
Use of a multi-application computer workstation in a clinical setting*

By William Hersh, M.D.
Biomedical Information Communication Center

David Hickam, M.D., M.P.H.
Health Services Research and Development

Department of Medicine
Oregon Health Sciences University
Portland, Oregon 97201

The goal of this study was to assess the usage frequency, user satisfaction, and quality of literature searches for a multi-application computer workstation in a university-based general medicine clinic. A computer with medical literature searching, textbook searching, and a decision-support program was deployed in the workroom of the clinic and made available for routine use. Data were collected for ten months. More than three quarters of the study participants used the computer, with use increasing by level of medical training. Despite physicians' known preferences for nonjournal sources of information, literature searching was the application used most frequently, followed by textbooks and decision support. The literature searches were replicated by experienced clinician and librarian searchers using first full MEDLINE and then text-word-only searching, to compare the quantities of relevant references retrieved. Novice searchers retrieved a larger number of relevant references than did the experienced searchers, but they also retrieved more nonrelevant references. For both groups of experienced searchers, the full MEDLINE feature set conferred little benefit over searching with only text words. These searching results call into question the value of traditional searching methods for both novice and experienced physicians.

Despite the allure of computerized access to the medical literature, acceptance of this tool has been modest. Physicians generate many unanswered questions in the course of patient care [1-3], yet most studies of computer-based retrieval systems in clinical settings have revealed a use rate of only one to six times per month [4-6]. If computers are to play a role in meeting physicians' information needs, they must be easier to use than current systems and offer more resources beyond the usual MEDLINE database. While many multi-application workstations have been described [7-9], only one study has examined their usage specifically in a clinical setting [10]. That study was limited by a relatively short period of observation (two weeks) and did not report the usage frequency for the different applications.

Despite the infrequent usage of these systems, overall satisfaction is generally high, with users reporting that systems are easy to use and provide useful information [11-14]. One study, however, found that after three years, one third of initial "early adopters" had stopped using search systems due to the difficulty of searching, the retrieval of material that did not meet information needs, and slow system speed [15]. Another study found that instituting user fees for a previously free system cut frequency of use by two thirds [16].

The infrequent use of computerized searching and conflicting assessments of its value have led researchers to attempt to measure the quality of user-system interaction. One stumbling block has been the definition of quality. Many have defined quality searches as those employing strategies similar to expert searchers, usually a medical librarian [17-20]. However, this definition is imperfect, because many studies have

* This work was supported in part by Grant no. LM05307 from the National Library of Medicine.

shown that even expert searchers rarely find more than half of the relevant references on a given topic [21-24], due to inconsistency in the assignment of indexing terms by indexers [25], variation in the language used by writers of medical texts [26], and different strategies for searching employed by even expert searchers [27]. Therefore, a more appropriate measure of quality may be the relevance of the articles actually retrieved to the needs at hand [28-31].

In the study reported here, the authors sought to measure the usage frequency, user satisfaction, and quality of searches conducted with a multi-application computer in the workroom of a university general medicine clinic. There were three aspects to this effort. The first involved measuring the use of a conveniently located multi-application computer system over a prolonged time period. The second involved evaluating a literature searching product with a user interface widely regarded as easy to use [32-33], especially for novices. This product features a natural language interface of the type long advocated by many information science researchers [34] but never evaluated in a nonlaboratory setting. The third aspect involved assessing the value of traditional MEDLINE searching, including the use of terms from the Medical Subject Headings (MeSH) vocabulary, subheadings, and term explosions, as compared to the simpler use of text words in the title, abstract, and subject heading fields.

METHODS

The General Medicine Clinic at Oregon Health Sciences University (OHSU) is the primary outpatient practice site for thirty-two internal medicine residents and nine general medicine faculty physicians. The workroom, an area of the clinic adjacent to patient exam rooms, is used for writing notes and discussing cases. Residents each see patients one half-day per week except when on an ambulatory care block, during which they use the clinic three to six half-days per week. Faculty members each see patients anywhere from one to six half-days per week, and most faculty members supervise residents one half-day per week. A Macintosh computer was installed in the workroom with four applications:

Knowledge Finder (KF). This is a CD-ROM retrieval system featuring a primary care subset of MEDLINE covering 270 journals over five years. KF features a "free-text" searching interface; users enter queries in natural language, and references are ranked in proportion to the number and frequency of search terms they contain. This process allows the retrieval of references that do not necessarily contain all the search terms, with the relative rank being an estimate of relevance.

