closely related. The approach was more informal and subjective than we would have liked. We would have preferred a method that would have allowed us to generate a logically-rigorous definition of the properties defining the semantic neighborhoods. However, objective measures of semantic locality , that would allow us to more-rigorously specify those properties, were not clearly applicable in this case. For example, we concluded that text corpus analysis such as word or term co-occurrence analysis [38,46] would not be applicable since we had no larger textual context for the set of terms extracted from the query log. The use of other empirical methods to establish semantic locality, such as focus groups, interviews, or surveys with laypersons and family physicians, was beyond the exploratory scope of this pilot study.

Step 4. We created extensions for the Metathesaurus, Eurodicautom, and the European Glossary with terms taken from DARE. We linked the DARE terms as a group to each vocabulary resource as near-synonyms of the general term *diabetes*. Even though there appear to be cases of folk-belief systems and associated vocabulary that are extremely precise in medical meaning [6], we adopted this conservative approach in order to avoid assigning inappropriately-specific meaning to the folk terms. However, this conservative approach may be inconsistent with the finding of [26] that in some cases persons who say they have *sugar* have very different health beliefs regarding their disease than do persons who say they have *sugar diabetes* or *diabetes*.

Step 5. Since we maintained the Metathesaurus and the European Glossary in local relational databases, we were able to easily automate searching of those databases. We used normalized string matches to search for the consumer terms and the physician terms in the DARE extensions of the Metathesaurus and the European Glossary. The normalized string associated with a term is produced by a rule-based transformation of the string of characters that carries the term, for example, the string of characters *eye doctors* carries the term *eye doctors*. According to the set of rules

1. Make plural singular
2. Put words in alphabetic order;

it may be normalized to the string *doctor eye*. Normalized string matches allow for more liberal term-to-term matches by preventing matches from being blocked by minor lexical variation. We created normalized-string indexes for the DARE extensions of the Metathesaurus and the European Glossary using the UMLS NORM program [39]. We then normalized the consumer and physician terms and matched them to strings in normalized-string indexes.

Step 6. The Eurodicautom is only available through a Web interface so we were not able to perform normalized string matches against it (according to an e-mail from J Vega J, official contact for questions and feedback about the Eurodicautom, 2001 Apr). Instead, we used the Eurodicautom interface to perform all-words-plus-context and truncate-only-if-no-match English-language searches against the Eurodicautom. In these searches, a term would match if all of its component words were contained in the Eurodicautom term. The use of truncation provides matches in cases of minor lexical variation. We created a table of DARE terms that we had linked to Eurodicautom terms, and a normalized string index. We searched the DARE extension of the Eurodicautom for the consumer and physician terms.

Step 7. We next collected pairs of consumer and physician terms that were contained either directly or indirectly in the DARE extensions of the 3 vocabulary resources. Some consumer and physician terms matched directly, via a string match, to a vocabulary resource, while some matched only indirectly by virtue of their semantic neighborhood. For example, the consumer term *diabetes diet* matched directly, while the consumer term *food for diabetics* only matched indirectly; *food for diabetics* matched indirectly because it was synonymous, nearly synonymous, or closely related to *diabetes diet*. We said that a consumer-term-physician-term pair was *directly covered* by a vocabulary resource when both of the following occurred:

- Both the consumer term and the physician term matched directly, via string matches, to the vocabulary resource
- The vocabulary resource, according to its internally-specified rules and structure, located the contained consumer term and the contained professional term in the same semantic neighborhood.

We considered the second condition satisfied for the Metathesaurus when either of the terms were associated with the same metaconcept, or were associated with distinct metaconcepts that are related according to the Metathesaurus related metaconcepts (MRREL) table. We considered the second condition satisfied for the Eurodicautom when the terms were matched to the same Eurodicautom entry - and similarly for the European Glossary.

A pair was *indirectly covered* by a vocabulary resource when both terms matched to terms in another pair that was directly covered, but at least one match was indirect. A vocabulary resource provided *partial indirect coverage* for a pair when only one member of the pair matched directly, and it provided *full indirect coverage* of the pair when neither member of the pair matched directly.

In order to provide preliminary results for our first research question, we compared the number of consumer and physician terms that directly and indirectly matched to the DARE extensions of the 3 resources, as well as the consumer-term-physician-term pairs that were directly and indirectly covered.

Step 8. In order to provide preliminary results for our second research question, coauthor T.B.P. selected 5 pairs that were provided partial-indirect coverage by the Metathesaurus, the Eurodicautom, or the European Glossary. For each of these pairs, the consumer term matched indirectly and the physician term matched directly. T.B.P. performed exact-phrase searches on www.altavista.com using the consumer term and physician term of each pair. Exact-phrase searches were used in order to better discriminate between the retrieval effects of the consumer term and the physician term. The results for the consumer term were compared to the results for the physician term. The results

were compared for their respective numbers of relevant sites maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations. Since anecdotal evidence suggests that consumers are likely to not look past the first few pages of results returned by a Web search engine, for any given search we examined only the first 20 results listed. We did not count a site if it was unavailable and we counted 2 or more pages from the same Web site as representing a single site.

Results

Results for research question 1

Our raw data consisted of 909 consumer terms (eg, *borderline diabetes*, *sugar dibetes*, *suger diabetes*) and 938 physician terms (eg, *insulin insensitivity* and *diabetic flow sheet*). We corrected spelling and merged duplicate terms. We then selected 86 consumer terms and matched them to 125 physician terms for a total of 225 pairs of terms. Table 1 shows example pairs of terms.

Table 1. Example Pairs of Consumer and Physician Terms

Consumer Term	Physician Term
diabetic food	healthy diabetic diet
diabetic food	diabetic diet plan
diabetes recipes	healthy diabetic diet
diabetes recipes	diabetic diet plan

We directly matched 27 consumer terms against the DARE extension of the Metathesaurus, 27 consumer terms against the DARE extension of the Eurodicautom, and 5 consumer terms against the DARE extension of the European Glossary. We indirectly matched 13 consumer terms against the DARE extension of the Metathesaurus, 11 consumer terms against the DARE extension of the Eurodicautom, and 1 consumer term against the DARE extension of the European Glossary.

We directly matched 39 physician terms against the DARE extension of the Metathesaurus, 23 physician terms against the DARE extension of the Eurodicautom, and 2 physician terms against the DARE extension of the European Glossary. We indirectly matched 29 physician terms against the DARE extension of the Metathesaurus, 7 physician terms against the DARE extension of the Eurodicautom, and no physician terms against the DARE extension of the European Glossary.

The DARE extension of the Metathesaurus directly covered 17 consumer-term-physician-term pairs, provided partial indirect coverage for 22 pairs, and full indirect coverage for 12 pairs. The DARE extension of the Eurodicautom directly covered 8 pairs, provided partial indirect coverage for 19 pairs, and full indirect coverage for 9 pairs. The DARE extension of the European Glossary directly covered 2 pairs, provided partial indirect coverage for 1 pair, and full indirect coverage for no pairs. Table 2 shows examples of directly and indirectly covered pairs of terms.

Table 2. Examples of Directly and Indirectly Covered Pairs of Terms

Consumer Term	Physician Term	Extended Metathesaurus	Extended Eurodicautom	Extended European Glossary
sugar diabetics	diabetes	partial indirect	partial indirect	partial indirect
sugar diabetes	diabetes	direct	direct	direct
diabetes	diabetes	direct	direct	direct
diabetes recipes	diabetic diet	partial indirect	partial indirect	not covered
diabetic food	healthy diabetic diet	full indirect	partial indirect	not covered
food to prepare for diabetic	healthy diabetic diet	full indirect	full indirect	not covered
diabetic diet	diabetic diet	direct	direct	not covered

Results for Research Question 2

Table 3 compares the results of Web searches on www.altavista.com for the consumer terms and physician terms. The *Quality (out of first 20 results)* column under Results contains the number of sites out of the first 20 search results that were maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations. In every case, the physician term produced results that contained more such sites than the consumer term.

Table 3. Comparison of Results of Consumer Term Searches and Physician Term Searches for Pairs with Partial Indirect Coverage

Consumer Term			Physician Term			Vocabulary Resource Providing Partial Indirect Coverage
Term	Results		Term	Results		
	Total	Quality* (out of first 20 results)		Total	Quality* (out of first 20 results)	
diabetes recipes	457	1	diabetic diet	83,778	2	Metathesaurus; Eurodicautom
foods for diabetics	100	0	diabetic diet	83,778	2	Metathesaurus; Eurodicautom
diabetes in children	2,016	1	juvenile diabetes	18,482	6	Eurodicautom
diabetic pregnancy	562	0	gestational diabetes	13,498	2	Metathesaurus
sugar diabetics	92	0	diabetes	1,038,268	6	Metathesaurus; Eurodicautom; European Glossary
organizations						

Discussion

The results for our first research question for the DARE extension of the Metathesaurus appear somewhat promising. The DARE extension of the Metathesaurus provided coverage, either direct or indirect, for approximately 23% of the natural language consumer-term-physician-term

pairs. The results provided by the Eurodicautom extension are less promising, since it provided coverage for only 16% of the term pairs. (The results for the European Glossary were negligible.) However, what we find somewhat promising overall is that both the Metathesaurus and the Eurodicautom extensions indirectly covered more pairs than they directly covered. A high percentage of covered term pairs, with more indirectly-covered pairs than directly-covered pairs, might constitute an efficient use of expensive controlled vocabulary resources for health care communications.

The results for the second research question are also somewhat promising. Notwithstanding the small sample size, in every case the natural language physician term produced better results than the consumer term for sites maintained by nonprofit health care professional organizations, academic organizations, or governmental organizations. This suggests that the DARE extensions of the Metathesaurus and Eurodicautom may be used to provide useful links from natural language consumer terms to natural language physician terms. It might be argued that these results are not particularly meaningful since

We performed our study using terms rather than actual consumer queries

Maintenance of a site by a nonprofit health care professional, academic, or governmental organization is not a certain indicator of the quality of the site.

