University of Massachusetts Medical School
eScholarship@UMMS

Population and Quantitative Health Sciences Publications | Population and Quantitative Health Sciences

1999-10-03

The art and science of searching MEDLINE to answer clinical questions. Finding the right number of articles

Jeroan J. Allison
University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/qhs_pp

Part of the Bioinformatics Commons, Biostatistics Commons, Epidemiology Commons, and the Health Services Research Commons

Repository Citation
Allison JJ, Kiefe CI, Weissman NW, Carter J, Centor RM. (1999). The art and science of searching MEDLINE to answer clinical questions. Finding the right number of articles. Population and Quantitative Health Sciences Publications. Retrieved from https://escholarship.umassmed.edu/qhs_pp/130

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Population and Quantitative Health Sciences Publications by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

International Journal of Technology Assessment in Health Care, 15:2 (1999), 281–296.
Copyright © 1999 Cambridge University Press. Printed in the U.S.A.

THE ART AND SCIENCE OF SEARCHING MEDLINE TO ANSWER CLINICAL QUESTIONS

Finding the Right Number of Articles

Jeroan J. Allison
Catarina I. Kiefe
Norman W. Weissman
Jerome Carter
Robert M. Centor
The University of Alabama at Birmingham

Abstract

The current medical environment makes information retrieval a matter of practical importance for clinicians. Many avenues present themselves to the clinician, but here we focus on MEDLINE by summarizing the current state of the art and providing an innovative approach for skill enhancement. Because new search engines appear rapidly, we focus on generic principles that can be easily adapted to various systems, even those not yet available. We propose an idealized classification system for the results of a MEDLINE search. Type A searches produce a few articles of high quality that are directly focused on the immediate question. Type B searches yield a large number of articles, some more relevant than others. Type C searches produce few or no articles, and those that are located are not germane. Providing that relevant, high-quality articles do exist, type B and C searches may often be improved with attention to search technique. Problems stem from poor recall and poor precision. The most daunting task lies in achieving the balance between too few and too many articles. By providing a theoretical framework and several practical examples, we prepare the searcher to overcome the following barriers: a) failure to begin with a well-built question; b) failure to use the Medical Subject Headings; c) failure to leverage the relationship between recall and precision; and d) failure to apply proper limits to the search. Thought and practice will increase the utility and enjoyment of searching MEDLINE.

Keywords: Medical Informatics, Information Storage and Retrieval, MEDLINE, Grateful Med, Sensitivity and Specificity

The evidence-based medicine movement, along with the current environment of expanding information and advancing technology, has transformed information

The authors wish to thank Marjorie Cahn of the National Library of Medicine and two anonymous reviewers for their helpful comments during the preparation of the manuscript.

This work was supported in part by the Agency for Health Care Policy and Research Grants HS-09446 and HS-08843.

retrieval into a matter of practical importance for clinicians (10;12;32;44;45). Several leading general medical journals now routinely carry articles on evidence-based medicine, but none has published an article focusing on general techniques useful to the clinician wishing to search MEDLINE. This topic is partly covered by several scattered articles and by two recent books (43;52). Here, we summarize the current state of the art and present a new approach for clinicians wishing to improve their skills. Our approach, based upon both personal use and teaching experience, continues to evolve (1;5;6).

ALTERNATIVES TO MEDLINE

Although we limit this discussion to MEDLINE, it is only one among many excellent information sources such as *Best Evidence* and the *Evidence-Based Medicine Journal* published by the American College of Physicians (20,39). The Cochrane Library provides comprehensive and scientific syntheses on an expanding array of topics (3;4).

These more specialized databases possess some important and unique features. First, the editors often apply scientific quality filters to the contents, eliminating work of low quality. Second, the search engines and search techniques assume an air of simplicity because of the limited number of entries in each database compared with MEDLINE. However, the restricted nature of these databases means that many topics of interest will be omitted.

ALTERNATIVE GATEWAYS TO MEDLINE

Maintained by the United States National Library of Medicine (NLM) and offered free of charge to the public on the Internet's World Wide Web, MEDLINE contains approximately 9 million articles from 3,900 journals at this writing. (These figures will be dated by publication of this manuscript.) The first articles are from 1960, and each database entry contains multiple fields such as title, abstract, author, institution, source, and publication type (59). Approximately 75% of articles after 1975 have abstracts, and 80% are in English. Users may search all fields, including abstracts. MEDLINE belongs to the MEDLARS family of databases, which contains more than 40 other databases such as HealthSTAR, TOXLINE, AIDSLINE, and CancerLit.

Many commercial vendors offer MEDLINE in a variety of formats, which usually contain a core of basic features that are more alike than different. These commercial offerings, based on MEDLINE downloads licensed by the NLM, are updated every 1 to 3 months. The NLM offers Internet access through two web sites, PubMed and Internet Grateful Med (IGM). PubMed (*http://www.ncbi.nlm.nih.gov/PubMed/*) offers a transparent search engine and linkages to sites for full-text journals. IGM (*http://igm.nlm.nih.gov/*) provides a more sophisticated search engine and offers access to several additional databases and to Loansome Doc, through which users may order full-text documents. Both PUBMED and IGM offer access to PREMEDLINE, where articles are added daily before their entry into MEDLINE. (PREMEDLINE is not indexed; therefore, the Medical Subject Headings [MeSH]-based techniques described below do not apply.) IGM offers a search refinement function that will automatically implement many of the strategies discussed in this paper. These interfaces garner much attention because they are available without charge and are easy to use, and both appear to be evolving in a such a manner as to make them more alike than different.