Stat!-Ref (SR). This is a CD-ROM collection of twelve textbooks, including *Scientific American Medicine*, *AHFS Drug Formulary*, and several volumes from the Lange series. At the time of the study, the SR system used the KF interface.

Yearbooks on Disk (YB). This is a CD-ROM containing volumes of the Yearbook series for the years 1989 to 1991, including yearbooks of *Cardiology*, *Dermatology*, *Diagnostic Radiology*, *Drug Therapy*, *Emergency Medicine*, *Family Practice*, *Medicine*, *Neurology and Neurosurgery*, *Obstetrics and Gynecology*, *Oncology*, and *Psychiatry*.

Quick Medical Reference (QMR). This is a diagnostic "decision support" system featuring profiles of more than 600 diseases or syndromes commonly seen in an internal medicine practice. Users can enter patient findings to generate differential diagnoses, rule out diagnoses, and obtain critiques of user-suggested diagnoses [35-36].

OHSU users were provided minimal system training, as little as ten or fifteen minutes, although they were offered more if they desired it. Each also was given an instruction packet with screen pictures and examples; this packet was also available, along with user manuals for the four applications, beside the computer. To obtain a password for the system, participants had to fill out a prestudy questionnaire about past computer experience and previous searching with MEDLINE and other resources. Once logged on to the system, users were required to enter brief information about the patient and search topic before using an application. They also had to fill out a questionnaire after the search. The system recorded each log-on, the pre- and postsearch questionnaire data, and the application used. In addition, KF saved search statement data for each MEDLINE search. All these data were collected over more than ten months (August 22, 1991, to June 30, 1992), covering a major portion of the clinical year. In the last month of the study period, a poststudy questionnaire was given, focusing on attitudes towards computers and this specific system and a comparison of this system with previously used systems.

Because the majority of searches were done with KF and because most expert searchers have the most experience with MEDLINE, the authors chose to limit the evaluation of search quality to results obtained with KF. After the data were collected, all KF searches were reviewed, and a subset was extracted consisting of all searches where the user entered adequate information for replication (i.e., more than three words each of patient and topic information). After duplicate topics and author searches were eliminated, the subset contained 106 searches. (All employed free-text searching only, as did nearly all nonauthor

searches with KF.) To replicate these searches, eleven medical reference librarians and eleven physicians experienced with MEDLINE were recruited. Each librarian had to have used MEDLINE at least several times per week. Each physician had to have used MEDLINE more than once a month for more than two years and have an active clinical practice in an ambulatory setting.

The replicated searches were randomized so that each was searched by two physicians and two librarians, with one of each pair searching with the full MEDLINE feature set and the other using text words only. (The full MEDLINE feature set is defined as all of the usual features available in a traditional MEDLINE searching system, including MeSH terms, subheadings, explosions, and Boolean operators.) Searching was done on the National Library of Medicine (NLM) ELHILL computer, with searchers allowed use of either the native ELHILL interface or the GRATEFUL MED front end. Searchers were required to access the most recent six years of the database (i.e., current MEDLINE file and first back file). A reference was considered retrieved if it was displayed to the screen or returned by their last search statement.

Searching quality was assessed with the measures of recall and precision, defined as follows:

$$\text{Recall} = \frac{\text{number of relevant references retrieved}}{\text{number of relevant references in database}}$$

$$\text{Precision} = \frac{\text{number of relevant references retrieved}}{\text{total number of references retrieved}}$$

Because the total number of relevant references in a large database is usually impossible to judge, the measure of relative recall was used, in which the denominator is the number obtained by pooling all relevant references retrieved by all searchers.

$$\text{Relative recall} = \frac{\text{number of relevant references retrieved}}{\text{number of relevant references found by all searchers}}$$

The test database for the calculation of recall and precision consisted of all MEDLINE references from the 270 journals in the current KF subset for 1987 through 1991, a total of 348,566 references. References retrieved in the MEDLINE searches that were not in the test database were discarded for the performance assessment. Every reference in the test database retrieved for a given query was judged for relevance by physicians who were clinically active and at or above the level of senior medical resident. The judges used the patient information and topic designation

provided by the original searcher as a statement of the query, and part of the MEDLINE record (title, source, authors, abstract, and publication type) as the reference. The assessors were encouraged to seek the original articles when relevance could not be determined based on the title and abstract. The reviewers were not told which original searcher retrieved which articles.

