Using the World Wide Web to Answer Clinical Questions

How Efficient Are Different Methods of Information Retrieval?

Mark A. Graber, MD; George R. Bergus, MD; and Christopher York
Iowa City, Iowa

BACKGROUND. The World Wide Web (Web) has the potential to revolutionize information retrieval in medicine. However, the best method of information retrieval from the Web is not known. The purpose of our study was to compare medical search engines, general-purpose search engines, medical meta-lists, and commercial sites on the Web with regard to their efficiency in retrieving medical information.

METHODS. Ten questions were identified from a database of questions posed by primary care clinicians. Authoritative answers were identified. Searches were performed using 1 commercial site, 4 general search engines, 9 medicine-specific search engines, and 2 medical meta-lists. The main outcome measures were the number of questions answered by each Web site, the correctness of the answers, the number of links followed to get an answer, and how well documented the answer was using the Health on the Net criteria.

RESULTS. MD Consult, a commercial site, answered 6 of 10 questions. Hardin MD (a meta-list) and Excite and HotBot (general search engines) each answered 5 questions. The medicine-specific search engines performed poorly, answering an average of only 1 question. MD Consult and HotBot required the least number of links to find an answer. MD Consult and Hardin MD had the best documented answers.

CONCLUSIONS. Medicine-specific search engines on the Web fare poorly in answering clinical questions when compared with general search engines. MD Consult, Excite, HotBot, and Hardin MD found the greatest number of answers.

KEY WORDS. World Wide Web [non-MESH]; information retrieval; computer communication networks. *(J Fam Pract 1999; 48:520-524)*

Physicians' information needs are well documented. More than 50% of patient visits in a nonacademic primary care setting generated a patient care question; only 33% of these were pursued and answered.[1] Other studies have shown a frequent need for information, and the anticipated ease of finding an answer predicted whether the question was pursued.[2-6]

The ideal source of information is rapidly accessible, broad enough to answer most questions, easy to use, and authoritative. The World Wide Web (Web) has the potential to fulfill all of these requirements. However, the current nature of the Web is limiting its use as a source of information. Hypertext markup language (HTML), used to create Web pages, was designed with the visual presentation of material in mind rather than to fit the needs of information retrieval.

Several information retrieval options are available to the physician, including medical meta-lists, medicine-specific search engines, general search engines, and commercial sites.

Medical meta-lists contain links to other sites on the Web. Generally, the meta-list site's coordinator (called the Webmaster) acts as a clearinghouse, collecting links to other sites of interest. A meta-list is analogous to a card catalogue in the library; it points the user toward the right place to find information, rather than containing information itself. Meta-lists are created by a person; an automated process is generally not involved. The primary advantage of a meta-list is that a person who has an interest in providing the user with salient information has reviewed each link. Thus, sites of questionable authorship and reliability should have been eliminated from consideration. The weakness of meta-lists is that they may reflect the webmaster's bias. Also, meta-lists may not be inclusive or up to date, since they are dependent on the dedication of the webmaster. Examples of meta-lists include Yahoo! (www.yahoo.com/Health) and the Hardin Library site (www.lib.uiowa.edu/hardin/md).

Search engines can also be used for information retrieval. There are 2 types of search engines that may be of interest to the physician: medicine-specific and general. Some medicine-specific search engines search sites for selected vocabularies. For example, Health on the Net's (HON) MedHunt uses a robot, Marvin (Multi-Agent Retrieval Vagabond on Information Networks), that searches the Web for any of 12,000 terms in its medical lexicon.[7] The strength of this approach is that the search engine will filter out any sites that are not (according to its programmed criteria) medical sites. Other medical search engines function as meta-lists and compile a list of sites that have been visited and

Submitted, revised, May 20, 1999.
From the departments of family medicine (M.A.G., G.R.B., C.Y.) and surgery (M.A.G.), University of Iowa College of Medicine, Iowa City. All requests for reprints should be addressed to Mark A. Graber, MD, Department of Family Medicine, 200 Hawkins Drive, 01290-D PFP, Iowa City, IA 52242-1097.

approved by human operators. The meta-list site will then search the compiled sites to return information.