Table 1. Classification of MEDLINE Search Results

Results	Characteristics of articles: Quantity	Quality	Relevance	Recommendations
Type A	Just right	High	High	Celebrate!
Type B	Large numbers, often too many	Variable	Variable	Improve precision Focus on review articles
Type C	Too few	Variable often low	Variable often low	Improve recall

Because new search engines are appearing rapidly, detailed reviews of the advantages and disadvantages of specific systems and interfaces are beyond the scope of this paper (23;53). (For technical details about PUBMED and IGM, please refer to the excellent on-line documentation.) Therefore, rather than dwelling on the mechanics unique to any single search program, which may soon be outdated, we instead present some important principles that can be applied universally. Application of these principles demands only a rudimentary familiarity with MEDLINE.

CLASSIFICATION OF MEDLINE SEARCH RESULTS

For instructional purposes, we propose the idealized classification system for MEDLINE search results as described in Table 1. No one search will produce results belonging exclusively to one search type. Instead, searches are classified by the user according to the reason for dissatisfaction (if any) with the results. Type A searches produce a few articles of high quality (scientific methodology and statistical technique) that are directly focused on the immediate question. Type B searches, which are appropriate under certain circumstances, yield a large number of articles, some more relevant than others. Type B searches may overwhelm the searcher by sheer volume. We leave the determination of what constitutes an excessive number of articles to the tolerance of the individual searcher, but we place our maximum at about 50. However, in specific situations where a high level of comprehensiveness is required, we must tolerate the increased number of irrelevant articles.

We can often seek refuge from the problem of excessive numbers in the review article. Review articles allow the physician to rely on others to carry out the complex and tedious task of information synthesis. With review articles, however, attention to quality is especially important. Reviews that confuse opinion with evidence may present a biased view from lack of attention to well-accepted scientific principles. A series on the systematic review, edited by Mulrow et al., appears in the *Annals of Internal Medicine* (41). Systematic reviews are "scientific investigations" based upon well-recognized "strategies that limit bias and random error" (9). Meta-analyses form a subset of systematic reviews. In a meta-analysis, statistical methods are used to combine the results of multiple studies. Cook et al. (9) offer a more elaborate discussion of the advantages and disadvantages of systematic reviews; further details fall beyond the scope of this paper.

As discussed below, Appendix 1 provides strategies for retrieving systematic reviews and meta-analyses. When a good review article is found, a follow-up MEDLINE search should cover the period of time subsequent to the most recent reference in the review. However, good review articles are not always available, and their absence mandates the use of other techniques.

Type C searches produce few or no articles, and those that are located are not germane. Providing that relevant, high-quality articles do exist, type B and C searches may often be improved with attention to search technique. Problems with these types of results stem from lack of precision (number of relevant articles retrieved/total number of articles retrieved) and low recall (number of relevant articles retrieved/number of relevant articles in the database). However, research demonstrates that searchers experience significant improvement in their skills with practice (21;22). For example, Pao et al. (48) found that between one and two MEDLINE searches per month during the first years of medical school improved the ability to find relevant items by a factor of eight.

The most daunting task facing the searcher lies in achieving the balance between too few and too many articles. Here, we address the following barriers to achieving this goal of effective and efficient searching: a) failure to begin with a well-built clinical question; b) failure to use the MeSH indexing system; c) failure to leverage the important relationship between recall and precision; and d) failure to apply proper limits to the search.

COMPONENTS OF EFFECTIVE MEDLINE SEARCHES

Start with a Well-built Clinical Question

The well-built clinical question forms the foundation of an effective search. The following elements comprise an essential core for clinical questions: a) patient; b) intervention; and c) outcome. The McMaster Evidence-based Medicine Working Group, in their *Users' Guides to the Medical Literature,* applies these three elements to several common clinical situations (11;13;14;15;16;17;25;27;28;29;31;42;46;47;49;50).

Proper entry of the first element, patient type, demands careful consideration of the characteristics of how the results of the search will be applied. For example, in searching for evidence on the treatment of hyperlipidemia in an 80-year-old woman, one must recognize that many studies on the treatment of hyperlipidemia exclude the elderly and/or women. A current controversy surrounds the extrapolation of the results of treating the young to treating the elderly (19;30). Analogous problems exist with gender. In addition, the third element, outcome, may mislead when surrogate markers are used in place of clinically meaningful outcomes (16). Our 80-year-old patient probably does not care whether an intervention will change the parameters of her lipid profile or modify her fibrinogen levels so much as whether the intervention will prolong her life, prevent a heart attack or congestive heart failure, or otherwise preserve her quality of life.