Reviewers were asked to judge relevance from the standpoint of a clinician seeking an answer to the question posed. Relevance was judged on a three-point scale: definitely relevant (article provided highly relevant information for clinician faced with the recorded patient data and information need), possibly relevant (article might provide useful information to the clinician), and not relevant (article did not provide any relevant information for this information need). Approximately 11% of the judgments were duplicated to assess interobserver reliability as measured by the kappa statistic.

To summarize, the study design provided results for five searches on each query: the original KF search, two searches (librarian and clinician) using the full MEDLINE feature set (MeSH terms and text words), and two searches (librarian and clinician) using text words only. Searching performance for each group was characterized by the mean relative recall (hereafter referred to as "recall") and precision of the queries. For determination of the means, queries with no definitely relevant references were discarded, in order to retain a common set of queries for measuring recall and precision at definitely relevant and definitely plus possibly relevant levels. Statistical analysis of the results was performed using repeated measures analysis of variance. Post-hoc *t* tests for paired observations were performed using the Bonferroni correction.

RESULTS

The study group consisted of all ten interns, eleven junior residents, eleven senior residents, and nine faculty members who used the clinic for practice and teaching. All but one subject completed the prestudy questionnaire and obtained a computer password, leaving a group of forty. While almost all (36 of 40) had used MEDLINE before, most did so only monthly (10 of 36) or a few times a year (18 of 36). The system used most frequently for access to MEDLINE was ORHION, a local MEDLINE subset available in the university library and over the campus computer network. Despite having experience with MEDLINE, only one third had heard of MeSH terms or subheadings, and only three (all interns) had heard of MeSH tree explosions. Only six subjects had used databases besides MEDLINE (such as BIOETHICSLINE), and only three ever had used an "expert system" such as QMR.

Table 1
Responses to poststudy questionnaire

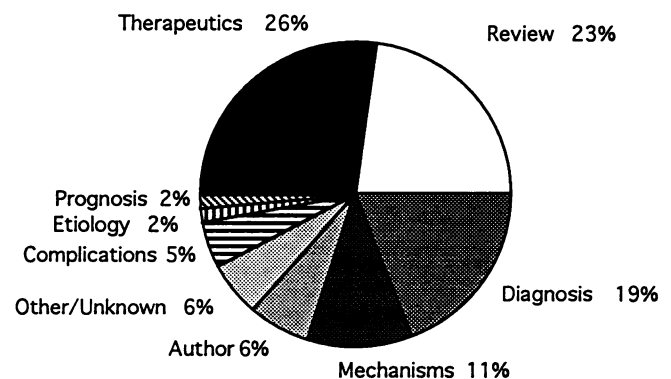
	Agree	Neutral	Disagree
Don't like to use computers	6	4	24
Computers too difficult to use	6	5	24
Workstation too difficult to use	2	3	26
Workstation too time-consuming to use	5	5	21
KF easier than other MEDLINEs	20	6	2
KF retrieves more relevant articles than other MEDLINEs	12	13	4
KF faster than other MEDLINEs	18	9	2
KF relevance ranking useful	20	6	3
More comfortable with MEDLINE than last year	21	6	2
Better at finding relevant articles than last year	21	8	1
Prefer KF MEDLINE to other MEDLINEs	19	7	3

Study subjects logged on to the workstation 301 times during the ten-month observation period. They started up applications 395 times. The majority of usage was with KF (270 start-ups), followed by modest use of SR (76), and minimal use of YB (8) and QMR (25). The system was used by more than three quarters of the subjects (31 of 40), and, for those who did use it, usage increased with level of medical training. The mean number of searches per person was 6.1 for interns, 12.3 for junior residents, 13.1 for senior residents, and 21.7 for faculty, for an overall mean of 12.7. Except for a slightly higher usage during the first two months and a drop-off in June, overall frequency of usage was relatively constant. The most common topic of searches was therapeutics, followed closely by diagnosis, general review (i.e., requests specifically designating a "review" or asking for information on more than one aspect of a topic, such as "diagnosis and management"), and mechanisms of disease (Figure 1).