In response to the first point, we point out that some of the terms we used in our www.altavista.com queries were extracted from longer consumer e-mail messages, and some were used as actual queries to a Web information site. These terms reflect the actual terms used by consumers, although admittedly only the latter can be said to represent actual consumer Web queries (subject to the limitation noted earlier in **Methods, Step 1** , concerning their pedigree as *layperson* terms). It is also true that we corrected the spelling of some consumer terms. However, the effect of misspellings on retrieval is not clear and is a subject of our current investigations. Finally, many Web search programs that accept natural language input extract terms from it and use them for retrieval as we did in our study.

In response to the second point, we agree that maintenance of a site by a professional, academic, or governmental organization does not guarantee that the site is of high quality, but, as stated in **Overview of the study** , we do think such maintenance increases the likelihood of high quality.

Limitations of the study

Although our results are somewhat promising, they are limited in several ways. One limitation of our study is that www.altavista.com, like most search engines, does not limit its exact-phrase results to results for true exact phrases, but will also report results where the words in the phrase are included in the document, but are not immediately next to each other. Thus, our results do not strictly discriminate between the retrieval effectiveness of the consumer terms and the physician terms. Our results do, however, reflect the situation actually faced by consumers, and to that extent are indicative of the relative effectiveness of the consumer and physician terms.

Another limitation of the study is that one of our sources of consumer terms, the Web query log, did not lend itself to term co-occurrence as a quantifiable measure of semantic locality, nor to other data that might be used to cluster terms into semantic neighborhoods. But, perhaps more importantly, we did not evaluate the final term pairs with respect to other quantitative evidence derived from consumer and physician surveys, nor did we evaluate them with respect to qualitative evidence derived from focus groups and interviews with consumers and physicians. Thus, even if the term pairs appear useful according to the standard of quality that we used in this project, they might not be more generally useful, because they might not constitute appropriate meaning-preserving (or

meaning-warping) steps from the language of consumers to the language of family physicians.

Finally, although this study may serve in part as a pilot for larger studies of access to consumer health care information, a general limitation of the study is that we focused on the Web, and did not address the use of other venues of medical information, such as newspapers, magazines, and television. We recognize that these other sources and formats of health information require similar investigation if consumers are to get the information they need to support their decisions about their personal health care. Indeed, as we said at the outset, the bidirectional communication problems we have been considering are more fundamental than the Web and the digital divide. More work needs to be done with physicians, patient-education professionals, and consumers to further articulate the extent of these problems and to develop methods for their resolution.

Acknowledgments

This project was supported in part by grant LM07089 from the National Library of Medicine. The *Dictionary of American Regional English* is supported financially by the National Endowment for the Humanities, and by National Science Foundation grants No. 89-08905, 94-08728, and 97-2906.

Conflicts of Interest

None declared.

References

1. ; United States Dept. of Health and Human Services. Healthy People 2010: Understanding and Improving Health, 2nd edition. Washington, DC: Dept of Health and Human Services; Nov 1, 2000.
2. ; National Cancer Institute. Centers of excellence in cancer communications research. Publication RFA-CA-01-019. URL: http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-01-019.html [accessed 2001 Jun]
3. ; National Cancer Institute. Scientific priorities for cancer research: extraordinary opportunities. URL: http://plan2002.cancer.gov/scpcomms.htm [accessed 2001 Jun]
4. Taylor H. The Harris Poll #44: Explosive growth of "cyberchondriacs" continues. 2000 Sep 11. URL: http://www.gsbc.com/harris_poll/index.asp?PID=104 [accessed 2001 May]
5. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the Internet, and the digital divide. Health Aff (Millwood) 2000 Dec;19(6):255-265 [FREE Full text] [Medline: 21033799] [doi: 10.1377/hlthaff.19.6.255]
6. Burnum JF. Dialect is diagnostic. Ann Intern Med 1984 Jun;100(6):899-901. [Medline: 84201224]
7. Snow LF. Folk medical beliefs and their implications for care of patients. A review bases on studies among black Americans. Ann Intern Med 1974 Jul;81(1):82-96. [Medline: 74266098]
8. Sugarman J, Butters RR. Understanding the patient: medical words the doctor may not know. N C Med J 1985 Jul;46(7):415-417. [Medline: 85296418]
9. Duffy FD. Dialogue: the core clinical skill. Ann Intern Med 1998 Jan 15;128(2):139-141 [FREE Full text] [Medline: 98085841]
10. Salmon P, Peters S, Stanley I. Patients' perceptions of medical explanations for somatisation disorders: qualitative analysis. BMJ 1999 Feb 6;318(7180):372-376 [FREE Full text] [PMC: 9933202] [Medline: 99132095]
11. Skelton JR, Hobbs FD. Descriptive study of cooperative language in primary care consultations by male and female doctors. BMJ 1999 Feb 27;318(7183):576-579 [FREE Full text] [PMC: 10037635] [Medline: 99156801]
12. Steel N. Communicating with patients. Specialist training should include communication skills. BMJ 1999 Jan 2;318(7175):60. [Medline: 99091487]
13. Freeman J, Loewe R. Barriers to communication about diabetes mellitus. Patients' and physicians' different view of the disease. J Fam Pract 2000 Jun;49(6):507-512. [Medline: 20378052]
14. Kefalides PT. Illiteracy: the silent barrier to health care. Ann Intern Med 1999 Feb 16;130(4 Pt 1):333-336 [FREE Full text] [Medline: 99149978]
15. Falvo DR. Effective Patient Education: a Guide to Increased Compliance, 2nd edition. Gaithersburg MD: Jones & Bartlett Publishers; Jan 15, 1994.
16. Leichter SB, Nieman JA, Moore RW, Collins P, Rhodes A. Readability of self-care instructional pamphlets for diabetic patients. Diabetes Care 1981 Dec;4(6):627-630. [Medline: 83026807]
17. Reid JC, Klachko DM, Kardash CA, Robinson RD, Scholes R, Howard D. Why people don't learn from diabetes literature: influence of text and reader characteristics. Patient Educ Couns 1995 Feb;25(1):31-38. [Medline: 95327574] [doi: 10.1016/0738-3991(94)00688-I]
18. Hardie GE, Janson S, Gold WM, Carrieri-kohlman V, Boushey HA. Ethnic differences: word descriptors used by African-American and white asthma patients during induced bronchoconstriction. Chest 2000 Apr;117(4):935-943 [FREE Full text] [Medline: 20231574] [doi: 10.1378/chest.117.4.935]

19. Weiland SK, Kugler J, Von Mutius E, Schmitz N, Fritzsch C, Wahn U, et al. [The language of pediatric asthma patients. A study of symptom description]. Monatsschr Kinderheilkd 1993 Nov;141(11):878-882. [Medline: 94111734]
20. Mahler DA, Harver A. Do you speak the language of dyspnea? Chest 2000 Apr;117(4):928-929 [FREE Full text] [Medline: 20231570] [doi: 10.1378/chest.117.4.928]
21. Moy ML, Lantin ML, Harver A, Schwartzstein RM. Language of dyspnea in assessment of patients with acute asthma treated with nebulized albuterol. Am J Respir Crit Care Med 1998 Sep;158(3):749-753 [FREE Full text] [Medline: 98402649]
22. Flynn CA, Barash A. Do African American asthmatics perceive and describe their asthma symptoms differently than white asthmatics? J Fam Pract 2000 Aug;49(8):688, 759. [Medline: 20401956]
23. Hornberger J, Itakura H, Wilson SR. Bridging language and cultural barriers between physicians and patients. Public Health Rep 1997 Oct;112(5):410-417. [Medline: 97464679]
24. Woloshin S, Bickell NA, Schwartz LM, Gany F, Welch HG. Language barriers in medicine in the United States. JAMA 1995 Mar 1;273(9):724-728. [Medline: 95156732] [doi: 10.1001/jama.273.9.724]
25. ; Texas Cancer Council. Practical guidelines for the development of audiovisual cancer education materials for African Americans. URL: http://tcdc.uth.tmc.edu/pgavemaa/6.4.html [accessed 2001 Jun]
26. Schorling JB, Saunders JT. Is "sugar" the same as diabetes? A community-based study among rural African-Americans. Diabetes Care 2000 Mar;23(3):330-334 [FREE Full text] [Medline: 20324712]
27. Sievert ME, Patrick TB, Reid JC. Need a bloody nose be a nosebleed? or, lexical variants cause surprising results. Bull Med Libr Assoc 2001 Jan;89(1):68-71. [PMC: 11209803] [Medline: 21077106]
28. ; Associated Press. Ferraro battling blood cancer. Columbia Daily Tribune 2001. 2001 Jun 19. p. Sect. 8A URL: http://archive.showmenews.com/2001/Jun/20010619News013.asp [accessed 2001 Aug]
29. ; Health Summit Working Group. Criteria for assessing the quality of health information on the Internet - policy paper. URL: http://hitiweb.mitretek.org/docs/policy.pdf [accessed 2001 May]
30. Wootton JC. The quality of information on women's health on the Internet. J Womens Health 1997 Oct;6(5):575-581. [Medline: 98019829]
31. Oermann MH, Wilson FL. Quality of care information for consumers on the Internet. J Nurs Care Qual 2000 Jul;14(4):45-54. [Medline: 20338358]
32. Adelhard K, Obst O. Evaluation of medical internet sites. Methods Inf Med 1999 Jun;38(2):75-79. [Medline: 99360179]
33. Lancaster FW. Vocabulary Control for Information Retrieval. Arlington, Va: Info Resources Press; Feb 1, 1986.
34. Brown PJ, Price C, Sonksen PH. Evaluating the terminology requirements to support multi-disciplinary diabetes care. Proc AMIA Annu Fall Symp 1997:645-649. [Medline: 98020576]
35. Brown PJB, Price C, Cox YM. Patient language - evaluating its relationship to a clinical thesaurus. Proc AMIA Annu Fall Symp 1997:917.
36. Buckland M, Chen A, Kim Y, Lam B, Larson R, et al. Mapping Entry Vocabulary to Unfamiliar Metadata Vocabularies. D-Lib Magazine. 1999. (1) URL: http://www.dlib.org/dlib/january99/buckland/01buckland.html [accessed 2001 Aug]
37. Jones JH. Bad BLood: The Tuskegee Syphilis Experiment, Revised Edition. New York: Free Press; Jan 15, 1993.
38. Bodenreider O, Nelson SJ, Hole WT, Chang HF. Beyond synonymy: exploiting the UMLS semantics in mapping vocabularies. Proc AMIA Symp 1998:815-819. [Medline: 99123237]
39. ; National Library of Medicine. UMLS Knowledge Sources, 11th edition 2000.
40. ; European Commission's Translation Service. Eurodicautom. URL: http://eurodic.ip.lu/cgi-bin/edicbin/EuroDicWWW.pl [accessed 2001 Aug]
41. ; European Commission. Multilingual glossary of technical and popular medical terms in nine European languages. URL: http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html [accessed 2001 May]
42. Cassidy FG. Dictionary of American Regional English. Cambridge, Ma: Belknap Press of Harvard University Press; 1985, Vol. 1.
43. Cassidy FG. Dictionary of American Regional English. Cambridge, Ma: Belknap Press of Harvard University Press; 1991, Vol. 2.
44. Cassidy FG. Dictionary of American Regional English. Cambridge, Ma: Belknap Press of Harvard University Press; 1996, Vol. 3.
45. ; DARE. Dictionary of American Regional English. URL: http://polyglot.lss.wisc.edu/dare/dare.html [accessed 2001 May]
46. Biber D, Conrad S, Reppen R. Corpus Linguistics : Investigating Language Structure and Use (Cambridge Approaches to Linguistics). Cambridge, UK: Cambridge University Press; Apr 23, 1998.