General search engines are designed to search the whole Web using an automated site selection process. This allows the inclusion of a larger range of pages than a site dependent on human input. The available general search engines include HotBot (www.hotbot.com) and Excite (www.excite.com). The downside for the physician is that general search engines search sites that are geared toward the general public and also search commercial sites that may contain information biased toward a particular product.

Online textbooks and full-text journals are another category of information retrieval. Many of these are available through a commercial site that the physician must pay a fee to use. This information has gone through the same peer-review process as its hard copy version. Commercial sites are sponsored by a publisher or organization and charge for information retrieval. MD-Consult (www.mdconsult.com) is an example of a commercial site.

The requirements for finding medical information are stringent. If the Web is to be useful as a source of information at the point of care, the information must be easily and quickly available, authoritative, and current. The reliability of information on the Web has been addressed elsewhere,[8,9] and instruments proposed for evaluating this information need further evaluation in terms of their efficacy and validity.[10] Although much has been written about using the Web for medical information retrieval,[11-14] limited information on the Web's ability to answer clinical questions has been generated.[15] The purpose of this study is to examine the suitability of the Web as a clinical information source and to compare its various information gateways with regard to their ability to answer questions.

METHODS

Medical meta-lists, general purpose search engines, and medical search engines were identified by colleague referral, advertisement, reference in publications, and by searching the Web. MEDLINE searches were excluded from the study unless the full text of the article was available online.

Ten questions were drawn from a database of more than 300 questions posed by the faculty and residents of a family medicine residency to specialists by E-mail.[16] Thus, the questions were representative of those that primary care clinicians encounter in the course of clinical work that have answers not easily obtainable in the clinic library. Several selection criteria were used to limit the number of questions to 10. Questions that required judgment or were a matter of opinion were eliminated from consideration; only questions with a factual answer were included. These questions were then stratified by the number of times they were asked. The 10 questions that were asked the most often were included. In the case of a tie in the number of times a question was asked, inclusion was based on our judgment of applicability to a medical practice. When several similar questions were asked, we phrased the final question to reflect the intent of all the questions. All questions were selected before the Web searches began. The questions included in the study are listed in Table 1.

For each question, a well-documented answer was derived from an authoritative source, such as an academic publication, a practice guideline, or a published consensus opinion. We first looked for evidence-based answers for each question. Other acceptable sources were defined before the searches. Examples of acceptable sources included the Centers for Disease Control and Prevention, the National Institutes of Health, and the consensus opinions of professional organizations, such as the American Heart Association. We also accepted independent research published in peer-reviewed journals and made an attempt to find the most recent publications that supplied an answer. In the case of studies, we judged the quality of the study before accepting it as a source for an answer. Finally, editorials in peer-reviewed journals were accepted if they were sufficiently referenced. Using only systematic evidence-based sources such as the Cochrane Collaboration would not have been practical, since the answers to many of the questions have not been addressed by systematic reviews.

Search terms were standardized ahead of time and represented a consensus opinion on how we would approach the search. Any bias introduced into the study by this method would be toward finding the information on the Web, because we are experienced with Web-based searching.

The same strategy was used with each search engine, and the search engines' default search mode was used. The first 5 Web pages identified by each search engine were examined, and up to 5 links were followed from each of these Web pages in an attempt to find an answer. It can be argued that this methodology is artificial and does not reflect the way an experienced user would use the Web. We chose this methodology because it allowed us to evaluate the efficacy of the search engine as opposed to the efficacy of the searcher. For sites that represented meta-lists where a search capacity was not available, up to 10 links were followed.

We limited the number of links followed to measure the efficiency of Web-based information retrieval and simulate the time pressures experienced by the practicing physician. Because the speed of Internet connections varies by site and connection type, the number of links followed was used as a surrogate marker for the time needed to find an answer. Once an answer was found, no further hits or links were followed.

All searches were performed in English and searches were limited to sites written in English. Any Web sites returned in languages other than English were not considered for our study.