The poststudy questionnaire was answered by thirty-seven (87.8%) of the subjects, with the results summarized in Table 1. Most users did not find computers in general or the workstation in particular too difficult or time-consuming to use. Most felt that KF was easier to use and faster than other MEDLINE systems and retrieved more relevant references. Approximately half of the group used a different MEDLINE searching system, almost exclusively ORHION, during the study period, although most used the other system less than they did KF.

For the replicated searches, two librarians and one physician used GRATEFUL MED, while the remainder used the native ELHILL interface. Table 2 shows the total number of references and average number of references per query retrieved by the five searching groups for all 106 queries, along with the proportion of those references that were in the test database. Only approximately one quarter of the references retrieved in the replicated searches were in the test

Figure 1
Search topics by subject



database. There were also a small number of references retrieved by KF (0.5% of the total) that were not in the test database, due to journals that were added or dropped over time on the KF CD-ROM. A total of 15,859 references (an average of 149.6 per query) were retrieved by all searchers from the test database, of which 12,565 (118.5 per query) were unique query-reference pairs.

There were 8,714 of the total (69.3%) references judged as not relevant, 2,053 (16.3%) judged as possibly relevant, and 1,798 (14.3%) judged as definitely relevant. The mean number of definitely relevant references per query was 17.0 (median, 9; range, 0-100). The mean number of possibly relevant references per query was 19.4 (median, 13; range, 0-81). There were 1,435 query-reference pairs (11.4%) judged in duplicate, with 1,003 agreements among the judges and 432 disagreements. The kappa score of 0.41 was comparable with other experiments of this type [37-39].

Recall and precision values for each group were calculated in two ways: first using only definitely relevant references and then using definitely plus possibly relevant references (Table 3). For both levels of relevance, the KF searchers obtained recall that was significantly higher ($P < .0001$) than that of other search groups. In turn, precision was significantly lower ($P < .0001$) than the non-KF groups. Librarians obtained better recall and precision using the full MEDLINE feature set than with text words, but the differences were not significant. Physicians obtained better recall with text words and better precision with the full feature set, but these differences also were not statistically significant. Librarians using the full feature set did have significantly ($P < .02$) higher recall than did physicians using the full feature set (but not text words) for definitely and possibly relevant articles.

Because the searchers using KF retrieved so many more references per search than did other groups, an

Table 2
References retrieved in test database

Group	References retrieved	References retrieved per search	% of references in test database
Clinic physicians	9,470	89.3	99.5
Librarians—full MEDLINE	7,032	66.3	27.1
Librarians—text words only	7,027	66.3	25.7
Physicians—full MEDLINE	4,550	42.9	25.3
Physicians—text words only	5,897	55.6	26.6

additional row was added to the table—KF searches cut off at a maximum of fifteen references per search (as opposed to the program's default operation of 100 references per search). (Fifteen was the average number of references retrieved per query in the test database by all the non-KF searchers). With this definition of retrieval, KF recall was not significantly different from that of either librarian or expert physician searchers. However, the precision of the KF searches based only on the top fifteen articles was significantly worse than searches using the full MEDLINE feature set conducted by librarians ($P < .01$ for definitely relevant articles, and $P < .0001$ for definitely plus possibly relevant articles) and by expert physicians ($P < .01$ for definitely relevant articles, and $P < .02$ for definitely plus possibly relevant articles).

Because a number of studies have shown that searchers tend to find nonoverlapping sets of relevant references, the authors also looked at the proportions of relevant references retrieved by one, two, three, four, or five searchers for each query. More than half of the 1,798 definitely relevant references were retrieved by only one of the five searchers (Table 4). Roughly one quarter were retrieved by only two searchers. Less than 10% of all definitely relevant references were retrieved by four or five searchers. When

Table 3
Mean percent recall and precision for the searching groups at different levels of relevance

Group	Re-trieved	Definitely relevant only		Definitely/possibly relevant	
		Re-call	Pre-cision	Re-call	Pre-cision
Clinic physicians—using KF	88.8	68.2	14.7	72.5	30.8
Clinic physicians—KF, top 15*	14.6	31.2	24.8	25.5	43.8
Librarians—full MEDLINE	18.0	37.1	36.1	30.8	59.4
Librarians—text words only	17.0	31.5	31.9	27.0	50.3
Physicians—full MEDLINE	10.9	26.6	34.9	19.8	55.2
Physicians—text words only	14.8	30.6	31.4	24.1	48.4

* Searches cut off at fifteen references.