Abbreviations

AP: Associated Press
CDC: Center for Disease Control
DARE: Dictionary of American Regional English
IRB: Institutional Review Board
NCI: National Cancer Institute

UMLS: Unified Medical Language System
US: United States

submitted 16.05.01; peer-reviewed by W Hersh, P Brown; comments to author 05.06.01; revised version received 20.07.01; accepted 08.08.01; published 28.08.01.

Please cite as:
Patrick TB, Monga HK, Sievert MC, Hall JH, Longo DR
Evaluation of Controlled Vocabulary Resources for Development of a Consumer Entry Vocabulary for Diabetes
J Med Internet Res 2001;3(3):e24
URL: http://www.jmir.org/2001/3/e24/
doi:10.2196/jmir.3.3.e24
PMID:11720966

Original Paper

Use of the Internet by Patients Before and After Cardiac Surgery: An Interdisciplinary Telephone Survey

Monica Murero[1], PhD; Giuseppe D'Ancona[2], MD; Hratch Karamanoukian[2], MD

[1]International Institute of Infonomics, Herleen, Netherlands

[2]Center for Less Invasive and Robotic Heart Surgery and Kaleida Health, Buffalo NY, USA

Corresponding Author:
Monica Murero, PhD
International Institute of Infonomics
POB 2606 NL 6401
DC Heerlen
The Netherlands
Phone: +31 45 570 76 90
Fax: +31 45 570 62 62
Email: murero@infonomics.nl

Abstract

Background: Little is known about to what extent patients who underwent medical treatment access the Internet and whether they benefit from consulting the Internet.

Objective: To understand if cardiopathic patients use the Internet for health-related information and whether they find retrieved information understandable and useful.

Methods: Telephone interviews, using a semi-structured questionnaire, were conducted with 82 patients who had undergone off-pump coronary-artery bypass grafting at the Center for Less Invasive and Robotic Heart Surgery in Buffalo, New York, USA. Study design was multidisciplinary, combining expertise of medical and communication science. Sources of medical information were identified (doctor, Internet, magazines, newspapers, television, radio, family members). Accessibility, quality, and readability of Internet medical information from the patients' point of view were investigated.

Results: Out of 82 patients, 35 (35/82, 42.7%) were Internet users. Internet users had a significantly higher education level than Internet non-users (college education: 42.9% of users, 10.6% of non-users; $P < .001$). Among the Internet users, 18 (18/35, 51.4%) had used the Internet for retrieving medical information; 17 (17/35, 48.6%) had not. No statistically significant differences in demographic data were found when comparing these 2 sub-groups of patients. Family-members' involvement was high (15/18, 83.3%). Internet medical information was rated helpful in most cases; readability was acceptable for only 3 patients (3/18, 16.7%). To improve on-line medical information, all patients interviewed suggested sites designed by their physicians.

Conclusions: Although 1 in 5 patients in our sample has used the Internet to retrieve medical information, the majority of them experiences difficulties comprehending the information retrieved. Health-care providers' should provide Internet medical information that is adequate for the non-medical public's needs.

(J Med Internet Res 2001;3(3):e27) doi:10.2196/jmir.3.3.e27

KEYWORDS

Internet/statistics and numerical data; Questionnaires; Data Collection; Attitude to Computers; Coronary Artery Bypass; Patient Education; Readability

Introduction

It is crucial that patients have adequate medical information to educate, sensitize, and prepare themselves before and after undergoing medical procedures. Despite this, many fields of medicine remain unexplored by and unknown to the general public. Open-heart surgery remains one of the least-known fields, where myths and realities still coexist in the imaginations of patients. Coronary artery disease is one of the leading causes of mortality in the industrialized world; cardiopathic patients should be informed about indications, risks, and benefits of cardiac surgery, and its possible alternatives. Recent innovations in cardiac surgery (minimally-invasive and robotic procedures) have significantly decreased morbidity and mortality rates [1,2]. In many cases, coronary surgery can be performed without arresting the heart and without using the heart-lung machine

[1]. Thanks to the contributions of a few pioneers, today open-heart surgery can be performed, in some instances, via small (5-cm long) incisions or even endoscopically, using robotic-assisted surgery [2].

Although clinicians and surgeons are aware of these new achievements, the majority of cardiopathic patients and their families, remain unaware of them. Adequate information on the Internet could increase patients' knowledge about these innovative surgical techniques and clarify their applications and limitations.

The introduction and popularization of the Internet is changing the way health-care providers and the general public search for and retrieve medical information. The Internet makes possible immediate access to and acquisition of encyclopedic information. In the field of cardiac surgery, for example, several authors have described and discussed the use of the Internet for educating cardiac surgery patients on-line [3,4,5,6]. However, little evidence is available about whether and to what extent this patient population actually uses the Internet. Although previous surveys have described the population of 'health-care seekers' that access and search Internet medical sites daily, they mostly do not distinguish between (a) patients that have already undergone or will undergo medical treatment and (b) the general public, including patients' family members, who may or may not have personal health-related problems [5]. Because most health resources available on the Internet are tailored for health-care providers, there is a question as to whether - even if patients access the Internet - patients understand and benefit from the information in these resources. We therefore focused our study on 2 central research questions: Is the Internet specifically used by cardiopathic patients to obtain health-related information? If this is the case, is on-line medical information understandable and useful for this non-medical public?

Methods

Eighty-two patients that had previously undergone coronary artery bypass grafting (CABG) were interviewed by telephone during March and April 2001. The interdisciplinary approach followed in our research is built around a group of experts in the field of medical and communication science. Working together, the cardiac surgeon who performed the operations and a communication scientist specializing in Internet studies designed, tested, and conducted the interviews. To collect in-depth, quantitative and qualitative data, we used a semi-structured questionnaire [8]. During a pilot phase 2 alternative forms of the telephone script were tested [9]. The final semi-structured questionnaire used for data collection is shown in the Appendix. The communication scientist was present during the interviews.

All 82 patients who participated in the study were operated on with an innovative surgical technique consisting of construction of coronary anastomoses on the beating heart, avoiding use of the heart-lung machine (OPCAB: off-pump coronary artery bypass grafting) at the Center for Less Invasive and Robotic Heart Surgery in Buffalo, NY. Average time from surgery to questionnaire completion was 45 days.

In the first part of the interview, patients were asked about their general health condition. The surgeon provided professional advice while answering all health-related questions that the patients asked during the interview.

In the second part of the interview, questions were asked to investigate which sources (doctor, Internet, magazines, television, radio, family members) patients accessed to retrieve medical information. Specific questions were asked to investigate the use of the Internet before and after the operation. Internet users were asked to characterize accessibility, quality, and readability of Internet-mediated medical information. The overall benefits obtained by retrieving on-line medical information were investigated. All the Internet users were asked for suggestions for improving existing medical Internet sites.

Statistical Analyses

Discrete and continuous variables were compared using the Pearson Chi-Square test and the Student t-test, respectively. Differences between variables were considered significant when the P value was less than .05.

Results

The sample consisted of 82 patients. Demographics are provided in Table 1. Average age was 64.7 with standard deviation of 10.08 (range, 42-86) years, with a median of 67. Fifty patients (61%) were male and 32 (39%) were female. Twenty participants (24.4%) had college education, thirty-three (40.2%) high school education, and the remaining 29 (35.4%) had primary school or secondary school education. Out of 82 participants, 35 (42.7%) were Internet users and 47 (57.3%) had never used the Internet.

Education levels of Internet users and non-users were significantly different (chi ^{2}= 29.78, with df = 2, P < .001). No other significant differences were observed in the demographics of the 2 groups (Table 1).

The most common means for acquiring medical information for patients (52/82, 63.4%) was via their family doctor, cardiologist, or cardiac surgeon (Table 2). The second most common means to acquire medical information was the Internet.

In the group of 35 Internet users, 2 subgroups were observed. Eighteen patients (51.4%) (subgroup A) had used the Internet to access medical information, and 17 (48.6%) (subgroup B) had never retrieved on-line medical information. No statistically-significant differences were observed when comparing the demographic characteristics (age, education, and gender) of these 2 subgroups (Table 3).