Use of Medical Subject Headings

The MeSH indexing system, maintained by the NLM, is a boon to all searchers of MEDLINE. MeSH is a controlled database indexing language designed to circumvent the imprecision of the synonyms and variations in medical jargon. For example, different authors may refer to the same concept of elevated serum lipids using varied terminology such as elevated cholesterol, abnormal lipid profile, hypercholesterolemia, hyperlipidemia, or dyslipidemia. Each MeSH descriptor represents a single medical concept, and all of the MeSH terms are contained within a hierarchy, which is represented by the MeSH tree (Figure 1). MeSH frequently has many "entry terms," or synonyms for concepts, that point to the same descriptor.

The MeSH tree structure begins with the broadest concepts at the highest level and progressively branches into more narrow concepts. Articles are indexed by

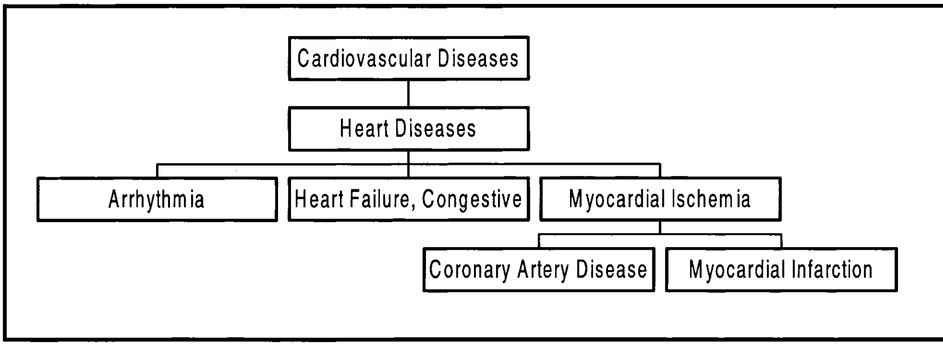

Figure 1. MeSH tree segment.

their most specific term. For example, if an article is indexed under "myocardial infarction," it will not be retrieved by using the MeSH heading of "cardiovascular diseases," although the former is, of course, a subset of the later. The *explode* feature, however, retrieves all articles indexed at a given level in addition to those below. Exploding "cardiovascular diseases" would therefore retrieve all articles under "myocardial infarction," in addition to all those below "cardiovascular diseases," such as "myocardial ischemia" and "coronary artery disease." Both PUBMED and IGM automatically explode terms now.

As each article is added to MEDLINE, the NLM applies an average of 10 MeSH terms to it according to standardized criteria. Each year the NLM updates the MeSH terminology as new medical concepts appear. The few MeSH terms that represent the main focus of the article receive a *major concept* designation. The *major concept* and *explode* features perform complementary but not opposite functions, and they may be appropriately used in combination. For example, suppose one is interested in searching the general topic of hyperthyroidism. As shown in Table 2, lack of familiarity with MeSH could lead to the erroneous conclusion that

Table 2. Illustrative MEDLINE Searches (1993–July 1997)[a]

Set	Search term(s)[b]	Results
1	Hyperthyroidism	1,119
2.	Graves' disease	1,321
3	1 and 2	130
4	Exp hyperthyroidism	2,310
5	*Exp hyperthyroidism	1,677
6	3 not 4	633
7	Myocardial infarction	11,017
8	(Treatment or therapy) (tw)	283,068
9	7 and 9	3,380
10	Myocardial infarction (sh/drug therapy, therapy)	4,015
11	Limit 10 to clinical trial	698
12	Sweet's syndrome	154
13	Lupus erythematosus, systemic	3,671
14	12 and 13	0
15	Sweet: (tw) or Sweet's syndrome	1,105
16	Lupus (tw) or lupus erythematosus, systemic	5,824
17	16 and 17	4

[a] See Appendix for key.
[b] All terms not otherwise labeled are MeSH terms.

the 1,119 articles under this term represent the complete picture (set 1). However, the MeSH term “Graves’ disease,” which produces 1,321 (set 2) articles, falls below “hyperthyroidism.” Of these 1,321 articles, only 130 (set 3) were captured by the higher order term of “hyperthyroidism.” Exploding “hyperthyroidism” captures all 2,310 articles (set 4). However, many of these articles are only peripherally related to our topic, and using the major concept designation to focus this already exploded search yields 1,667 articles (set 5) by eliminating 633 articles (set 6).

In addition, each article is tagged with publication type and subheading designations, useful in further refining a search. Examples of publication types include *randomized controlled trial, review, review of reported cases,* and *editorial*. Subheadings are general qualifiers, such as *complications, diagnosis, treatment, epidemiology,* and *economics,* that are linked with a specific MeSH term. Subheadings in themselves lack a precise meaning; however, when combined with a MeSH term, they efficiently isolate specific aspects of a broad subject area. The searcher can use certain “pre-exploded” subheadings to include all related subheadings. These groups of logically related subheadings exist in exploded form in the MeSH system. Most search engines allow the user to pick from a list of available publication types and subheadings.

Of all the features of the many MEDLINE product lines, searchers benefit most from mapping. Mapping translates textual information entered by the searcher into the most appropriate MeSH heading. Often the mapping function provides an array of matching terms and asks the searcher to choose the most relevant. This frees the physician from acquiring a detailed knowledge of MeSH. Some search engines automatically invoke mapping, even without the awareness of the searcher. Some engines allow the searcher to activate or inactivate the mapping function. As the NLM now devotes considerable resources to the Unified Medical Language System with its metathesaurus, mapping capabilities continue to improve. (For details please refer to the NLM home page at *http://www.nlm.nih.gov/*.)