Table 4
Number of searchers who retrieved relevant references

No. of searchers	Relevant references retrieved by	
	Full KF retrieval (%)	Reduced KF retrieval (%)
0	n.a.	429 (23.9)
1	957 (53.2)	788 (43.8)
2	474 (26.4)	355 (19.7)
3	190 (10.6)	121 (6.7)
4	99 (5.5)	73 (4.1)
5	42 (2.3)	32 (1.8)

the KF retrieval set was limited to fifteen references per search, the number of definitely relevant references not retrieved by any searcher was 23.9%.

When relevant references retrieved by only one searcher are categorized by the search group that retrieved them, the utility of KF becomes apparent. For the full KF retrieval set, more than one quarter of definitely relevant references were retrieved by physicians using KF alone. With the reduced KF set, librarians ranked highest for unique retrievals.

DISCUSSION

In general, the multi-application computer system was well-accepted and provided quality searches. Frequency of use was comparable with that of other clinical computer systems [40–42]. Direct comparison of usage rates would not be meaningful, because the other studies took place in different practice settings or with different specialties. Furthermore, the residents in this study used the clinic only one half-day per week, while the faculty used it only one to six half-days per week. The most important aspect of the usage rate was its low level in comparison to the high number of unanswered questions generated in clinical practice [43–45]; this divergence indicates that computer information systems still have a long way to go in playing a substantial role in meeting the physician's needs for information.

A major finding was that MEDLINE searching is still the most frequently used application, despite the availability of electronic versions of information resources more commonly used in clinical practice, namely textbooks and drug compendia. Their modest usage in this study indicates that electronic versions, despite being up-to-date and accessible, may not yet approach their paper counterparts in terms of convenience and ease of use. There was also only modest usage of the decision-support application QMR, most likely because it takes a long time to master and entering patient data is time-consuming. Both textbooks and decision-support applications may be accepted to a greater degree when computer use by physicians is routine, perhaps when (and if) more patient infor-

mation is available on computers [46] or physicians enter patient data directly [47]. In the latter case, applications running concurrently may be able to interact with physicians at the time of data entry, through either direct loading of patient findings into applications such as QMR (eliminating the need for retyping) or automatic generation of searches based on patient findings or hypothesized diagnoses [48-49].

The high user satisfaction with the system was consistent with results obtained with other clinically based systems. User satisfaction is certainly important but must be correlated with quality of results. This study looked at the quality of searching from two perspectives that are unique in the research literature on this topic: first, comparing use of a system such as KF in the hands of novice searchers to use of traditional tools by more experienced searchers and, second, comparing use of the full MEDLINE feature set to searches based only on text words.

The significantly higher recall obtained by the clinic physicians, even at the cost of diminished precision, shows that end users can access the medical literature effectively with tools such as KF. Certain aspects of the study may offer reasons for the significantly better recall obtained by novice searchers, beyond the value of KF's innovative approach. To begin with, both the experienced clinician and librarian searchers faced the handicap of starting with limited information, usually a brief sentence on the patient and information need. It has been shown that librarians perform better searches when they can interview the information seeker and obtain more detail on the exact information request [50]. Second, the experienced searchers might have searched in a different manner if their database contained only the references on the KF CD-ROM. While the 42.9 to 66.3 references retrieved per search indicated that strategies were appropriate for the database they were searching (six years of full MEDLINE), the experienced searchers might have broadened their searches if the initial retrievals had produced just the 10.9 to 18.0 references per search present in the test database. Broadening the search could have led to higher recall. Indeed, the higher recall of the novices may have been a function of their much larger retrieval sets, as shown by their more modest recall levels when the retrieval set was limited to the average size of the retrieval sets of the experienced searchers.

The differences in terms of results between the use of the full MEDLINE feature set and the use of text words by the experienced searchers were small and statistically insignificant, indicating that at least for the types of searches done in this study, the full MEDLINE feature set did not confer any advantage in retrieving relevant references. When comparing physicians with librarians using the full feature set, there

was a statistically significant advantage in recall for librarians, suggesting that advanced MEDLINE features are a tool of most benefit to librarians.

