Comparing Web search engine performance in searching consumer health information: evaluation and recommendations*

By Gang Wu, M.S.
Information Services Librarian

Shiffman Medical Library
Wayne State University
4325 Brush Street
Detroit, Michigan 48201

Jie Li, M.L.S., AHIP
Information Services Librarian

Biomedical Library
University of South Alabama
2451 Fillingim Street
Mobile, Alabama 36617

Identifying and accessing reliable, relevant consumer health information rapidly on the Internet may challenge the health sciences librarian and layperson alike. In this study, seven search engines are compared using representative consumer health topics for their content relevancy, system features, and attributes. The paper discusses evaluation criteria; systematically compares relevant results; analyzes performance in terms of the strengths and weaknesses of the search engines; and illustrates effective search engine selection, search formulation, and strategies.

INTRODUCTION

The Web can be used as a quick and direct reference source to answer many consumer health questions regarding medical conditions, syndromes, disorders, medical news, rare diseases, health products, and drug information [1]. In addition, the Web is a source of information about health topics in the early stages of research. However, information found on the Web needs to be filtered and may include voluminous misinformation or nonrelevant information [2]. One of the best sources of valuable, reliable consumer health information is a U.S. government-sponsored, quality Web site such as healthfinder† and MEDLINE*plus*‡, or sites sponsored by academic institutions such as HealthWeb§ [3], Hardin Meta Directory**, and BioSites††. However, consumers may not be aware of these Web sites or want to distill specific information on a topic quickly, using an available search engine.

Finding useful health information quickly on the Internet can challenge both the consumer and the information professional. Though the performance of currently available search engines has been improving continuously with powerful search capabilities of various types, lack of comprehensive coverage [4], inability to predict the quality of retrieved results [5], and absence of controlled vocabularies [6] make it difficult for users to use search engines effectively. The use of the Internet as a consumer health information source needs to be carefully evaluated as no traditional quality standards or control have been applied to the Web [7]. Librarians need to be able to provide informative recommendations to their clientele regarding the selection of search engines and effective search strategies.

* Based on a presentation at the Ninety-eighth Annual Meeting of the Medical Library Association, Philadelphia, Pennsylvania, May 25, 1998.

† heathfinder, the federal gateway Web site for health information, is available at http://www.healthfinder.gov/.

‡ MEDLINE*plus*, the National Library of Medicine's Web site for consumer health information, is available at http://www.nlm.nih.gov/medlineplus/.

§ HealthWeb, the source of evaluated health information, is available at http://healthweb.org.

** Hardin Meta Directory is available at http://www.lib.uiowa.edu/hardin/md/index.html.

†† BioSites is available at http://galen.library.ucsf.edu/biosites/.

A literature review revealed a number of articles discussing search engine performance and features [8–9], search strategies [10–11], search engine precision [12], and coverage, but none attempted to compare the leading search engines regarding retrieval content and relevancy in the consumer health area. Several authors examined health information resources on the Internet [13–14]. One interesting study presented by Anderson, Allee, Chung, Westra, and Lingle [15] compared twenty-five comprehensive or authoritative health information Web sites, focusing on their selection policies, updating frequency, response time for e-mail inquiries, and features such as the use of tables, frames, Java, animated graphics interchange formats (GIFs), navigation tools, and search engines. Criteria used in their evaluation included administration and quality control, meta-content, and design. Conclusions drawn from the study enabled librarians to assist their clients in selecting the best sites for addressing specific questions and general information needs, and might assist Web designers in developing more useful health information products. However, the study did not include a comparison of leading search engines nor the reliability and relevancy of the information included within the sites.

This evaluation study of search engine performance was conducted using seven leading search engines to (1) analyze the search results, (2) identify strengths and features of each search engine, (3) discuss the variables comprising effective search strategies, and (4) rank the search engines according to their performance in searching specific consumer health topics. The comparison was designed to (1) identify which search engines are likely to yield the most useful and relevant consumer health information through the use of realistic search questions, and (2) attempt to ascertain which type of consumer health questions each search engine answers best.