Table 1. Demographic Data for 82 Patients Interviewed

	All Patients Interviewed (n=82)	Group of All Non-users of Internet (n=47)	Group of All Users of Internet (n=35)	Statistical Tests and P Value
% of all patients interviewed	100%	57.3%	42.7%	
Average age, years Standard deviation	64.74 10.8 (range, 42-86)	66.40 8.88	62.51 11.24	P: NS* (t-test)
Gender: Male: % (number) Female: % (number)	61% (50) 39% (32)	63.8% (30) 36.2% (17)	57.1% (20) 42.9% (15)	χ^2= .377 df = 1 P: NS*
Education: College: % (number) High School: % (number) Primary School or Secondary School.: % (number)	24.4% (20) 40.2% (33) 35.4% (29)	10.6% (5) 29.8% (14) 59.6% (28)	42.9%(15) 54.3% (19) 2.9% (1)	$^2\chi 2$ = 29.78 df = 2 P < .001

* NS = not significant

Table 2. Sources of Medical Information for 82 Patients Interviewed

Source of Medical Information	Number of Patients Using the Source	% of Patients Interviewed Using the Source
Doctor	52	63.4
Internet	18	22.0
Magazines	16	19.5
TV	12	14.6
Family	5	6.1
None	4	4.8

Table 3. Demographic Data for 35 Patients with Internet Access

	Subgroup A: Used Internet to Retrieve Health-related Information (n=18)	Subgroup B: Did Not Use Internet to Retrieve Health-related Information (n=17)	Statistical Tests and P Value
% of patients with Internet access	51.4%	48.6%	
Average age, years Standard Deviation	60.21 2.37	64.4 9.67	P: NS* (t-test)
Gender: Male: % (number) Female: % (number)	61.1% (11) 38.9% (7)	47.1% (8/17) 52.9% (9/17)	χ^2= .238 df = 2 P: NS*
Education: College: % (number) High School: % (number) Primary/Sec School: % (number)	50% (9) 50% (9) -	35.3% (6) 58.8% (10) 5.9% (1)	χ^2= 1.625 df = 2 P: NS*

* NS = not significant

Details for subgroup A: patients who used the Internet to retrieve health-related information

Among the 18 patients who had obtained medical information from the Internet (subgroup A), the Internet was accessed before surgery in 9 cases (9/18, 50%), after surgery in 7 cases (7/18, 38.9%), and both before and after surgery in 2 cases (2/18, 11.1%).

Preoperative Internet use was mainly intended to retrieve general information about cardiovascular disease (7/18, 38.9%), general information about cardiac surgery (5/18, 27.7%), and specific information about innovative cardiac surgery procedures (eg, OPCAB) (3/18, 16.7%). Postoperative Internet use was more oriented towards general health and prevention issues (6/18, 33.3%) and towards information about cardiovascular drugs prescribed after the operation (3/18, 16.7%).

Fifty percent (9/18) of the patients involved at least 1 relative (other than a younger family member) in their search; in 3 cases (16.7%) a younger family member searched for the patient, and in 6 cases (33.3%) patients searched alone. Internet users in subgroup A, easily accessed the information using popular search-engines such as Yahoo! and AOL. In some cases (6/18, 33.3%) Internet users complained about 'overwhelming' information retrieved during the search.

Readability of the Internet-mediated medical information was found to be acceptable in only 3 cases (3/18, 16.7%). The majority (15/18, 83.3%) encountered difficulties in fully understanding the information. In spite of this readability barrier, the majority of patients in subgroup A (17/18, 94.4%) found the on-line medical information rather helpful to cope better with stress and anxiety during the preoperative and postoperative period.

The patients in subgroup A were asked for suggestions for improving the existing medical Internet sites. To increase the credibility of the medical information available on-line, all patients (18/18, 100%) indicated that they would benefit from access to Internet-sites developed by their immediate health-care provider (cardiologist or surgeon). Other suggestions are reported in Table 4. Fifteen patients agreed that medical information on the Internet should be reported using easier language in order to make it more understandable to the majority of the general public.

Table 4. Patients' Suggestions for Improving Existing Medical Internet Sites

Patients' Suggestions	Number of Patients	% of Patients*
Internet sites designed by the patient's doctor	18	100
Easier language	15	83.3
Patients' forum	3	16.7
Surgeons' results published	3	16.7
More links	2	11.1

* Percentages are calculated on a total of 18 patients that did use the Internet to retrieve health-related information.

Discussion

The Internet is a powerful tool in the hands of health-care providers and is also available to health-care users. However, many patients that could actually benefit from proper use of the Internet are unfamiliar with new technologies (eg, computer technologies) and, in the majority of cases, are not even on-line [10].

Fifty-seven percent of the interviewed patients have no Internet access. Based on age and gender no significant differences were observed comparing the groups of Internet users (35/82) and Internet non-users (47/82). However, statistically significant differences ($P < .001$) for levels of education were observed between the 2 groups. Well-educated patients use the Internet significantly more than poorly educated patients. Among Internet users, 42.9% (15/35) had a college education and more than half (54.3%, 19/35) had a high-school education. Among Internet non-users only 5 patients (10.6% 5/47) had graduated from college, while the majority had a primary school or secondary school education (59.6%, 28/47). This indicates that the more patients are educated the more they tend to be Internet users. This is consistent with the Stanford Institute for the Quantitative Study of Society study [10]. However, the level of education is not a predictor of the propensity to use the Internet for medical information retrieval: almost half (16/35, 45.7%) of the Internet users had either a high school or a college education and did not use the Internet for medical-information retrieval.

There were no significant differences in the demographic data (education, age, and gender) of the two subgroups (A and B) of Internet users. This brings up the question: why did about half (18) of the Internet users search for medical information and about half (17) did not? Behavioral and attitude differences in these 2 groups can be better explained by the social science literature [11,12]. Avoidance of using the Internet to retrieve information related to medical diseases could be interpreted, in part, as an attempt by patients to defensively prevent anxiety, since, being highly involved with a problem, the patients tend to avoid other sources of stimuli that could generate further dissonance, anxiety, and stress [13]. In this regard, direct consultation with the health-care provider may be more reassuring and less distressing. However, the Internet could be a potent means for retrieving information about conventional and innovative clinical and surgical treatments. In an original study entitled ROCEP (Rural On-line Cardiac Education Project), Scherrer et al [3] analyzed the effectiveness of the Internet, and other traditional methods of medical-education, in a group of patients waiting for cardiac surgery. The Internet-based method offered increased social support, decreased anxiety, and improved lifestyle - and facilitated positive attitudes towards the impending surgery. These findings are supported by a recent pilot study funded by NIH (National Institutes of Health) in which Leaffer et al [5] demonstrated that senior citizens can easily acquire computer and Internet skills to search for medical information on the Internet. Moreover, the elderly use this tool to assume an active role in their personal health care.

Berland et al [14] investigated accessibility, quality, and readability of Internet-based health information, and concluded that: access to health information is not efficient, coverage of key clinical information is poor and inconsistent, and high reading levels are required to understand Internet-based health information. Our findings support in part Berland et al's [14] conclusions from the patients point of view. Although the majority of patients could easily find general medical

information, only a few patients (3/18, 16.7%) were actually able to retrieve specific information concerning innovative surgical treatments. Moreover, in most cases, the readability of information was poor.

Regarding the effects of using the Internet, patients reported an overall benefit from on-line information. However, significant concerns should be raised about the quality, credibility, and origin of health-related information available on the Internet. Although the Internet is a powerful tool for improving the health-care decision-making process, users should be aware of the potential for misinformation present in unprofessional Internet sites, and should always assess the source of information provided [15]. All patients agreed that, to increase the credibility of on-line medical information, access to Internet-sites developed by their immediate health-care-provider should be available. We believe that there is a growing need for objective, reproducible, widely accepted criteria that can form the basis for regulation of publication of medical material on the Internet. Internet users should be aware of the potential for on-line misinformation and should always be able to verify the source of information. Criteria for evaluating health information on the Internet have been presented in a policy paper supported by the Health Information Technology Institute [16].

Limitations and Future Studies

This study represents the initial stage of an interdisciplinary project that includes ongoing research on larger cohorts of patients. The number of participants involved in the present study is limited because only a minority of cardiac-surgery candidates are currently treated with minimally-invasive surgical techniques (eg, OPCAB). In spite of the small size of the sample, demographic findings are consistent with the data published in the American Society of Thoracic Surgeons National Database [17].

The multidisciplinary approach of the present study has helped us to better understand the phenomenon under study and to address issues that stand in the nature of semi-structured questionnaire methodology. Under the supervision of the communication scientist, the interviewer (cardiac surgeon) avoided indirectly or directly suggesting or influencing patients' answers, thus preventing the 'Hawthorne effect' [18]. The "presence" of the interviewer (cardiac surgeon) may have increased cooperation rates and it also made it possible for respondents to get immediate clarification of health-related issues from their physician.

In conclusion, we believe that additional interdisciplinary studies should be encouraged to include not only cardiopathic patients, but also subjects affected with different pathologies, to investigate the relationship between type of disease and propensity towards Internet medical information retrieval.

Conflicts of Interest

None declared.

Appendix 1

Semi-structured questionnaire used in the telephone interview

First part:

a) Surgeon's introduction and explanation of the reason for the interview

b) Check demographic data (name, age and education, gender)

c) Check actual health status of the patient

(-- Surgeon's feedback on post-surgery issues --)

Second part:*

d) Check all sources of information used when dealing with health-related issues

e) Check if Internet user or not-user.

f) Use of the Internet for medical-related topics (operation) (Yes/No)

- data access technique (web browsers, specific URLs, etc.)

- when the search was performed (before/after the operation)

- with whom (family-members' involvement in the search)

g) Quality of the data retrieved: (record patients' words) credible, clear (language)

h) Was the information retrieved helpful? (Yes/No)

Why? (check pros and cons of on-line medical information)

i) Suggestions for improving actual medical web sites'

* = Repeat key-questions (i.e., quality of information, pros and cons, problems encountered, benefits).