TABLE 1

Questions Included in the Study

Question: What are the current recommendations for administration of the pneumococcal vaccine?
Terms in Search: Pneumococcal vaccine, schedule, recommendation
Source of Answer: MMWR 1997; 46:RR-8

Question: What is the risk of transmission of hepatitis C by sexual intercourse?
Terms: Hepatitis C, transmission, sexual
Source: McCashland TM. Am J Gastroenterol 1996; 91:2069-70

Question: What are the current recommendations for screening and treatment of group B strep in women and their infants?
Terms: Group B streptococcus, pregnancy, screening
Source: MMWR 1996; 45:RR-7

Question: What are the diagnostic criteria for hemochromatosis, and how is it treated?
Terms: Hemochromatosis, criteria, diagnosis
Source: americanhs.org/ah00008.html

Question: What are the recommendations for screening of patients with DES exposure?
Terms: DES, (Pap smear), prevention
Source: www.dcpc.nci.nih.gov/pceb/pubs/DES_Pubs/DES_Daughters/pelvicexam.html

Question: What is the dosing regimen for itraconazole for tinea capitis?
Terms: Tinea, capitis, itraconazole
Source: Elewski B. Dermatol Clin 1996; 14:23-31

Question: What is the possibility of a child of a patient with Parkinson's disease developing the same problem?
Terms: Parkinson's disease, genetics, heredity
Source: home.mdconsult.com/das/journal/view/N/821064?ja=60322&PAGE=1.html&sid=879256&source=HS

Question: When do the fontanelles close?
Terms: Pediatric, fontanelle, (closure age)
Source: www.vh.org/Providers/ClinRef/FP Handbook/Chapter 10/11-10.html

Question: What is appropriate malaria prophylaxis in a 17-month-old child?
Terms: Pediatrics, malaria, prophylaxis
Source: jupiter.who.ch/yellow/table.3.htm#mef

Question: What level of PSA is a significant elevation in a 45-, 55-, 65-, and 75-year-old person?
Terms: PSA, elevation, age
Source: Richardson TD. Urol Clin North Am 1997; 24:339-51

Each site was evaluated using several criteria including: (1) whether the site professed to offer an answer to the question asked; (2) the number of Web page links that had to be followed to retrieve the answer; and (3) the accuracy of the information (using explicit predefined criteria). One researcher used the predefined criteria to judge the accuracy of the information. Sites that contained an answer to the question were compared with the HON criteria (Table 2), which are voluntary guidelines aimed at insuring the reliability and objective nature of health information on the Web. That a site meets these criteria is not a guarantee that the information on the site is correct. However, the criteria are provided as a way to increase the accountability of the site owner. And, several of the criteria reflect information that the physician will need to judge the quality of the information. These include the source of the information, whether any commercial financing is involved, whether any of the information is promotional material, and how to find additional information on the topic.

One of the authors (M.G.) was able to find an answer to all of the questions on the Web using various search strategies, although the searches required more than 3 hours.

RESULTS

One commercial site, 4 general search engines, 9 medicine-specific search engines, and 2 medical meta-lists were included in our study. The commercial site, MD Consult, was the top performer with regard to questions answered. This publisher-sponsored site retrieved answers to 6 of the 10 questions. Of the other top performers, 2 were general purpose search engines (HotBot and Excite) and 1 was a medical meta-list (Hardin MD). Each of these sites found answers to 5 questions. None of the medicine-specific search engines performed as well. With the exception of Medical World Search, which found 4 of 10 answers, and HON, which found 3 of 10 answers, the medicine-specific search engines rarely found the requested information. Of the remaining med-

TABLE 2

Summary of the Health on the Net (HON) Criteria

- All information is provided by medically or health-trained individuals unless otherwise specified.
- The information is not provided to replace a visit between a patient and a health care practitioner.
- If applicable, patient confidentiality and data will be maintained.
- If applicable, references to the source data and links will be available on the site.
- Any claims made about the performance or benefits of a product will be balanced and referenced.
- Contact addresses will be available for visitors trying to find additional information, and the Webmaster's address will be prominently displayed.
- Commercial and noncommercial organizations that have contributed funding or materials for the site will be clearly identified.
- Advertising or promotional materials will be clearly identified as such.