When mapping is inactivated, the searcher must manually enter the correct MeSH heading or perform a text word search. Often systems process text word searches differently. For example, with one commercial system a text word search queries the MEDLINE fields of title and/or abstract for articles containing the word(s) entered by the searcher. However, IGM queries, among others, the title, abstract, MeSH, MeSH subheading, and personal name fields. For the novice, text word searching may lead to poor results.

We regard mapping to MeSH as the best starting point for clinicians and medical educators seeking answers to clinical questions. As described below, we view various text-based strategies as a second-line, and often unnecessary, approach in the refinement of a search. One notable exception includes newly emergent concepts, which may not have yet been incorporated into MeSH. Many investigators have examined the relative merits of searching by text word or by a controlled-language indexing system such as MeSH (24;40;57). Clearly, in experienced hands an array of strategies is useful. However, we agree with Lowe and Barnett, who aptly stated:

> The lesson here is that the MeSH indexing performed by the NLM is a form of intelligent preprocessing that should be taken advantage of whenever possible. Failure to do so is an important reason why MEDLINE searches fail. (33)

Most programs provide a mapping tool to ease the navigation through MeSH. A typical tool allows one to accomplish the following maneuvers with minimal difficulty: a) map any text word to the equivalent MeSH term; b) view any MeSH

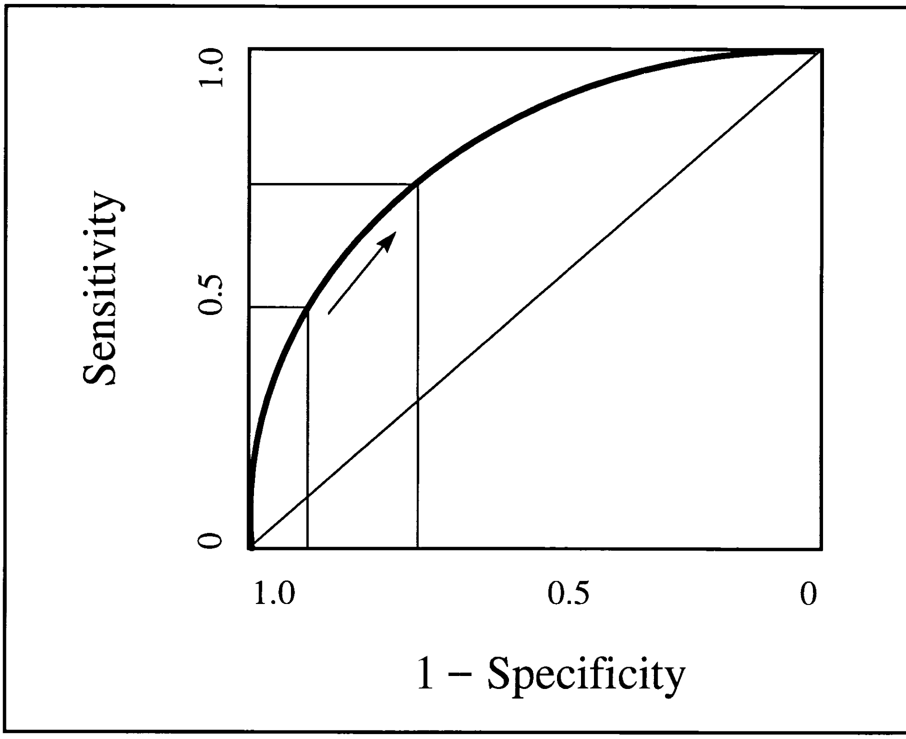

Figure 2. Generic receiver operator characteristic (ROC) curve.

term in the context of its place in the MeSH tree structure; c) choose from MeSH terms that are higher or lower in the hierarchy to broaden or narrow the search; d) explode the search; e) focus the search to include only those articles with MeSH terms tagged as major concepts; f) apply publication types and subheadings; and g) display "scope notes," or help text explaining a MeSH term.

Leverage the Relationship between Recall and Precision

The general problem of signal detection, or distinguishing noise from information, is relevant to searching MEDLINE (2;8;18). A similar problem surfaced in World War II when the military sought to use radar to distinguish between geese and bombers. As the sensitivity on the radar was increased, few bombers were missed. However, there were many false alarms. As the specificity of the radar was increased, almost all alarms were true alarms, but at the expense of missing several bombers. Centor (7) reviewed the graphical representation of such behavior as receiver operator characteristic (ROC) curves. ROC curves describe the behavior of many phenomena involving diagnostic tests and information retrieval systems. As shown in Figure 2, moving along the ROC curve in the direction of the arrow increases sensitivity and decreases specificity.

For the MEDLINE searcher, the problem is culling the irrelevant and poor-quality articles without eliminating those of high relevance and quality. Unfortunately, as the ROC curves show, it is not possible to maximize both the recall and the precision of a MEDLINE search. *Herein lies the art of searching MEDLINE: knowing the appropriate balance to achieve for a specific purpose.* For example, a resident in training seeking an article for morning report or an educator looking for a good review article for rounds might sacrifice recall. In contrast, when preparing a systematic review for publication or writing a grant application, recall assumes more importance.