METHODOLOGY

The search engines selected for comparison were Medical World Search‡‡, AltaVista§§, HotBot***, Excite†††, Infoseek‡‡‡, Northern Light§§§, and Yahoo****. Metasearch engines—which translate searches into syntax understandable by discrete search engines, conduct searches against multiple search engines simultaneously, and display a composite result (e.g., Dogpile, or MetaCrawler) [16]—were excluded from this comparison. Five topics were identified and searched using the seven search engines. The topics were selected from questions posed to librarians at two reference desks in two different academic health sciences libraries and represented categories of commonly asked questions by consumers. The topics selected focused on diseases, drugs, medical news, rare disorders, and natural products. Table 1 shows the query topics used in the study. Default setting was used for all search engines. To achieve the best results and increase comparability, the search strategies were adjusted to take advantage of each search engine's capabilities. Queries were formulated using keywords linked by Boolean operators or non-Boolean symbols (e.g., + or " ").

Table 1
Sample query topics

1. Migraine headache (disease)
2. Adverse effects of tamoxifen use in breast cancer (drug)
3. Montelukast for children with asthma (medical news)
4. Kawasaki syndrome (rare disorder)
5. Dong Quai for menopause (natural product)

A "criteria checklist" was developed to evaluate results according to these factors: relevance, source reliability, currency, and duplicate and inactive links. Relevance was defined as results that provided information covering many aspects of the topic and would be considered useful for the general public users. Nonrelevant results included results that provided bibliographic information or other potentially useful information about a topic, but were neither complete nor fully informative. A source's reliability was judged by its authorship, source of its content, disclosure, and currency. For example, government agencies, academic institutions, professional associations or research foundations, hospital systems or managed care organizations, and watchdog organizations generally provided relevant, useful, reliable, and scientifically sound health information [17]. Materials published or updated within one year were considered current. If the same Web document was retrieved more than once by a search engine within the first thirty citations, it was counted as a duplicate. Inactive links included "404 errors," where the server was contacted but the path was not found; errors stating that access to the page was forbidden; pages announcing that the desired page had moved; and "603 errors," where a site's server did not respond. The inactive link ratio was an indication of how frequently and thoroughly the engine checked the links in its database for currency.

Questions were searched by two librarians using each test search engine within a two-week timeframe in spring 1998. Because of the different features of the search engines, the searchers tailored strategies individually for each search engine. Strategies were dou-

‡‡ Medical World Search is available at http://www.mwsearch.com.
§§ AltaVista is available at http://www.altavista.com.
*** HotBot is available at http://www.hotbot.com.
††† Excite is available at http://www.excite.com.
‡‡‡ Infoseek is available at http://infoseek.go.com.
§§§ Northern Light is available at http://www.northernlight.com.
**** Yahoo is available at http://www.yahoo.com.

Table 2
Number and percentage of relevant hits retrieved per queries topic (n = 30)

	Infoseek	AltaVista	Excite	Yahoo	Northern Light	HotBot	Medical World Search
1. Migraine	22 (73%)	19 (63%)	16 (53%)	11 (37%)	12 (40%)	6 (20%)	19 (63%)
2. Tamoxifen	14 (47%)	13 (43%)	7 (23%)	2 (7%)	7 (23%)	2 (7%)	2 (7%)
3. Montelukast	12 (40%)	5 (17%)	21 (70%)	14 (47%)	4 (13%)	15 (50%)	7 (23%)
4. Kawasakai	11 (37%)	18 (60%)	15 (50%)	11 (37%)	11 (37%)	6 (20%)	9 (30%)
5. Dong Quai	16 (53%)	18 (60%)	12 (40%)	20 (68%)	14 (47%)*	9 (30%)	0 (—)

* Northern Light figures do not count the "Special Collection" documents because only bibliographic information is provided. To read full-text journal or newspaper articles, users need to order them online.

ble-checked by a third searcher to improve retrieval for each query. Because the search engines retrieved hundreds of hits (i.e., links or pages of information), the comparison of relevancy, source reliability, and duplicate and inactive links was conducted based on the first thirty results returned from queries, as 80% of users only view the first two pages of results [18]. The results were examined by the two searchers through e-mail to reach agreement on the relevancy of the materials.

RESULTS

Table 2 illustrates the percentage of relevant hits per topic. Infoseek retrieved the largest number of relevant hits for topics regarding migraine and the adverse effect of the drug tamoxifen. Excite retrieved the largest number of relevant hits for topic 3 regarding the use of montelukast to treat asthma in children. AltaVista retrieved the largest number of relevant hits for topic 4 regarding the rare Kawasaki syndrome. Yahoo produced the greatest relevant number for the natural product Dong Quai used as drug. Medical World Search retrieved no documents on using Dong Quai as a drug. AltaVista retrieved the highest number of total hits and Medical World Search the lowest number for the five topics.