Note: Full text available at www.hon.ch/HONcode/Conduct.html.

TABLE 3

Numbers of Questions Answered

Site	Questions Answered
MD Consult (www.mdconsult.com)	6
HotBot (www.hotbot.com)	5
Excite (www.excite.com)	5
Hardin MD (Hardin Meta-Directory) (www.arcade.uiowa.edu/hardin/md/)	5
Medical World Search (www.mwsearch.com)	4
Alta vista (www.altavista.digital.com)	3
HON (www.hon.ch/cgi-bin/find)	3
Yahoo! (www.yahoo.com/Health)	2
Medscape (www.medscape.com)	2
Webcrawler (www.webcrawler.com)	1
Achoo (www.achoo.com)	1
Web doctor (www.gretmar.com/webdoctor/home.html)	1
Medical Matrix (www.medmatrix.org/index.asp)	0*
Medguide (www.medguide.net)	0
Sixsenses (www.sixsenses.com)	0
Medweb (www.cc.emory.edu/WHSCL/medweb.html)	0
Sleuth (www.isleuth.com/medi.html)	0
MD Gateway (www.mdgateway.com	†
Medaccess (www.medaccess.com)	‡

HON denotes Health on the Net
*Using search engine.
†Unable to connect to site.
‡Unable to connect to site using multiple alternative addresses (www.medaccess.net and www1.medaccess.com).

icine-specific search engines, 1 answered 1 question and the other 5 were unable to answer any of the 10 questions. We could not connect to 2 of the sites despite multiple attempts using several alternative uniform resource locators (URLs).

Table 3 contains the results from all sites. Table 4 provides additional data about the 4 sites that answered the most questions. All of the answers found were correct. With regard to efficiency, HotBot required following the fewest links for an answer (average 2.6); MD Consult was almost as efficient, requiring an average of 2.8 links. Hardin MD was the least efficient of the top 4 sites, requiring an average of 5 links to find an answer.

Both MD Consult and Hardin MD scored an average of 8 on the HON criteria. Excite and HotBot did less well with documenting the source of information. This is inherent in the nature of their design: General search engines do not limit returned sites to those of professional organizations, journals, and so forth.

TABLE 4

Summary of Top Retrieval Methods

Engine	Number of Answers	Average Number of Web Pages Viewed Before Finding an Answer	Average HON Criteria Met
MD Consult	6	2.8	8
HotBot	5	2.6	6.8
Excite	5	3.8	7.4
Hardin MD	5	5.0	8

HON denotes Health on the Net.

DISCUSSION

The Web is beginning to revolutionize information retrieval in medicine. However, it remains a difficult tool to use. Finding an authoritative answer to all of the 10 questions in our study on the Web was time consuming and challenging. One might argue that the difficulty arose because the Web was not designed as a tool for medical use. However, the medicine-specific search engines were especially poor performers. This does not mean that these sites are not useful for providing continuing medical education, keeping the user apprised of new developments, facilitating discussions, and so forth, but at this time their role in answering clinical questions is limited. The next-generation Web language, eXtensible Markup Language (XML), will likely remedy this situation to some degree. XML allows the use of meta-tags that describe the information in a document to a search engine, structuring it for retrieval.[17] These descriptions will make information searching and retrieval more efficient. Another development that should aid in information retrieval is the development of medical vocabularies, such as the Uniform Medical Language System (UMLS).[18,19]

Although not formally studied, a difference between using search engines and meta-lists became clear. The use of a meta-list requires medical knowledge and an intuition that is not needed with a search engine. To find an answer with a meta-list, the searcher has to make appropriate choices. The subjectivity of this approach was demonstrated during this project. One of the researchers (C.Y.), who has only a basic knowledge of medical terminology, was far less successful at finding answers with meta-lists than a physician investigator (M.G.) (eg, Hardin site = 3 vs 5 answers). Individual variations in information retrieval using meta-lists require further study.