Performing a MEDLINE search bears marked similarity to performing a medical test in that both convert a crude preprocedure estimate into a more refined postprocedure estimate. With the MEDLINE search, the preprocedure estimate is the proportion of relevant articles in the entire database and the postprocedure estimate is the proportion of relevant articles in the results of the search. For the medical test, the pre-procedure estimate is the baseline probability of disease for a given patient, and the post-procedure estimate is the more accurate assessment of the probability of disease obtained after the test.

Following this analogy, the precision of a search (number of relevant articles retrieved/ number of articles retrieved) is analogous to the positive predictive value (number of patients with disease/number of patients testing positive) of a medical test, and recall (number of relevant articles retrieved/number of relevant articles in the database) is analogous to sensitivity (number of patients testing positive for disease/number of patients with disease). Extending Baye's theorem to this problem implies that precision of a search improves with increasing sensitivity, increasing specificity (number of irrelevant articles not retrieved/number of irrelevant articles in the database) and increasing prevalence (number of relevant articles in entire database/number of articles in entire database) (51). Table 3 summarizes this terminology.

Table 4 summarizes several techniques for increasing either recall or precision of a previously completed MEDLINE search. MeSH contains built-in features allowing the searcher to adjust both recall and precision. We have already discussed many of these. The hierarchical structure of the MeSH tree allows the searcher to easily broaden or narrow the search. In addition, reviewing the MeSH terms applied to an article of known relevance and then performing additional searches with these MeSH terms increases recall (54).

For text-based searches, an increased number of synonyms from reviewing articles of known relevance or brainstorming will obviously increase recall. Truncation instructs the search engine to query specific MEDLINE database fields for semantic variations. For example, in many applications, "pharmac:" retrieves pharmacy, pharmacology, and pharmacological.

The Boolean operators OR, AND, and NOT are useful in combining the results of separate searches. Using OR creates a new set containing all articles from the original sets, exclusive of duplicates. Using AND creates a new set containing only articles found in all original sets. Many searches fail because users combine too many terms with AND or NOT in a single search, thus producing overly restrictive criteria for which there are few (or no) matches. Using NOT eliminates a known source of irrelevant articles. However, in eliminating known irrelevant articles, some relevant articles in a particular set will also be eliminated. In general, MeSH subheadings provide better alternatives than complex Boolean searches (33).

Proximity searches involve searching for multiple words within a specified distance in either the title or abstract fields. For example, one might search for articles where "cholesterol" and "elderly" appear within five words of each other in the abstract. As the distance between the words is increased, the search becomes more sensitive and less specific. Searching for juxtaposed words attains maximum precision.

Table 3. Definitions of Terminology

Search terminology		Diagnostic testing terminology	
Item	Definition	Item	Definition
Precision	$\frac{\text{No. of articles relevant retrieved}}{\text{No. of articles retrieved}}$	Positive predictive value	$\frac{\text{No. of patients with disease}}{\text{No. of patients testing (+)}}$
Recall	$\frac{\text{No. of relevant articles retrieved}}{\text{No. of relevant articles in database}}$	Sensitivity	$\frac{\text{No. of patients testing (+)}}{\text{No. of patients with disease}}$
Specificity	$\frac{\text{No. of irrelevant articles not retrieved}}{\text{No. of irrelevant articles in database}}$	Specificity	$\frac{\text{No. of patients testing (−)}}{\text{No. of patients without disease}}$
Prevalence	$\frac{\text{No. of relevant articles in database}}{\text{No. of articles in database}}$	Prevalence	$\frac{\text{No. of patients with disease}}{\text{No. of patients in population}}$

Table 4. Refinement Techniques for Search Strategies

Technique	Increasing recall	Increasing precision
MeSH	Explode	Major concept
	Broaden	Narrow
	Consult known articles for additional terms	Subheadings
		Publication types
Text	Synonyms	
	Truncation	
	Consult known articles for additional terms	
Boolean	Increase terms with OR	Decrease terms with OR
	Decrease terms with AND, NOT	Increase terms with AND, NOT
	Proximity search	Proximity search
Mixed	Methodological filters	Methodological filters

Methodological filters, developed by the McMaster team, now appear in several search engines, including the clinical queries feature of PubMed. These filters depend upon combinations of text words, MeSH headings, and Boolean operators (34;35;36;37;38;55) and have known recall rates and precision. These rates generally improve as more recent years of MEDLINE are searched (24;56;58). Mostly, we find that MeSH provides the best starting place for a search, but we turn to these more complex methodological filters when a MeSH-based search is inadequate or for other special cases. The PubMed search engine transparently guides the user in invoking these methodological filters, but without such assistance the filters are often difficult to apply.

A discussion of methodological filters falls beyond the scope of this paper, but because of their usefulness, we have included them in full in Appendix 1. The filters for systematic reviews warrant special attention because the MeSH indexing process does not yet distinguish between systematic and nonsystematic reviews (26).

A key to MEDLINE abbreviations is included in Appendix 2.