References

1. D'ancona G, Karamanoukian HL, Soltoski P, Salerno TA, Bergsland J. Changing referral pattern in off-pump coronary artery bypass surgery: a strategy for improving surgical results. Heart Surg Forum 1999;2(3):246-249. [Medline: 21174418]
2. Tang LW, D'ancona G, Bergsland J, Kawaguchi A, Karamanoukian HL. Robotically assisted video-enhanced-endoscopic coronary artery bypass graft surgery. Angiology 2001 Feb;52(2):99-102. [Medline: 21118816]
3. Scherrer-bannerman A, Fofonoff D, Minshall D, Downie S, Brown M, Leslie F, et al. Web-based education and support for patients on the cardiac surgery waiting list. J Telemed Telecare 2000;6 Suppl 2(6 Suppl 2):S72-S74. [Medline: 20430809] [doi: 10.1258/1357633001935662]
4. D'Ancona G, Murero M. Is the Internet a useful tool to educate cardiac surgery patients? Heart Surg Forum [in press].
5. Leaffer T, Gonda B. The Internet: an underutilized tool in patient education. Comput Nurs 2000;18(1):47-52. [Medline: 20138771]
6. Chitwood WR. The Internet and the thoracic surgeon: a "virtual" future. Ann Thorac Surg 1996 Jun;61(6):1603-1606. [Medline: 96225226] [doi: 10.1016/0003-4975(96)00224-X]
7. ; The Pew Internet & American life Project. The online health care revolution: how the Web helps Americans take better care of themselves. 2000 Nov 26. URL: http://www.pewinternet.org/reports/reports.asp?Report=26&Section=ReportLevel2&Field=Level2ID&ID=123 [accessed 2001 Feb 26]
8. Kerlinger FN, Lee HB. Foundations of Behavioral Research. New York: Holt Rinehart and Winston; Oct 1, 1985.
9. Nunnally JC. Introduction to Psychological Measurement. New York: McGraw-Hill Companies; 1970.
10. ; Stanford Institute for the Quantitative Study of Society. Internet and Society. 1999. URL: http://www.stanford.edu/group/siqss/Press_Release [accessed 2001 Sep 28]
11. Appley MH. In: Trumbull R, editor. Dynamics of Stress: Physiological, Psychological, and Social Perspectives (Plenum Series on Stress and Coping). New York: Plenum Publishing Corporation; Dec 1, 1986.
12. Goldstein M. The relationship between coping and avoiding behavior and response to fear-arousing information. Journal of Abnormal and Social Psychology 1959 Mar;58:247-252.
13. Festinger L. Theory of Cognitive Dissonance. Stanford: Stanford University Press; Jun 1, 1957.
14. 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]
15. Winker MA, Flanagin A, Chi-lum B, White J, Andrews K, Kennett RL, et al. Guidelines for medical and health information sites on the internet: principles governing AMA web sites. American Medical Association. JAMA 2000;283(12):1600-1606 [FREE Full text] [Medline: 20197445] [doi: 10.1001/jama.283.12.1600]
16. ; Mitretek Systems, Health Information Technology Institute. Criteria for assessing the quality of health information on the Internet. 1999. URL: http://hitiweb.mitretek.org/docs/criteria.html [accessed 2001 Sep 28]
17. ; Society of Thoracic Surgeons. STS data analyses - Annual trends and summaries. 1999 Jan. URL: http://www.ctsnet.org/doc/3160 [accessed 2001 Mar 15]
18. Franke RH, Kaul JD. The Hawthorne experiments: first statistical interpretation. Am Sociol Rev 1978;43:623-643.

Abbreviations

CABG: Coronary Artery Bypass Grafting
OPCAB: Off-Pump Coronary Artery Bypass Grafting
ROCEP: Rural On-line Cardiac Education Project
NIH: National Institutes of Health

peer-reviewed by T Arvanitis, JN Browndyke; comments to author 14.09.01; revised version received 21.09.01; accepted 24.09.01; published 30.09.01.

Please cite as:
Murero M, D'Ancona G, Karamanoukian H
Use of the Internet by Patients Before and After Cardiac Surgery: An Interdisciplinary Telephone Survey
J Med Internet Res 2001;3(3):e27
URL: http://www.jmir.org/2001/3/e27/
doi:10.2196/jmir.3.3.e27
PMID:11720969

Original Paper

Using the Internet to Teach Health Informatics: A Case Study

David Parry[1], MSc; Alec Holt[2], BSc; John Gillies[3], FRACP

[1]Business Faculty, Auckland University of Technology, Auckland, New Zealand

[2]Department of Information Science, University of Otago, Dunedin, New Zealand

[3]Wellington School of Medicine, Wellington, New Zealand

Corresponding Author:
David Parry, MSc
School of information Technology (CJ)
Business Faculty
Auckland University of Technology
Private Bag 92006
Auckland 1020
New Zealand
Email: parry@aut.ac.nz

Abstract

Background: It is becoming increasingly important for health professionals to have an understanding of health informatics. Education in this area must support not only undergraduate students but also the many workers who graduated before informatics education was available in the undergraduate program. To be successful, such a program must allow currently-employed students with significant work and family commitments to enroll.

Objectives: The aim was to successfully create and teach a distance program in health informatics for the New Zealand environment.

Methods: Our students are primarily health professionals in full time employment. About 50% are doctors, about 25% nurses, and the rest include dentists, physiotherapists, and medical managers. Course material was delivered via the World Wide Web and CD-ROM. Communication between students and faculty, both synchronous and asynchronous, was carried out via the Internet.

Results: We have designed and taught a postgraduate Diploma of Health Informatics program using the Internet as a major communication medium. The course has been running since July 1998 and the first 10 students graduated in July 2000. About 45 students are currently enrolled in the course; we have had a dropout rate of 15% and a failure rate of 5%. Comparable dropout figures are hard to obtain, but a recent review has suggested that failure-to-complete rates of 30% to 33% may be expected.

Conclusions: Internet technology has provided an exciting educational challenge and opportunity. Providing a web-based health informatics course has not been without its frustrations and problems, including software compatibility issues, bandwidth limitations, and the rapid change in software and hardware. Despite these challenges, the use of Internet technology has been interesting for both staff and students, and a worthwhile alternative for delivering educational material and advice to students working from their own homes.

(J Med Internet Res 2001;3(3):e26) doi:10.2196/jmir.3.3.e26

KEYWORDS

Medical Informatics; Education, Distance; Internet; Universities; New Zealand

Introduction

Health informatics is not just information science for health professionals. One definition of health or medical informatics is that it: "Comprises the theoretical and practical aspects of information processing and communication, based on knowledge and experience derived from processes in medicine and health care." [1] It has been said that: "Computers are to health informatics, what stethoscopes are to cardiology." [2] However, computers remain an essential part of any teaching about health informatics. As Edward Shortliffe has said: "Schools need to look beyond computer literacy concerns to develop formal informatics curricula that meet the needs of future practitioners who will function as users and creators of data and information." [3] Internet technology is starting to be used in the medical-education process in Australia and New Zealand and, increasingly, information-technology skills are needed for good

medical practice. This paper addresses issues relating to the use of the Internet in teaching health informatics.

The Internet was chosen as a major medium for delivery of the course because New Zealand has a long and narrow shape and a sparse population. Although distances between centers are not great, travel in New Zealand is inconvenient, especially for the people who have been identified as likely to participate in the Diploma Course. A distance method of teaching is needed to allow more potential students to take part. As Hersh [4] states, potential students for courses like this are likely to be working, have limited time, and be unwilling to move to complete a traditional course. In particular, the authors and other members of the development team believe that the use of Internet communication would significantly increase the quality of the educational experience of the students, allowing better communication between participants and better support for students.

Methods

The Course [5] is aimed at increasing the number of people skilled in health informatics in New Zealand. Although there is significant interest in the topic, health workers in New Zealand have uneven skill levels and often work outside their own personal skill level in isolated environments. The primary goals of the course are:

- Raising health professionals' understanding of health informatics and computer technology, including: the effective use of common software, communication tools, and some of the concepts underlying the use of computers in health care.
- Providing academic recognition of informatics skills, by awarding a diploma.
- Developing a network of expertise, comprising alumni and the current students and staff, to promote informatics in the New Zealand health system.
- Reaching people working in relatively-remote environments.

The developers of the course chose to deliver the course by electronic methods, especially because these methods are flexible and the students need to receive and work on the course material at times that suit their lifestyle.

The diploma is one of a number offered by the University of Otago, the second-largest university in New Zealand. Entry is open to graduates with at least a 3-year bachelors degree and to people with equivalent education (such as nurses trained before the bachelor-of-nursing degree was introduced) who are working in the field of healthcare or have a strong interest in it. Fees vary but are about US $4000 for the whole diploma for New Zealand residents. Until now, advertising for this course has been confined to New Zealand.

Students have participated from all of the provinces of New Zealand and from overseas (Australia, Brunei, Papua New Guinea, and the United Kingdom) but all face-to-face meetings have occurred in New Zealand. Our students are primarily health professionals in full-time employment. About 50% are doctors and about 25% are nurses; the rest include dentists, physiotherapists, and medical managers. The students study outside of regular working hours. Because the need for significant travel would preclude students' participation in a traditional face-to-face course, the authors and the development team hoped the use of Internet technologies for communication would increase the participation and interest of the students in the course, as well as giving the students an insight into the use of these communication technologies, which are rapidly becoming vital tools of health professionals.

The authors and the development team believe that this form of education is likely to be widely used in the future and that the benefits of electronic teaching include:

- Demonstration of the advantages of instruction using electronic media such as: hyperlinks, multimedia, and all the tools that computer-based courses can provide.
- Potential increase in the students' confidence in their ability to learn new software packages, based on their use of the course material, which is a software package, albeit one that is designed to be easy to use.
- Advice on the most suitable software applications from a teaching staff experienced in this form of course delivery.
- A flexible working pattern for the students, with support for different learning styles [4].

The authors met weekly via Internet technologies while developing the educational material - and continue to meet weekly using these technologies. In addition, the authors have met face-to-face. The authors use Internet technologies for direct supervision of the students.

A number of communication software packages have been tested and evaluated. We expect to continue to try out new technologies as they appear, and will use newer ones when they offer superior reliability, flexibility, or both.

Overview of the Diploma of Health Informatics Course

The New Zealand Education University system includes a period of study called a "paper." Points for a paper are based on the hours of study per 15-week semester. Each point is roughly equivalent to 2 hours of study per week. For this diploma, the student must obtain 40 points to graduate. To earn 40 points, a student must complete and pass 4 papers, since each paper with a passing (above 50%) grade is worth 10 points and no points are awarded for a failed paper.