LIMITATIONS

There are several limitations to this study. First, the Web is not static, and changes to site organization and search engines occur frequently. Even though the study results were valid in August 1998, they may not be

predictive of the future. Two of the search engines suggested by colleagues were no longer online at the time we started the study (mdgateway and mdaccess). Second, there are other medical sites on the Web, and medical sites and general search engines continue to multiply. Our study represents a cross section of sites. And although we tried to represent the most popular sites in our study, it may not apply to all Web medical sites. Additionally, individual searchers may get better results than those obtained in our study for 2 reasons. A searcher familiar with a specific search engine or metalist may have learned strategies that maximize returns from that site. Also, an advanced searcher may change strategies when an initial search does not yield results. However, using multiple search strategies is time consuming, which may limit the usefulness of the Web at the point of care.

CONCLUSIONS

Medicine-specific search engines are not as successful in returning information as general search engines. Hardin MD and MD Consult stand out from other medicine-specific sites in the number of questions answered and the number of HON criteria met for each answer. Excite and HotBot were the top 2 general search engines. There is the potential for other information gateways to improve. However, maintaining a comprehensive source of prescreened information is time consuming, and there has to be motivation to impel online services to take the time to do this properly. Finally, even though all of the answers we found were correct, there is undoubtedly incorrect information on the Web. Using peer-reviewed materials reviewed by a commercial site is one way to avoid this pitfall.

ACKNOWLEDGMENTS

Support for this project was provided by the University of Iowa Department of Family Medicine.

REFERENCES

1. Gorman PN, Helfand M. Information seeking in primary care: how physicians and other health care practitioners choose which clinical questions to pursue and which to leave unanswered. Med Decis Making 1995; 15:13-9.
2. Covell DG, Uman GC, Manning PR. Information needs in the office practice: are they being met? Ann Intern Med 1985; 103:596-9.
3. Ely JW, Burch RJ, Vinson DC. The information needs of family physicians and other health care practitioners: case-specific clinical questions. J Fam Pract 1992; 35:265-9.
4. Dee C, Blazek R. Information needs of the rural physician: a descriptive study. Bull Med Libr Assoc 1993; 81:259-64.
5. Chambliss ML, Greensboro JC. Answering clinical questions. J Fam Pract 1996; 43:40-4.
6. Osheroff JA, Forsythe DE, Buchanan BG, et al. Physicians and other health care practitioners' information needs: analysis of questions posed during clinical teaching. Ann Intern Med 1991; 114:576-81
7. Health on the Net Foundation. Marvin Project. Available at www.hon.ch/MedHunt/Marvin.html, accessed January 7, 1999.
8. Impicciatore P, Pandolfini C, Casella N, et al. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. BMJ 1997; 314:1875-9.
9. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: caveat lector et viewor — let the reader and viewer beware. JAMA 1997; 277:1244-5.
10. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? JAMA 1998; 279:611-4.
11. Yamamoto LG. Accessing and using the Internet's World Wide Web for emergency physicians. Am J Emerg Med 1996; 14:302-8.
12. Spooner SA. The pediatric Internet. Pediatrics 1996; 98:185-92.
13. McKeown MJ. Use of the Internet for obstetricians and gynecologists. Am J Obstet Gynecol 1997; 176:271-4.
14. Sikorski R, Peters R. Internet anatomy 101: accessing information on the World Wide Web. JAMA 1997; 277:171-2.
15. Hersh WR, Gorman PN, Sacherek LS. Applicability and quality of information for answering clinical questions on the Web. JAMA 1998; 280:1307-8.
16. Bergus GA, Sinift, SD, Randall CS, et al. Use of an e-mail curbside consultation service by family physicians. J Fam Pract 1998; 47:1-4.
17. Gottesman BZ. Why XML matters. PC Magazine Online, available at www.zdnet.com/pcmag/features/xml98/intro.html, accessed January 7, 1999.
18. Suarez HH, Hao X, Chang IF. Searching for information on the Internet using the UMLS and medical world search. Proc Annu AMIA Fall Symposium 1997:4-8.
19. Humphreys BL, Lindberg DA, Schoolman HM, et al. The Unified Medical Language System: an informatics research collaboration. JAMA 1998; 5:1-11.