Another important finding was the small overlap in relevant articles retrieved by the different searchers. This phenomenon is well known to information scientists, one of whom has suggested that the best way to double recall in a search is to have another person repeat it [51]. Even experienced searchers use markedly different strategies for searching on the same topic, thereby producing different sets of relevant references [52]. It is important that searchers realize that, no matter how complete their results seem, there are probably additional relevant references to be found.

The use of recall and precision for assessing search system performance has several limitations. First, relevance can depend on the context of the search. For example, on a given topic, one searcher may just look for a good review article, while another may look for a particular type of study, such as a randomized controlled trial. Second, recall and precision are quantitative measures that say little about the quality of the references retrieved. The busy clinician often is seeking just a small number of references that he or she hopes will contain the information desired. As one study participant stated, "I don't want 40 relevant references, I want the answer." Finally, the clinical significance of differences in recall and precision is unknown. They do provide an estimate of the proportion of topically relevant articles each searcher retrieved, but it is unclear how large a difference is necessary to affect the overall quality of the search.

CONCLUSION

This study demonstrates that although computer information systems still have a limited role in the clinical setting, inexperienced searchers effectively can retrieve a large quantity of relevant information using innovative, easy-to-use software products such as KF. Study results suggest that efforts to make computerized medical information directly accessible in clinical settings will be beneficial. Further research not only must verify these findings but also must identify methods for improving access to the increasing wealth of electronic information and define how that information is best delivered to assist clinicians.

REFERENCES

1. COVELL D, UMAN G, MANNING P. Information needs in office practice: are they being met? *Ann Intern Med* 1985; 103:596-9.
2. OSHEROFF J, FORSYTHE D, BUCHANAN B, BANKOWITZ R ET AL. Physicians' information needs: analysis of questions posed during clinical teaching. *Ann Intern Med* 1991;114: 576-81.