Table 3 displays the total retrieval and percentage of all five queries for relevancy, reliability, duplicate links, and inactive links. Infoseek retrieved 75 relevant items of the first 150 hits, followed by AltaVista and Excite at 73 and 71, respectively. Medical World and Northern Light retrieved more reliable items in comparison with their total relevant hits. Infoseek and Excite had the lowest number of duplicate and inactive links, while HotBot had the highest number of duplicate links; Yahoo and Northern Light had the highest number of inactive links.

DISCUSSION

Search engine features

Each search engine has its own focus and features. Table 4 itemizes the features and special attributes of each search engine. To achieve better precision with search results, users need to take advantage of those special features and select different search engines according to their needs. For instance, Northern Light, with its feature of delivering full-text articles at a relatively low price in response to topic 5 regarding using Dong Quai to treat menopause, offers options to retrieve information either from the Web or from its special collection, which includes a wide spectrum of academic and popular literature articles that cannot be found in MEDLINE or the Health Reference Center consumer health database. For the layperson, although the search results may include some inactive links and out-of-date materials, Northern Light offers some

Table 3
Comparison of retrieval for the first thirty hits of all query topics (n = 150)

	Infoseek	AltaVista	Excite	Yahoo	Northern Light	Hotbot	Medical World Search
Relevance	75	73	71	58	48	38	37
Percentage	50%	48.7%	47.3%	38.7%	32%	25.3%	24.7%
Reliability	48	59	61	37	43	29	35
Percentage	32%	39.3%	40.7%	24.7%	28.7%	19.3%	23.3%
Duplicate links	4	14	5	7	6	25	7
Percentage	2.7%	9.3%	3.3%	4.7%	4%	16.7%	4.7%
Inactive links	5	11	4	21	21	8	13
Percentage	3.3%	7.3%	2.7%	14%	14%	5.3%	8.7%

Table 4
Search engine feature and attribute chart

Search engines	Features	Comments
AltaVista	Limit to particular date or language Perform proximity, Boolean, or field search Refine search result by selecting relevant topics suggested Use * for truncation and parentheses to group search expressions Default operator OR Include subject index Natural language searching Limit translation capabilities to 6 languages	Best for proximity and nesting searches Advanced search turns off automatic relevance ranking Results can be ranked by criteria user specifies for advanced search Ignore Boolean operators in the simple search Duplicate and inactive links are a problem
Excite	Concept-based search retrieves information with query words and related words Find more related information based on a particular interesting document retrieved Perform Boolean search Use parentheses to group search expressions Refine search result by select relevant topics suggested Sort the results by sites Include subject index Default operator FUZZY LOGIC (matches contain all or at least one of the search term)	No field searching Automatic stemming Use Boolean operators AND/OR/AND NOT to turn concept-based search off and find the documents containing the exact words you are looking for
HotBot	Limit to date, media type, location, or domains Retrieve all of the words, any of words, phrase, URL, person Include subject index	Offer user-friendly advanced search Good for searching multimedia files No truncation Can easily change the default search AND to OR Duplicate links are a problem
Infoseek	Limit to location or subject categories Natural language search Group related hits Field searching Default operator OR Automatic truncation Include subject index	Offer user-friendly advanced search Put the most important term first Terms in brackets are searched as phrase with no order within 100 words
Medical World Search	Perform UMLS-based search Automatically maps and explodes term Refine search options Groups relevant hits No truncation Default operator OR	Only search UMLS phrase Cannot issue Boolean operator at the initial search Not good for journal publisher Web sites or natural products using as a drug Good starting place for professional information, though data may not be updated frequently as other search engines Comparably small index database Able to send the query formulation to other search engines No advanced search option
Northern Light	Limit to language, date, source, and type of information Customize hits by categories (e.g., subject, source of information, type of information) Field searching Default operator AND Automatic search for word ending variants	Special collection offers an option to order full-text articles Offer advanced search option Inactive links are a problem
Yahoo	Limit search to subcategories Include subject index Default operator FUZZY LOGIC (matches contain all or at least one of the search term)	Offer advanced search option Can send the query formulation to other search engines if not found in Yahoo Inactive links are a problem