Apply Search Limits

The application of search limits is important but does not form an integral part of the search strategy. A well-executed search benefits from limits applied post hoc to remove unwanted documents. For clinical questions, many choose to apply the common limits of "Human," "English," and "Abridged Index Medicus" (AIM). The AIM contains a core set of 125 journals selected from the more than 3,900 journals covered by MEDLINE. Although the AIM has not been updated in 20 years, for the most part articles not in the AIM are useful for specialized research topics or for investigators conducting systematic reviews.

Examples

Two examples of searches in MEDLINE (1993–97) serve to illustrate key points. First, suppose that we are interested in the treatment of myocardial infarction. From Table 2, the MeSH term (set 7) produces 11,107 articles. This is obviously a type B search. One way of increasing the precision of the search involves a strategy of first identifying all articles that contain the words *treatment* or *therapy* in their title or abstract, as shown in set 8. Set 9 is then produced using the Boolean AND to derive the intersection of sets 7 and 8. This procedure is more logistically difficult and less advisable than the strategy of choosing the MeSH subheadings of *drug therapy, therapy,* and *surgery* from a list provided by the MEDLINE thesaurus (set 9). To improve the quality of the evidence presented in the articles, we next limited set 10 to only those dealing with clinical trials.

In a second example, we desire to know if there is an association between Sweet's syndrome and lupus. As noted in Table 3, the first MeSH-based strategy produced type C results with no relevant articles (set 14). Next, we employed a text-based strategy of adding synonyms to increase the recall of the search. Truncation was used to identify such variations as Sweet's syndrome, Sweet syndrome, Sweet's disease, etc. This approach was more successful (set 17). (Actually, the yield was increased even further when this strategy was pursued in an older MEDLINE database.)

CONCLUSION

Our personal experience suggests that searching MEDLINE leads to increased physician satisfaction. With some practice and attention to basics, searching skills can be rapidly increased. However, we challenge those who view designing a MEDLINE search as a nonintellectual activity. We hope that physicians improve their patient care, education, and research by practicing the art and science of searching MEDLINE.

REFERENCES

1. Allison, J., Box, B., & Kiefe, C. *Teaching internal medicine residents to effectively search the medical literature*. New Orleans, LA: Southern Society of General Internal Medicine Workshop, 1995.
2. Beck, J. R., & Shultz, E. K. The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology and Laboratory Medicine*,1986, 110, 13–20.
3. Bero, L., & Rennie, D. The Cochrane Collaboration. Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA*, 1995, 274, 1935–38.
4. Bulger, R. J. Archie Cochrane revisited. *JAMA*, 1988, 260, 1284.
5. Carter, J. H. Medical informatics in postgraduate training: A way to improve office-based practitioner information management. *Journal of General Internal Medicine*, 1991, 6, 349–54.
6. Carter, J., Allison, J., Herbert-Carter, J., et al. Using computers to help manage clinical information: An ACP workshop. Birmingham, AL: American College of Physicians, 1995.
7. Centor, R. M. Signal detectability: The use of ROC curves and their analyses. *Medical Decision Making*, 1991, 11, 102–06.
8. Centor, R. M., & Schwartz, J. S. An evaluation of methods for estimating the area under the receiver operating characteristic (ROC) curve. *Medical Decision Making*, 1985, 5, 149–56.
9. Cook, D. J., Mulrow, C. D., & Haynes, R. B. Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, 1997, 126, 376–80.
10. Crawley, L. Evidence-based medicine: A new paradigm for the patient. *JAMA*, 1993, 269, 1253–54.
11. Drummond, M. F., Richardson, W. S., O'Brien, B. J., Levine, M., & Heyland, D. Users' guides to the medical literature, XIII: How to use an article on economic analysis of clinical practice, A: Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA*, 1997, 277, 1552–57.
12. Gorman, P. Does the medical literature contain the evidence to answer the questions of primary care physicians? Preliminary findings of a study. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1993, 571-75.