The course includes 2 compulsory introductory papers, each lasting 15 weeks. Students are sent a questionnaire before the course to assess their level of computing expertise. They are then classified into 1 of 3 groups based on level of expertise (Table 1).

Students have rarely had any programming experience and have rarely previously completed any university-level courses in computing. Even students scoring at quite high levels of expertise are interested and stimulated by the introductory paper, especially by learning the techniques required for group interaction.

Table 1. Expertise levels of students, based on skills

Expertise level	Skills
Advanced	programs from the Internet or CD-ROM
Medium	Office Professional software
Beginner	Has experience with word processing. Can use e-mail

The initial paper (701) is an introduction to the computer and to standard office-suite software. Microsoft Office Professional is used as the teaching tool, along with graphics packages such as Paint Shop Pro and HTML-development tools such as Composer or FrontPage. Although the students may have used these packages before, we expect the students to be able to gain a deeper understanding of the use of these tools and ultimately to develop simple programs using Visual Basic for Applications. The second paper (702) provides an overview of the discipline of health informatics.

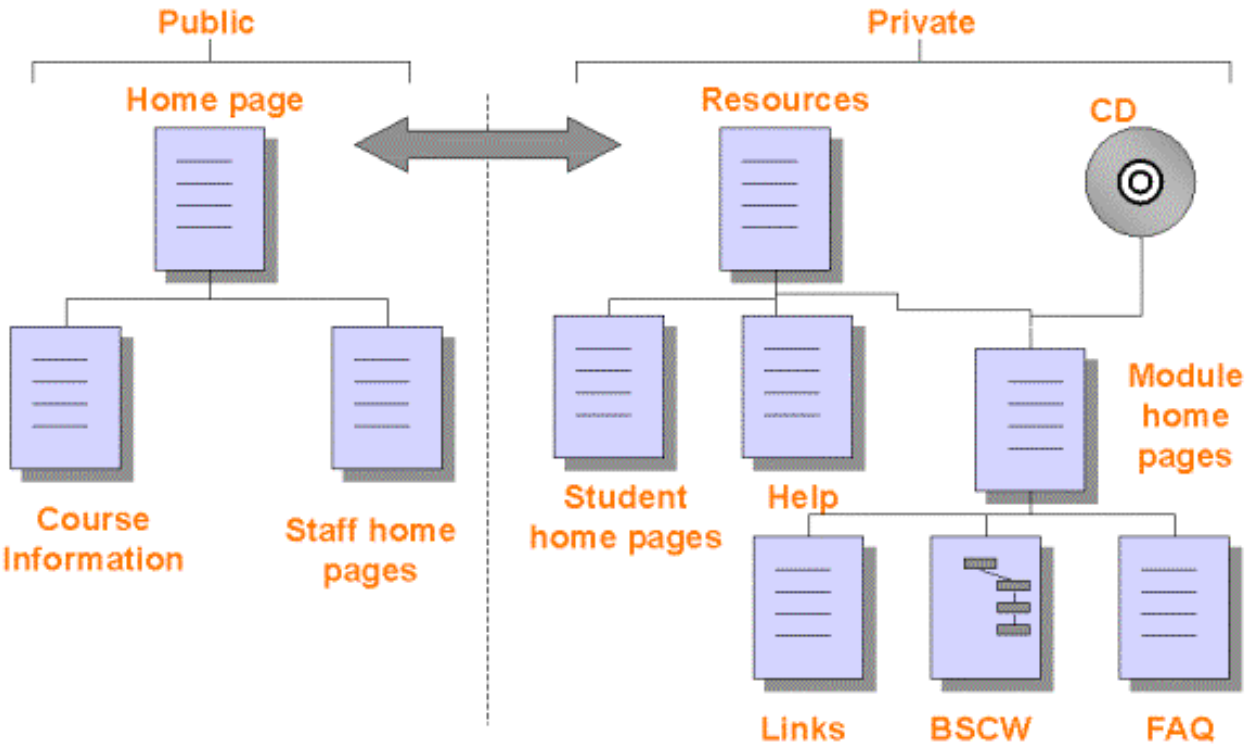
There are 5 other papers, from which the student selects 2 to complete the requirements for the diploma. These papers examine in more detail specific aspects of health informatics. The full list is:

- 701 Essential information-management skills
- 702 Principles of health informatics
- 703 Project development and software design
- 704 Evidence-based practice
- 705 Computer-aided learning
- 706 Research methods and statistics in the health sector
- 707 Personal project in the area of health informatics

For each paper, each student is given: a CD-ROM with the main course material, a list of recommended textbooks, and a reading list with between 20 and 30 journal articles or book chapters.

Other relevant material, including hyperlinks to more resources, is posted on a Web-based file-sharing system, Basic System for Cooperative Work (BSCW) [6],. BSCW provides areas for asynchronous discussion, file storage, and hyperlinks, with an extensive and powerful security model that allows rights to be assigned to individual users and groups. Asynchronous discussion is similar to discussion by e-mail except that all related postings are grouped together in "threads," as is the case in Usenet newsgroups. For example, a student may post a question, and both staff and students may reply to it. Because communication is asynchronous, these replies may be spread over a number of days. Interested people can view the progress of the discussion and contribute to it at times of their choosing. In contrast, synchronous discussion is similar to a discussion by telephone, requiring that all parties are present at the same time. Some areas of BSCW are set up to allow posting of files and hyperlinks by staff, some by students, and some by both. Details of the electronic resources, including BSCW, are shown in Figure 1. BSCW is used primarily for material and files that are updated during the semester; at the end of the semester, this material is archived and, if appropriate, included on the Web-site or CD-ROM the next time the paper is offered. The World-Wide-Web-based materials in the private area are password protected. Errata and additional materials to support the CD-ROM are also published in the private area.

Figure 1. Communication links and electronic resources for the diploma of health informatics



Details of Teaching Techniques

Workshop

The authors consider a face-to-face workshop that is held at the beginning of all papers essential to the effective functioning of the participants after they have returned to their own working environment. When participants interact from their remote sites, each participant requires a mental picture of other participants. Initially, attendance at these workshops was not compulsory, but non-attendees found difficulty in participating fully in the group discussions, so this attendance was made compulsory. For these workshops, subsidized travel and accommodation is provided for students. Until now, all workshops have been held in New Zealand, but if overseas participation increases, workshops may be held overseas in, for example, the United Kingdom.

The goals of the workshops are for the students to:

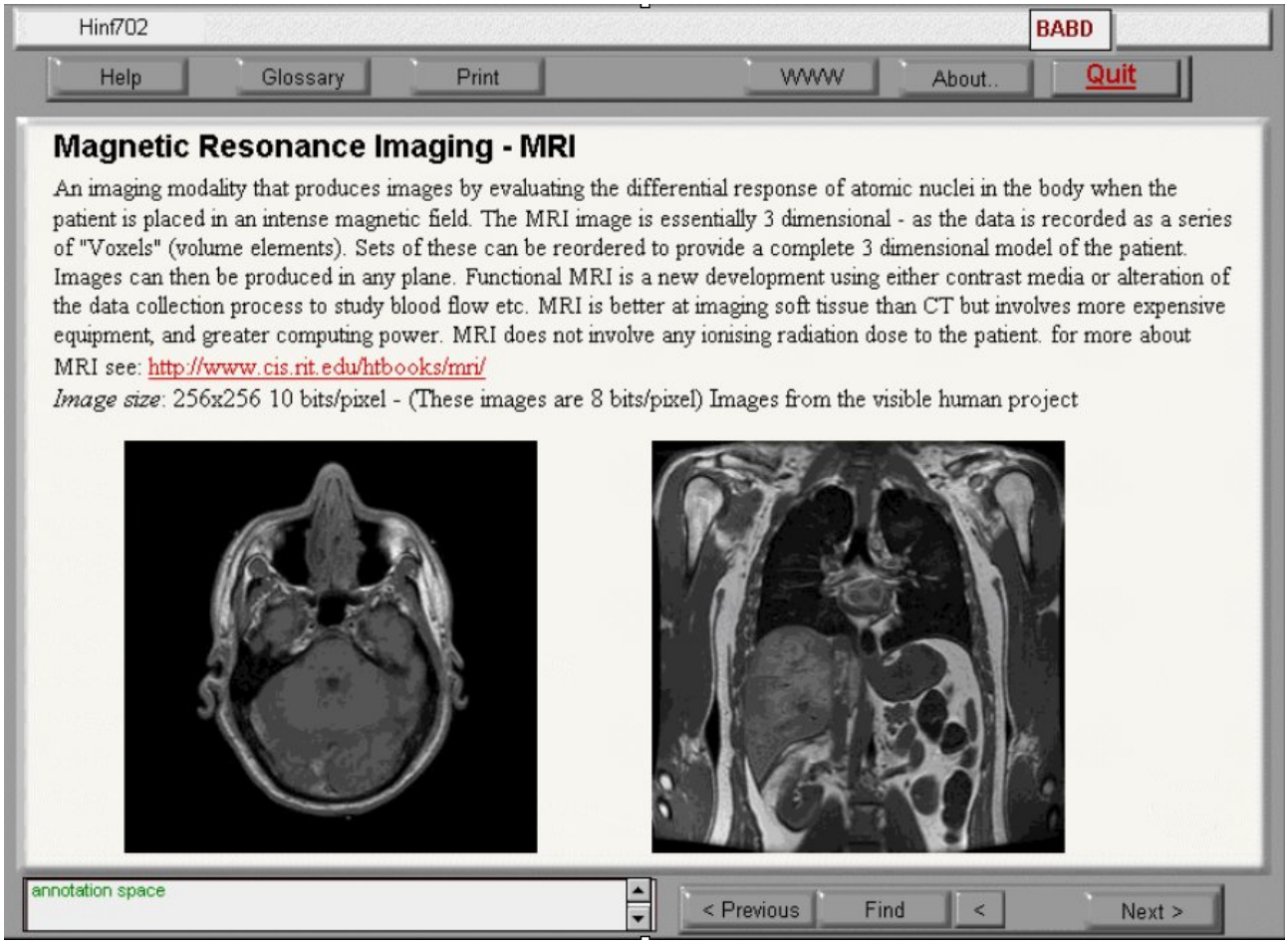
- Meet and interact with fellow participants and the teaching staff.
- Start work on their group projects (described below, in Coursework and Communication), with face-to-face meetings.
- Learn about the communications software (described below, in Coursework and Communication).
- Learn about aspects of the course and meet experts in the field.
- Have fun.