unique sources and attractive features such as document ordering and "customer folders," which give a complete overview of major topics, types of documents retrieved, and source of information (e.g., government, personal pages, or commercial) at a single glance. Medical World Search uses the National Library of Medicine's Uniform Medical Language System (UMLS) for term mapping and automatically explodes query terms based on the UMLS hierarchical structure; in other words, users can use medical terminology to expand or narrow their searches. This capability may be valuable to obtain comprehensive results on a given topic. For example, when users type in the term "breast cancer" in the query box of Medical World Search, the system will, in addition to finding documents mentioning breast cancer, retrieve documents indexed by potentially related and specific terms such as male breast cancer, breast cancer stage,

cellular diagnosis, ductal breast carcinoma, and mammary Paget's disease.

Database content and coverage

In addition to search engine features and attributes, their database content and coverage are other important factors to consider for search engine selection and to improve the comprehensiveness or precision of search results. For disease-related information, medical news, drug information, and rare disease syndromes, all the search engines investigated find some relevant materials, though the results may overlap to some extent. Based on this study, HotBot followed by AltaVista have larger consumer health information collections based on the total number of results returned for the sample. Medical Word Search and Infoseek are recommended for disease searches and AltaVista for drug information. Excite may be considered as a good starting place for medical news. Yahoo, AltaVista, and Northern Light are recommended for searching natural products such as drugs and alternative medicine resources. Northern Light's special collection offers bibliographic information including some unique and quality resources for topic 5 on Dong Quai for menopause. Materials retrieved for topic 5 from the Web sites investigated include extensive advertising. Therefore, users need to be especially aware of the need to evaluate Web information critically. Medical World Search, though with a comparably small Web index database, focuses on clinical medicine and provides reliable information from sources such as the University of Iowa's Virtual Hospital†††† and The University of Pennsylvania's OncoLink‡‡‡‡ Web sites. This search engine returns good results about diseases.

Query formulation

Search query formulation will directly affect search results. For complex searching, using Boolean commands to perform adjacent or nested searches is better. For example, in topic 2 on adverse effects of tamoxifen in breast cancer, using Boolean operators to combine the keywords—such as (tamoxifen AND (adverse effects OR side effects), AND (breast cancer OR breast neoplasm)—will improve the precision of the search results. For simple searching, in most cases, users may use a simple search form and put "+" in front of a word or phrase. To improve precision, users should consider conducting a phrase search using quotation marks and enter the most important word or phrase first (e.g., +"Dong Quai" +menopause). The more specific the term used, the greater the precision achieved.

Medical terminology was more likely to yield bibliographic citations in the evaluation. There was little overlap between trade and generic drug name searches. Generic drug names used as search terms appeared to produce more specific information. Other useful search tips used in the evaluation included: use synonyms and other variations for search terms; keep searches simple; use advanced or power search functions to limit a search to a particular date, language, media type, source type, or location; and try more than one search engine because no single engine indexes more than one-third of the "indexable Web" based on the Lawrence and Giles study [19].

Meta-search engines and other health-related megasites are available for users to conduct a quick search in addition to the ones compared in this study. Because all search engines complement each other and meta-searches will conduct simultaneous searches against different search engines, users may try a meta-search first for some simple queries. If satisfactory results are not produced or the search query is complicated, users can select those search engines discussed above using their powerful features. Megasites (e.g., HealthLinks§§§§, Magellan*****, OMNI†††††) focused specifically on medical topics are additional good sources for users to find health-related information.

CONCLUSION

In this study, Infoseek, AltaVista, and Excite rank as the top search engines with highest relevant percentage of returns (50%, 48.7%, and 47.3%, respectively), with overall good performance for their currency, sources of information, and advanced search features. Yahoo and Northern Light's special collection are good sources for alternative medicine and natural products. Medical World Search provides information that may be of interest to health professionals due to its UMLS thesaurus-based design, though limited coverage and inactive links are a problem. HotBot has comprehensive coverage and advanced search features. It is a good source for images and multimedia file searching, though duplicate links are a problem. This study has determined that, using the best search engines, only half of the retrieval would be relevant.