3. ELY J, BURCH R, VINSON D. The information needs of family physicians: case-specific clinical questions. *J Fam Pract* 1992;35:265-9.
4. HOROWITZ G, JACKSON J, BLEICH H. PaperChase—self-service retrieval. *JAMA* 1983;328:2495-500.
5. HAYNES R, MCKIBBON K, WALKER C, RYAN N ET AL. Online access to MEDLINE in clinical settings. *Ann Intern Med* 1990;112(1):78-84.
6. COLLEN M. Full-text medical literature retrieval by computer: a pilot test. *JAMA* 1985;254:2768-74.
7. RODERER N, CLAYTON P. IAIMS at Columbia-Presbyterian Medical Center: accomplishments and challenges. *Bull Med Libr Assoc* 1992;80:253-62.
8. BROERING N, BAGDOYAN H. The impact of IAIMS at Georgetown: strategies and outcomes. *Bull Med Libr Assoc* 1992;80:263-75.
9. CIMINO C, BARNETT G. Standardizing access to computer-based medical resources. In: Miller R, ed. *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care*, November 4-7, 1990, Washington, DC. Los Alamitos, CA: IEEE, 1990:33-7.
10. OSHEROFF J, BANKOWITZ R. Physicians' use of computer software in answering clinical questions. *Bull Med Libr Assoc* 1993;81:11-9.
11. LUDWIG L, MIXTER J, EMANUELLE M. User attitudes toward end-user literature searching. *Bull Med Libr Assoc* 1988;76:7-13.
12. DALRYMPLE P. CD-ROM MEDLINE use and users: information transfer in the clinical setting. *Bull Med Libr Assoc* 1990;78:224-32.
13. SLINGLUFF D, LEV Y, EISAN A. An end-user search service in an academic health sciences library. *Med Ref Serv Q* 1985;4:11-21.
14. MARKERT RJ, PARISI AJ, BARNES HV, COHEN S ET AL. Medical student, resident, and faculty use of a computerized literature searching system. *Bull Med Libr Assoc* 1989;77:133-8.
15. MARSHALL J. The continuation of end-user online searching by health professionals: preliminary survey results. Abstracts of the Ninetieth Annual Meeting of the Medical Library Association May 18-24, 1990, Detroit, MI. Chicago: The Association, 1990.
16. HAYNES R, RAMSDEN M, MCKIBBON K, WALKER C. Online access to MEDLINE in clinical settings: impact of user fees. *Bull Med Libr Assoc* 1991;79:377-81.
17. SEWELL W, TEITELBAUM S. Observations of end-user online searching behavior over eleven years. *J Am Soc Info Sci* 1986;37:234-45.
18. MITCHELL J, JOHNSON E, HEWETT J, PROUD V. Medical students using Grateful Med: analysis of failed searches and a six-month follow-up study. *Comput Biomed Res* 1992;25:43-55.
19. WALKER C, MCKIBBON K, HAYNES R, RAMSDEN M. Problems encountered by clinical end users of MEDLINE and GRATEFUL MED. *Bull Med Libr Assoc* 1991;79:67-9.
20. KIRBY M, MILLER N. MEDLINE searching on Colleague: reasons for failure or success of untrained users. *Med Ref Serv Q* 1986;5:17-34.
21. LANCASTER F. *Evaluation of the MEDLARS Demand Search Service*. Bethesda, MD: National Library of Medicine, 1968.
22. KIRPALANI H, SCHMIDT B, MCKIBBON K, HAYNES R ET AL. Searching MEDLINE for randomized clinical trials involving care of the newborn. *Pediatrics* 1989;83:543-6.
23. HAYNES, Online access.
24. HERSH W, HICKAM D. A comparison of retrieval effectiveness for three methods of indexing medical literature. *Am J Med Sci* 1992;303:292-300.
25. FUNK M, REID C. Indexing consistency in MEDLINE. *Bull Med Libr Assoc* 1983;71:176-83.
26. EVANS D. Pragmatically-structured, lexical-semantic knowledge bases for unified medical language systems. In: Greenes R, ed. *Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care*, November 1988, Washington, DC. New York: IEEE, 1988:169-73.
27. MCKIBBON KA, HAYNES RB, DILKS CJ, RAMSDEN MF ET AL. How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. *Comput Biomed Res* 1990;23(6):583-93.
28. LANCASTER, op. cit.
29. KIRPALANI, op. cit.
30. HAYNES, Online access.
31. HERSH, op. cit.
32. DALRYMPLE, op. cit.
33. MCCARBERG B. MEDLINE Knowledge Finder. *JAMA* 1989;261:1812.
34. SALTON G. Developments in automatic text retrieval. *Science* 1991;253:974-80.
35. MILLER R, MASARIE F, MYERS J. Quick Medical Reference (QMR) for diagnostic assistance. *MD Comput* 1986;3:34-48.
36. MILLER R, POPLE H, MYERS J. INTERNIST-1: an experimental computer-based diagnostic consultant for general internal medicine. *New Engl J Med* 1982;307:468-76.
37. HAYNES, Online access.
38. HERSH, op. cit.
39. HERSH W, HICKAM D. A comparison of two methods for indexing and retrieval from a full-text medical database. *Med Decis Making* 1993;13:220-6.
40. HOROWITZ, op. cit.
41. HAYNES, Online access.
42. COLLEN, op. cit.
43. COVELL, op. cit.
44. OSHEROFF, Physicians' information needs.
45. ELY, op. cit.
46. DICK R, STEEN E. *The computer-based patient record: an essential technology for patient care*. Washington, DC: National Academy Press, 1991.
47. CIMINO J, BARNETT G. The physician's workstation: recording a physical examination using a controlled vocabulary. In: Stead W, ed. *Proceedings of the Eleventh Annual Symposium on Computer Applications in Medical Care*, Washington, DC. New York: IEEE, 1987:287-91.
48. CIMINO J, AGUIRRE A, JOHNSON S, PENG P. Generic queries for meeting clinical information needs. *Bull Med Libr Assoc* 1993;81:195-206.
49. POWSNER S, MILLER P. Linking bibliographic retrieval to clinical reports: Psychtopix. In: Kingsland L, ed. *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*, November 1989, Washington, DC. New York: IEEE, 1989:431-5.
50. SARACEVIC T, KANTOR P. A study of information seeking and retrieving. III. Searchers, searches, and overlap. *J Am Soc Info Sci* 1988;39:197-216.

51. SARACEVIC T. Individual differences in organizing, searching, and retrieving information. In: Griffiths JM, ed. Proceedings of the Fifty-fourth Annual Meeting of the American Society for Information Science. Washington, DC:

Learned Information, 1991:82-6.

52. MCKIBBON, op. cit.

Received February 1994; accepted May 1994