In the first workshop, the aim is to provide the student with a thorough background in use of the conferencing software and the BSCW server. Most students arrive at the workshop feeling a little unsure about their capabilities; it is vital that the students gain confidence while still at the workshop, as this significantly reduces technical problems once they have returned home. In addition to practical demonstrations and exercises on their own, we provide a printed handbook for each participant.

Course Materials

At the workshop, students are given a CD-ROM containing the course materials, along with a paper booklet of photocopied readings. The CD-ROM was constructed using Macromedia Authorware 5.1 and includes graphics, text, and videos (Figure 2). Students can save annotations to the CD-ROM on their own machines' hard drive; and a log is kept on their machine of the "pages" they have visited. The CD-ROM includes software needed for the course (eg, for zipping/unzipping [compressing/uncompressing files] and for browsing), a printable PDF (Portable Document Format) image of the course material, and extra example-application files. Each CD-ROM has about 10 modules, with a competency test for each module (see below, in Coursework and Communication). A textbook such as van Bemmel's book [1] is supplied for each paper.

Figure 2. Screen shot from a course CD-ROM



Coursework and Communication

Of the total marks assigned in each paper, 60% are given for individual tests and 40% are given for a group project.

The group project is an interesting and exciting challenge for all participants, including the staff members. In each paper, there is a group project that aims to reinforce and expand the topics discussed in that paper. Group projects have resulted in reports subsequently published in this journal [7]. In addition to the group projects students submit competency tests, for example databases or Web-pages they have developed, at set times during the semester. The group meetings (described in more detail later in this section) are also used as tutorials and discussions about the current competency test. Although this imposed timetable reduces flexibility, it allows students with problems to be identified early in the semester, before they are too far behind.

After the initial work in the face-to-face meetings at the workshop, the group meets using Internet-conferencing software. The authors and development team have tried synchronous packages and asynchronous packages, and each group decides upon the package, or combination of packages most suitable for their needs. Most choose NetMeeting and e-mail as their primary communication methods and BSCW as their location for posting documents. Where there is difficulty with the Internet-based audio-conferencing such as NetMeeting some groups have elected to use one of a variety of text-chat software options (eg, ICQ).

Internet-based audio-conferencing is presently difficult to manage, and the technology is often unreliable. We have established a set of guidelines for running Internet-based audio-conferences, to maximize the benefits of time spent on-line. Students become frustrated when they lose contact with the meeting or when they cannot hear; at present, these are recurring issues with Internet-based audio-conferences. Computer hardware is important to reliability, since some sound cards are not fully compatible with the software. The card must have full-duplex sound capability (the ability to support sound transmission and reception at the same time), to allow efficient communication. A 56K or faster modem allows much higher quality of communication.

The guidelines for running Internet-based meetings are:

- Participants must have met in person.
- Agenda is prepared and circulated in advance.
- Chairperson controls meeting and writes minutes.
- Method of voting is agreed on - eg, it is agreed that silence equals consent.
- Chairperson summarizes at the end of each agenda topic.
- Minutes are circulated the same day as the meeting.

Asynchronous communication tools.

The majority of the asynchronous work was done using e-mail which, for the University of Otago system, is limited to

messages less than or equal to 1 Megabyte (MB). Although we set up some newsgroups, the majority of document sharing and threaded discussions took place via BSCW, BSCW was developed by the German Institute for Information Technology [8] and is free for use by educational institutions. BSCW is a collection of scripts written in the Python programming language that reside on a World Wide Web (WWW) server and allow secure file storage that is accessed via Web pages. One of the most useful features of BSCW is the ability to "version" a document. This allows changes to be made, while still allowing access to the original document and intermediate versions. Access control allows the staff to limit what each user of BSCW can see - for example, discussion groups, a software archive, and the user's project area -with appropriate rights of access (for example, read, write, or delete). The teaching staff has access to all areas including the definitive course-production documents, student results, and administrative areas. A typical BSCW screen shot is shown in Figure 3.

We are running the Windows NT version of BSCW 3.2, using Microsoft's Internet Information Server 4.

Synchronous communication tools.

Table 2. Assessment of synchronous communication tools

Tool Name (Tool Company)	Tool Type	Number of Participants	Cost	Degree of Use	Comments
VoxChat (Voxware)	Audio chat and text chat; half-duplex*	Up to 5	Free	Used by all staff and students when the course was launched	longer supported
Internet Conference Professional (VocalTec)	Audio chat and text chat, whiteboard†	No limit imposed by software, but needs 4 KB/second bandwidth per audio connection to server (eg, for 10 users the server's connection to the Internet must be at least 40 KB/second of bandwidth)	About US $40 per seat‡ plus an unknown cost for server	Widely used, when first purchased; subsequent use forestalled by licensing problems	Easy-to-use and reliable product let down by a complicated mode for connecting to a meeting; no longer supported
ConferenceRoom (WebMaster)	Text chat	Up to 200	US $100	Used as a backup system if other systems not available	A simple-to-use, Web-based text chat system, with few features
NetMeeting (Microsoft) and OnLive server (White Pine)	Text chat and audio chat, whiteboard†, application sharing	5 free; up to 25 with licensed server	OnLive (the server): US $2,500 for 25 seats‡ for educational use;NetMeeting (the client): free	Used by all groups since 1999	OnLive extends NetMeeting by allowing audio chat between more than 2 participants
ICQ (Miribalis)	Text chat	No limit	Free	Used by 1 group	Easy-to-use and well-designed

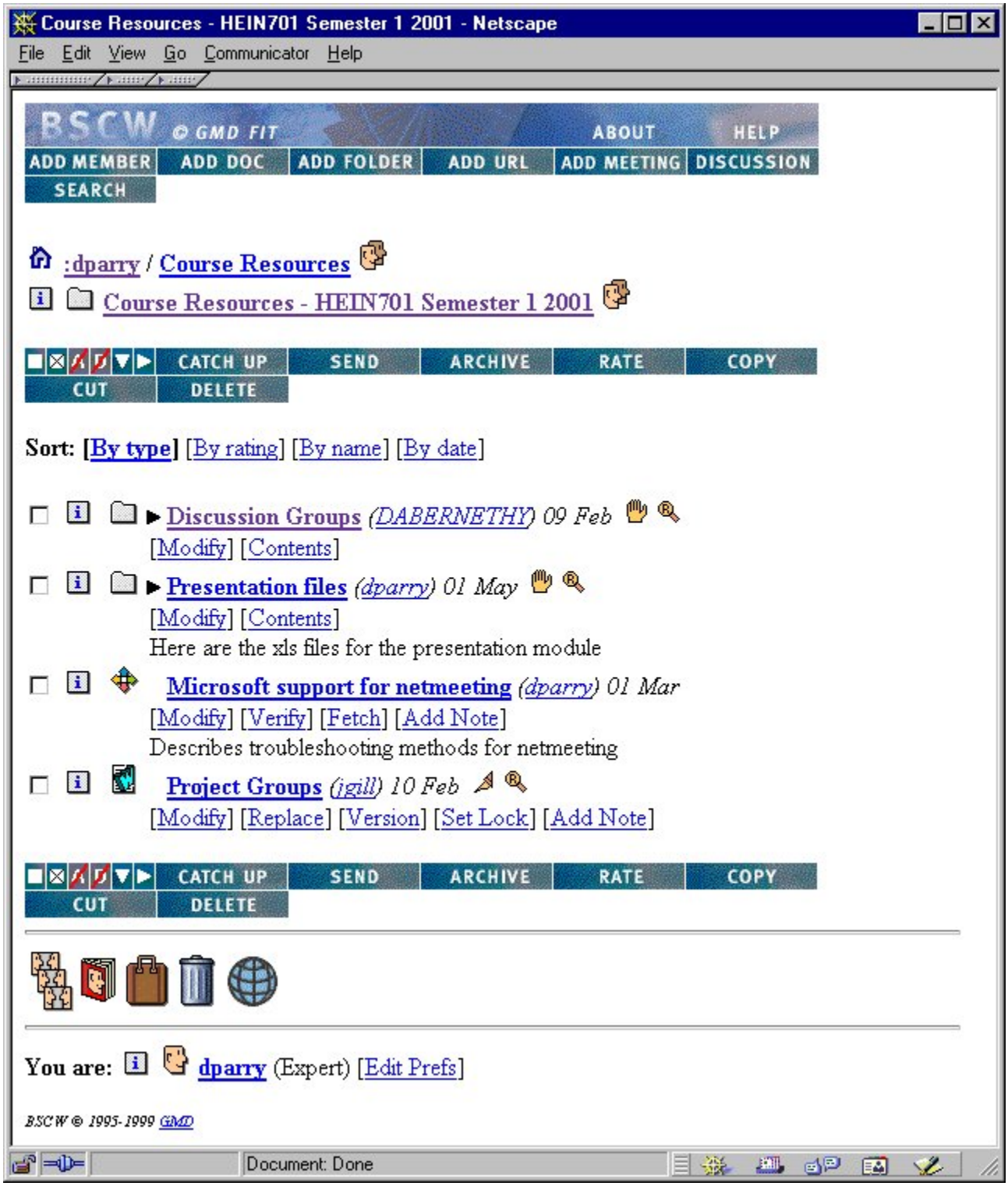
* Half-duplex: user has to wait until sound reception stops before user can transmit sound.

† Whiteboard: a common graphical area that can be seen and edited by all participants.

‡ Each computer requires a seat

We assessed some synchronous communication tools (Table 2), and have a policy of exploring new technologies as they become available. We appraised each system based primarily on cost, number of participants, and reliability. Currently we primarily use the NetMeeting/OnLive combination.