13. Guyatt, G. H., Naylor, C. D., Juniper, E., et al. Users' guides to the medical literature, XII: How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. *JAMA*, 1997, 277, 1232–37.
14. Guyatt, G. H., & Rennie, D. Users' guides to the medical literature. *JAMA*, 1993, 270, 2096–97.
15. Guyatt, G. H., Sackett, D. L., & Cook, D. J. Users' guides to the medical literature, II: How to use an article about therapy or prevention, A: Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA*, 1993, 270, 2598–601.
16. Guyatt, G. H., Sackett, D. L., & Cook, D. J. Users' guides to the medical literature, II: How to use an article about therapy or prevention, B: What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA*, 1994, 271, 59–63.
17. Guyatt, G. H., Sackett, D. L., Sinclair, J. C., et al. Users' guides to the medical literature, IX: A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA*, 1995, 274, 1800–04.
18. Hanley, J. A., & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 1982, 143, 29–36.
19. Havel, R. J., & Rapaport, E. Management of primary hyperlipidemia. *New England Journal of Medicine*, 1995, 332, 1491–98.
20. Haynes, R. B. Current awareness and current access: ACP Journal Club goes electronic. *ACP Journal Club*, 1995, 123, A14.
21. Haynes, R. B., Johnston, M. E., McKibbon, K. A., Walker, C. J., & Willan, A. R. A program to enhance clinical use of MEDLINE. A randomized controlled trial. *Online Journal of Current Clinical Trials*, 1993.
22. Haynes, R. B., McKibbon, K. A., Walker, C. J., et al. Online access to MEDLINE in clinical settings. A study of use and usefulness. *Annals of Internal Medicine*, 1990, 112, 78–84.
23. Haynes, R. B., Walker, C. J., McKibbon, K. A., Johnston, M. E., & Willan, A. R. Performances of 27 MEDLINE systems tested by searches with clinical questions. *Journal of the American Medical Informatics Association*, 1994, 1, 285–95.
24. Haynes, R. B., Wilczynski, N., McKibbon, K. A., Walker, C. J., & Sinclair, J. C. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *Journal of the American Medical Informatics Association*, 1994, 1, 447–58.
25. Hayward, R. S., Wilson, M. C., Tunis, S. R., Bass, E. B., & Guyatt, G. Users' guides to the medical literature, VIII: How to use clinical practice guidelines, A: Are the recommendations valid? The Evidence-Based Medicine Working Group. *JAMA*, 1995, 274, 570–74.
26. Hunt, D. L., & McKibbon, K. A. Locating and appraising systematic reviews. *Annals of Internal Medicine*, 1997, 126, 532–38.
27. Jaeschke, R., Guyatt, G., & Sackett, D. L. Users' guides to the medical literature, III: How to use an article about a diagnostic test, A: Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA*, 1994, 271, 389–91.
28. Jaeschke, R., Guyatt, G. H., & Sackett, D. L. Users' guides to the medical literature, III: How to use an article about a diagnostic test, B: What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*, 1994, 271, 703–07.
29. Laupacis, A., Wells, G., Richardson, W. S., & Tugwell, P. Users' guides to the medical literature, V: How to use an article about prognosis. Evidence-Based Medicine Working Group. *JAMA*, 1994, 272, 234–37.
30. Leaf, D. A. Lipid disorders: Applying new guidelines to your older patients. *Geriatrics*, 1994, 49, 35–41.
31. Levine, M., Walter, S., Lee, H., et al. Users' guides to the medical literature, IV: How to use an article about harm. Evidence-Based Medicine Working Group. *JAMA*, 1994, 271, 1615–19.

32. Lindberg, D. A., Siegel, E. R., Rapp, B. A., Wallingford, K. T., & Wilson, S. R. Use of MEDLINE by physicians for clinical problem solving. *JAMA*, 1993, 269, 3124–29.
33. Lowe, H. J., & Barnett, G. O. Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches. *JAMA*, 1994, 271, 1103–08.
34. McKibbon, K. A., & Walker-Dilks, C. J. Beyond ACP Journal Club: How to harness MEDLINE for diagnostic problems. *ACP Journal Club*, 1994, 121, A10–12.
35. McKibbon, K. A., & Walker-Dilks, C. J. Beyond ACP Journal Club: How to harness MEDLINE for therapy problems. *ACP Journal Club*, 1994, 121, A10–12.
36. McKibbon, K. A., & Walker-Dilks, C. J. Beyond ACP Journal Club: How to harness MEDLINE to solve clinical problems. *ACP Journal Club*, 1994, 120, A10–12.
37. McKibbon, K. A., Walker-Dilks, C., Haynes, R. B., & Wilczynski, N. Beyond ACP Journal Club: How to harness MEDLINE for prognosis problems. *ACP Journal Club*, 1995, 123, A12–14.
38. McKibbon, K. A., Walker-Dilks, C. F., Wilczynski, N. L., & Haynes, R. B. Beyond ACP Journal Club: How to harness MEDLINE for review articles. *ACP Journal Club*, 1996, 124, A12–13.
39. Miles, A. Evidence based medicine. *Journal of Evaluation in Clinical Practice* will start publication in September. *BMJ*, 1995, 311, 258–59.
40. Mitchell, J. A., Johnson, E. D., Hewett, J. E., & Proud, V. K. Medical students using Grateful Med: Analysis of failed searches and a six-month follow-up study. *Computers and Biomedical Research*, 1992, 25, 43–55.
41. Mulrow, C. D., Cook, D. J., & Davidoff, F. Systematic reviews: Critical links in the great chain of evidence. *Annals of Internal Medicine*, 1997, 126, 389–91.
42. Naylor, C. D., & Guyatt, G. H. Users' guides to the medical literature, XI: How to use an article about a clinical utilization review. Evidence-Based Medicine Working Group. *JAMA*, 1996, 275, 1435–39.
43. Osheroff, J. A. (ed.). Computers in Clinical Practice. Philadelphia, PA: *American College of Physicians*, 1995.
44. Osheroff, J. A., & Bankowitz, R. A. Physicians' use of computer software in answering clinical questions. *Bulletin of the Medical Library Association*, 1993, 81, 11–19.
45. Osheroff, J. A., Forsythe, D. E., Buchanan, B. G., et al. Physicians' information needs: Analysis of questions posed during clinical teaching. *Annals of Internal Medicine*, 1991, 114, 576–81.
46. Oxman, A. D., Sackett, D. L., & Guyatt, G. H. Users' guides to the medical literature, I: How to get started. The Evidence-Based Medicine Working Group. *JAMA*, 1993, 270, 2093–95.
47. Oxman, A. D., Cook, D. J., & Guyatt, G. H. Users' guides to the medical literature, VI: How to use an overview. Evidence-Based Medicine Working Group. *JAMA*, 1994, 272, 1367–71.
48. Pao, M. L., Grefsheim, S. F., Barclay, M. L., et al. Effect of search experience on sustained MEDLINE usage by students. *Academic Medicine*, 1994, 69, 914–20.
49. Richardson, W. S., & Detsky, A. S. Users' guides to the medical literature, VII: How to use a clinical decision analysis, A: Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA*, 1995, 273, 1292–95.
50. Richardson, W. S., & Detsky, A. S. Users' guides to the medical literature, VII: How to use a clinical decision analysis, B: What are the results and will they help me in caring for my patients? Evidence Based Medicine Working Group. *JAMA*, 1995, 273, 1610–13.
51. Sackett, D. L., Haynes, R. B., & Tugwell, P. *Clinical epidemiology: A basic science for clinical medicine*, 2nd ed. Boston, MA: Little and Brown, 1991.
52. Sackett, D. L., Richardson, W. S., Rosenberg, W., & Haynes, R. B. *Evidence-based medicine:* How to practice and teach EBM. Secaucus, NJ: Churchill Livingstone, 1996.
53. Schoonbaert, D. SPIRS, WinSPIRS, and OVID: A comparison of three MEDLINE-on-CD-ROM interfaces. *Bulletin of the Medical Library Association*, 1996, 84, 63–70.
54. Srinivasan, P. Retrieval feedback in MEDLINE. *Journal of the American Medical Informatics Association*, 1996, 3, 157–67.