The quality of materials on the Web is clearly more variable than the quality of results retrieved through traditional database searching [20]. A large amount of materials retrieved from the Web must be examined and carefully evaluated, thus users cannot predict the quality and timeliness of search results. However, searching the Web does enable users to find extremely

†††† Virtual Hospital is available at http://www.vh.org.
‡‡‡‡ OncoLink is available at http://cancer.med.upenn.edu.

§§§§ HealthLinks is available at http://www.hslib.washington.edu.
***** Magellan's health megasite is available at http://magellan.excite.com/health.
††††† OMNI is available at http://omni.ac.uk.

current information and images, illustrations, medical conferences and products, current statistics, news, drugs, and full-text articles, and complements retrieval through traditional bibliographic databases.

ACKNOWLEDGMENTS

The authors are very grateful to Ellen Marks at the Wayne State University Shiffman Medical Library for her support of this study, her valuable comments on the manuscript, and her editorial assistance.

REFERENCES

1. DICKSTEIN R, GREENFIELD L, ROSEN J. Using the World Wide Web at the reference desk. Comput Libr 1997 Sep;17(8): 61–5.
2. STONE VL, FISHMAN DL, FRESE DB. Searching online and Web-based resources for information on natural products used as drugs. Bull Med Libr Assoc. 1998 Oct;86(4):523–7.
3. REDMAN PM, KELLY JA, ALBRIGHT ED, ANDERSON PF, MULDER C, SCHNELL EH. Common ground: the HealthWeb project as a model for Internet collaboration. Bull Med Libr Assoc 1997 Oct;85(4):325–30.
4. LAWRENCE S, GILES CL. Searching the World Wide Web. Science 1998 Apr 3;280:(5360):98–100.
5. MCCLUNG HJ, MURRAY RD, HEITLINGER LA. The Internet as a source for current patient information. Pediatrics 1998 Jun;101(6):E2.
6. NOTESS GR. Internet search techniques and strategies. Online 1997 Jun/Aug; 21:(4)63–6.
7. HEALTH SUMMIT WORKING GROUP. Criteria for assessing the quality of health information on the Internet. [Web Document]. Health Information Technology Institute, Nov 1998 [rev 5 May 1999; cited 21 Jun 1999]. <http://hitiweb.mitretek.org/docs/policy.html>
8. WOOTTON JC. Search engines. J Womens Health 1997 Jun; 6(3):345–7.
9. SU LT, CHEN HL, DONG XY. Evaluation of Web-based search engines from the end-user's perspective: a pilot study. Proc ASIS Annu Meet 1998;35:348–61.
10. DOWNES PK. Successful Web searching. Br Dent J 1998 Oct 24;185(8):393–9.
11. LITTLETON D. A review of strategies for finding health information on the World Wide Web. Med Ref Serv Q 1998 Summer;17(2):51–5.
12. CHU HT, ROSENTHAL M. Search engines for the World Wide Web: a comparative study and evaluation methodology. Proc ASIS Annu Meet 1996;33:127–35.
13. SILBERG WM, LUNDBERG GD, MUSACCHIO RA. Assessing, controlling, and assuring the quality of medical information on the Internet: caveant lector et viewer—let the reader and viewer beware. [Editorial]. JAMA 1997 Apr 16;277(15): 1244–5.
14. Medical help on the Internet. Consumer Rep 1997;62:27–31.
15. ANDERSON PF, ALLEE N, CHUNG J, WESTRA B, LINGLE V. Comparison of health information megasites. [Web Document]. Ann Arbor, MI: University of Michigan. [24 May 1998; cited 21 Jun 1999]. <http://www.lib.umich.edu/megasite/abstract.html>.
16. NOTESS GR. Internet "onesearch" with the mega search engines. Online Nov–Dec 1996;20(6):36–9.
17. How to evaluate medical information found on the Internet. [Web Document]. California Medical Association Online Network of Services (CMAnet), Apr 1998. [25 Feb 1999; cited 21 Jun 1999]. <http://www.cmanet.org/publicdoc.cfm/60/0/GENER/99>.
18. JANSEN BJ, SPINK A, BATEMAN J, SARAEVIC T. Real life information retrieval: a study of user queries on the Web. SIGIR Forum 1998;32(1):5–17.
19. LAWRENCE, op. cit.
20. STONE, op. cit.

Received March 1999; accepted June 1999