Figure 3. BSCW screen shot



Results

We have designed and taught a Postgraduate Diploma of Health Informatics program using the Internet as a major communication medium [5]. The first 10 students graduated from the course in July 2000. About 45 students are currently (July 2001) enrolled in the course because we have had a dropout rate of 15%, and a failure rate of 5%. Comparable dropout figures are hard to obtain, but a recent review has suggested that failure-to-complete rates of 30-33% may be expected [9]. Dropouts have tended to occur either just after the workshop, when fees can be refunded, or due to outside events, such as change in employment or family circumstances. Gender distributions and percentages of students qualified as a doctor are shown in Table 3. (Date-of-birth data is not presented because it is not recorded in our system.)

Student evaluation of the course, using the standard University of Otago evaluation system (a 5-point scale; 1 is worst, 5 is best) has averaged above 4 for satisfaction; the University-wide average is about 3.5. Unfortunately, this evaluation system does

not allow comparison with other papers offered by the University.

We hope to make the papers available to a wider audience, both as a diploma and as a series of short courses. Initially this is being pursued in collaboration with the University of Tasmania and the University of Auckland. In addition, the 701 and 702 papers form the basis for some undergraduate teaching at the Wellington School of Medicine. Four of the diploma graduates are currently studying towards Masters degrees in health informatics and 1 is continuing on to a PhD. Some publications, for example [7], have been produced by students and staff as part of the course.

Table 3. Student Gender and Professional Background, up to June 2001

Paper	Cumulative Number of Students	Female%	Male%	Medical Degree%
701 (Introduction; compulsory)	65	43.1	56.9	43.1
702 (Overview; compulsory)	54	40.7	59.3	44.4
703 (Database Systems)	14	28.6	71.4	64.3
704 (Evidence Based Practice)	12	25.0	75.0	66.7
705 (Computer Aided Learning)	10	70.0	30.0	30.0
706 (Research Methods)	14	64.3	35.7	21.4
707 (Research Project)	6	50.0	50.0	66.7
All papers	**175**	**43.4**	**56.6**	**45.1**

Technical difficulties were encountered in 3 main areas:

1. **Software compatibility.** Although the course was based on the Windows platform, many of the synchronous communication tools had problems with different hardware and driver configurations. This was a particular problem with sound cards. Because the students used home PCs there was little control of the hardware setup or of other software that could cause potential problems. Even use of a standard office suite caused problems with different installations, especially when the software was pre-installed.
2. **Bandwidth limitations.** At the beginning of the course, a 33K modem was the standard for home use. Although 56K modems are now commonplace, many areas of New Zealand have difficulty obtaining connections at data transfer rates above 4 KB/second, because of telephone-line quality. In addition, the university network is chronically congested despite regular upgrades. This has often limited use of the audio tools for discussion. Files larger than 100 KB are routinely zipped to reduce bandwidth overheads.
3. **Software and Hardware Obsolescence.** This is a major issue due to the current lifetime of software. Since the course has started there have been 2 generations of Windows operating systems and 3 generations of Microsoft Office suites. Although changes from generation to generation are often cosmetic, changes in, for example, screen appearance, menu location, or menu labels impact strongly on the usefulness of videos or labeled diagrams that demonstrate particular actions needed to perform a task. File formats have also changed; this presents particular problems for Microsoft Access, as the earlier versions cannot read later files. Generally, older versions of software are not available after new versions are released, so the approach of sticking with an old version for a long time is not feasible.

These technical difficulties cause greater problems when students first begin work for the diploma than later on. As the students become more confident and skilled, they can overcome these difficulties, and indeed see this as one of their greatest achievements.

Discussion

Teaching an electronic course requires a shift in attitude from teaching in a face-to-face environment. The majority of the interaction with the students is done via text-based methods, so the communication style is different from that of a classroom situation. Care has to be taken to avoid confusion, and feedback from the students has to be encouraged to confirm understanding. The delay in feedback with asynchronous communication can cause frustration and in some cases irritation. The deadlines for competency tests are on a Wednesday, so that students with difficulties working over the weekend can contact staff and receive a reply before the assignment is due.

We have used a combination of different packages to teach our course. It seems that at present this approach is superior in cost and flexibility to an approach based on a single solution [10]. Other courses in health and medical informatics (for example, those from Monash University in Melbourne, Australia [11] and Oregon Health Sciences University, USA [12]) have used techniques similar to ours although as graduate certificates they require less study than a diploma. The recent announcement that MIT is to make large amounts of material freely available on the WWW [13] suggests to us that course materials are not the only part of a course that has value to students. The

announcement implies confidence on MIT's behalf that students will continue to enroll in the courses offered, and pay the substantial fees, for the numerous benefits of participation in a course as compared with solo study.

The overwhelming impression we have gained from our students is the enthusiasm they have for this subject and the method of teaching. Despite many technical difficulties, this teaching approach seems to succeed in inspiring the participating health professionals. As Internet methods become more reliable and potential students more accustomed to using them, these techniques will be available for wider continuing professional education and even for undergraduate training. With the increase in biomedical research and information, the demand for "just-in-time" learning is increasing, and Internet-based learning is one of the few practical methods of supplying this to professionals that do not have easy access to face-to-face instruction.

In the future, we hope to:

- Shift more course material from the CD-ROM to the Internet, possibly by using distributed databases to store the material, both on our server and on the student's hard drive.
- Apply these techniques to other subjects.
- Expand the course, through alliances to other countries.
- Use intelligent information processing to support students better - especially in terms of setting up the communication tools; for example, by automatically calculating the student's connection speed when they use the synchronous tools and by creating a database of frequently asked questions that can be searched using natural language queries.

For those wishing to set up a course on similar lines there is a wide range of sources of information; an excellent book is "Learning Networks: A Field Guide to Teaching and Learning Online" [14]. The Association for the Advancement of Computing in Education (AACE) [15] publishes a number of useful journals. The journal Academic Medicine [16] produces a yearly list of advanced projects using new technologies.

Acknowledgments

The authors would like to acknowledge the enormous work of previous members of the development team including David Abernethy, Richard Pascoe, Sophie Cockcroft, and Alice Breton. The patience, good humor and enthusiasm of the students are also acknowledged. Funding has come from The New Zealand Ministry of Health; Capital Coast Health, Wellington; and the University of Otago.

Conflicts of Interest

All the authors are involved in the teaching and production of the postgraduate diploma.

Multimedia Appendix

Downloadable slides [PowerPoint ppt file, 821 kB - jmir_v3i3e26_app1.ppt]

Multimedia Appendix

Video (large file!): Use of BSCW [AVI file, 28 MB - jmir_v3i3e26_app2.avi]

References

1. Van Bemmel J, Musan M. The Handbook of Medical Informatics: Springer; 1997.
2. Coiera E. Guide to Medical Informatics, the Internet and Telemedicine. London: Arnold Publication; Jan 15, 1997.
3. Shortliffe E. Medical informatics meets medical education. JAMA 1995;273(13):1064-1065. [doi: 10.1001/jama.273.13.1061]
4. Hersh W. Distance learning techniques for medical informatics. Proc AMIA Symp 2000:1170 [FREE Full text] [Medline: 0011080024]
5. ; University of Otago. Postgraduate Diploma in Health Informatics. 2001 Apr 6. URL: http://hein.otago.ac.nz/ [accessed 2001 Apr 10]
6. ; OrbiTeam Software GmbH. Download BSCW software. URL: http://bscw.gmd.de/Download.html [accessed 2001 Sep 24]
7. Campbell Brebner RJ, Jayne K, Wendy M, Graham P, Alec H. Electronic medical consultation: A New Zealand perspective. J Med Internet Res 2001;3(1):e13 [FREE Full text]
8. Bentley R, Appelt W, Busbach U, et al. Basic Support for Cooperative Work on the World Wide Web. Int J Hum Comput Stud 1997;46(6):827-846. [doi: 10.1006/ijhc.1996.0108]

9. Phipps R, Merisotis J. What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education. The Institute for Higher Education Policy. 1999. URL: http://www.ihep.com/Publications.php?parm=Pubs/Abstract?9 [accessed 2001 Sep 26]
10. Downes S. The Future of Online Learning. 1998. URL: http://www.atl.ualberta.ca/downes/future/ [accessed 2001 Apr 10]
11. ; Monash University. Graduate Certificate in Health Informatics. 2000. URL: http://www.monash.edu.au/pubs/handbooks/distance/de0122.htm [accessed 2001 Apr 10]
12. ; Oregon Health & Science University. Distance Learning in Medical Informatics. 2001. URL: http://www.ohsu.edu/bicc-informatics/distance/ [accessed 2001 Apr 10]
13. ; Massachusetts Institute of Technology. MIT to make nearly all course materials available free on the World Wide Web. 2001. URL: http://web.mit.edu/newsoffice/nr/2001/ocw.html [accessed 2001 Apr 10]
14. Harasim L, Hiltz SR, Teles L, Turoff M. Learning Networks: A Field Guide to Teaching and Learning On-Line. Cambridge, Mass: The MIT Press; Sep 5, 1995.
15. ; Association for the Advancement of Computing in Education. Home page. URL: http://www.aace.org/ [accessed 2001 Sep 26]
16. ; Academic Medicine. Home page. URL: http://www.academicmedicine.org/ [accessed 2001 Sep 26]

Abbreviations

AACE: Association for the Advancement of Computing in Education
BSCW: Basic System for Cooperative Work
ICQ: I Seek You
KB: Kilobyte
MB: Megabyte
PDF: Portable Document Format
WWW: World Wide Web

peer-reviewed by C Boyer, N Olivieri; comments to author 09.02.00; revised version received 10.04.01; accepted 29.08.01; published 29.09.01.

Please cite as:
Parry D, Holt A, Gillies J
Using the Internet to Teach Health Informatics: A Case Study
J Med Internet Res 2001;3(3):e26
URL: http://www.jmir.org/2001/3/e26/
doi:10.2196/jmir.3.3.e26
PMID:11720968

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