55. Walker-Dilks, C. J., McKibbon, K. A., & Haynes, R. B. Beyond ACP Journal Club: How to harness MEDLINE for etiology problems. *ACP Journal Club*, 1994, 121, A10–11.
56. Wilczynski, N. L., Walker, C. J., McKibbon, K. A., & Haynes, R. B. Assessment of methodologic search filters in MEDLINE. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1993, 601–05.
57. Wilczynski, N. L., Walker, C. J., McKibbon, K. A., & Haynes, R. B. Quantitative comparison of pre-explosions and subheadings with methodologic search terms in MEDLINE. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1994, 905–09.
58. Wilczynski, N. L., Walker, C. J., McKibbon, K. A., & Haynes, R. B. Reasons for the loss of sensitivity and specificity of methodologic MeSH terms and textwords in MEDLINE. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1995, 436–40.
59. Wood, E. H. MEDLINE: The options for health professionals. *Journal of the American Medical Informatics Association*, 1994, 1, 372–80.

APPENDIX 1
METHODOLOGICAL FILTERS[a,b]

Problem type	Search strategy[c]	Sensitivity	Specificity
Therapeutic			
Strategy 1	Placebo: (tw) OR [Double (tw) AND Blind: (tw)]	57%	97%
Strategy 2	Randomized Controlled Trial (pt) OR Drug Therapy (sh) or Therapeutic Use (sh) OR Random: (tw)	99%	74%
Diagnostic			
Strategy 1	Exp [Sensitivity and Specificity] OR Predictive (tw) OR Value: (tw)	55%	98%
Strategy 2	Exp [Sensitivity and Specificity] OR Diagnosis& (px) OR Diagnostic Use (sh) OR Sensitivity (tw) OR Specificity (tw)	92%	73%
Etiologic			
Strategy 1	Cohort Studies OR Case-control Studies	40%	97%
Strategy 2	Exp Cohort Studies OR Exp risk OR [Odds (tw) and Ratio: (tw)] OR [Relative (tw) and Risk (tw) OR [Case (tw) and Control: (tw)]	82%	70%
Prognostic			
Strategy 1	Prognosis OR Survival Analysis	49%	97%
Strategy 2	Incidence OR exp Mortality OR Follow-up Studies OR Mortality (sh) OR Prognosis: (tw) or Predict: (tw) OR Course (tw)	92%	73%
Review[d]			
Strategy 1	Meta-analysis (pt) OR [review (pt) AND MEDLINE (tw)]		
Strategy 2	Meta-analysis (pt) OR Overview (tw) OR Meta-analysis (tw) OR Meta-analysis (tw) OR [review (pt) AND MEDLINE (tw)]		

[a] See Appendix 2 for key to MEDLINE abbreviations.
[b] Based upon work of Haynes et. al., where search terms were compared to manual searches (gold standard).
[c] All terms not otherwise labeled are MeSH terms.
[d] Comparison with gold standard not yet performed.

APPENDIX 2
KEY TO MEDLINE ABBREVIATIONS

Term	Meaning
default	Medical Subject Heading (MeSH)
sh	MeSH subheading
pt	Publication type
exp	Explode MeSH term
px	Pre-explode MeSH term
*	Focus MeSH term
/	Signifies that MeSH subheadings follow
tw	Text word
&	Wild character, replaces one character or none
:	Wild character, replaces none, one, or many