Searching the medical literature

G Gore

Practice makes perfect

Searching the medical literature for information can be quite confusing without a basic understanding of how the literature is organized or indexed, and without guidance about how the most popular search systems work. In this commentary, I will use Medline to illustrate some basic principles that are widely applicable, but the sad fact is that each database follows different rules—especially in subject searching.

INDEXED DATABASES

What is an indexed bibliographic database? This is a question authors may want to answer because journal articles are only useful to other researchers if they can be retrieved with relative ease. A bibliographic database is a structured collection of descriptive information used to identify publications, such as journal articles. This descriptive information is organized, or indexed, into searchable fields (such as author, title, source, or subject). In light of the heterogeneity in the arena of injury prevention, researchers may need to become familiar with several of these databases—for example, CINAHL, ERIC, PsycInfo, or PAIS International. Most health-oriented investigators will certainly need to know how to navigate the popular medical giants—Medline, Current Contents, and Embase.

MEDLINE

When Medline, the online database counterpart of Index Medicus (produced by the National Library of Medicine in the USA), began indexing the contents of Injury Prevention, this meant that a reference to every paper published in this journal could be found in the database. The most popular sections of an article that Medline indexes are the author’s name, the title of the article, the source, the subject of the article, and text words from a combination of fields (often including the abstract, but not the full text). Expert indexers read each such paper to identify key search terms.

OTHER INDEXES OR DATABASES

Two other important indexes covering the biomedical literature are Current Contents (a product of ISI, a Thomson Company) and Embase (a product of Elsevier). These mammoth databases include journals using different selection criteria. As Pless notes, Injury Prevention has just made the giant leap into Current Contents, from which researchers will be able to retrieve references to articles published here along with tables of contents. Injury Prevention is not currently indexed in Embase but does appear in CINAHL, the Cumulative Index to Nursing and Allied Health. Articles from the journal are indexed in a particular database depending on their relevance to the database’s subject coverage. This means that only articles in Injury Prevention relevant to the field of nursing and/or allied health will appear in CINAHL.

If you are confused about bibliographic indexes, you are not alone. A common problem, even for librarians, is figuring out which indexes cover a particular journal. One free method of checking this is through JAKE (jointly administered knowledge environment), at http://jake.lib.sfu.ca. It is not perfect, but it will often give you a good indication of the indexes you should search for references to articles from a given journal.

SEARCH SYSTEMS (FINDING THINGS)

There are several methods used to find articles on a given topic. Browsing, or leafing through print issues, is limited because it is now impossible to cover the sufficient amount of the relevant literature by this method. Another time honoured method is selecting references from articles already at hand. It may be honoured by time but it is not well regarded by reviewers and editors. This method should never be relied upon for many reasons—not least of which that references will be at least as old if not older than the article from which they are being retrieved.

Today most researchers complement these methods with computer mediated searches. To do so efficiently they need to understand the structure of the most popular indexes and appreciate, for example, the difference between Medline and Current Contents and how each is best searched.

Medline can be accessed through a variety of subscription based search systems (for example, Ovid Online) but PubMed is the most popular—and not simply because it is free. PubMed is available to any researcher with an internet connection. It is a world wide web platform (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi). Developed by the National Center for Biotechnology Information at the National Library of Medicine, PubMed has been making Medline freely available to the international community since 1997—an impressive act of generosity when one considers the cost of subscribing to other search systems. Remember, PubMed is the search system, Medline a database “behind” it.

MEDICAL SUBJECT HEADINGS (MeSH)

Before we look at the search system, the reader may need a primer on MeSH (an acronym for medical subject headings), the controlled vocabulary thesaurus used by the National Library of Medicine. MeSH creates a standardized use of language in an attempt to unify differences in terminology into single concepts. It is hierarchically organized in a tree structure that allows for varying levels of subject specificity. The National Library of Medicine employs a specialized staff of indexers who read each article and who then choose and assign each one the most specific and appropriate MeSH terms. These terms, a mystery to many researchers, are commonly used by health sciences librarians. The 21,973 available can be made more precise by combining them with subheadings or qualifiers to further describe a particular aspect of a MeSH concept.

For example, the MeSH term “injuries and wounds” can be combined with the subheading “prevention and control”. This heading/subheading combination ends up sounding much like another MeSH term, the infamous “accident prevention”, and illustrates one of the limitations of controlled vocabularies. Admittedly imperfect, an exhaustive search will require that you use many different search methods and this means becoming familiar with MeSH.

Because of the confusion that often arises when using MeSH terms, many researchers prefer to use “natural language” vocabulary. For example, they may enter the term “injury prevention” into the general Pubmed query box. This involves a leap of faith: the researcher is assuming (or hoping) that articles dealing with this subject will contain this sequence of words within at least one of the fields that the system searches. This makes it important for authors to include terminology—when creating a title or writing an abstract—that will make retrieval of the article as easy and as obvious as possible for fellow researchers.

www.injuryprevention.com
To be efficient, researchers also need to become skilled with the search system they are using—in this case, PubMed. This system looks simple enough on the surface but is actually quite sophisticated. Hence, the more you play with it, the more skilled you will become. Explore using the “limits” (for example, to limit articles by language, publication date, human or animal). Get used to performing your searches in the “preview/index” mode: this allows you to see how many results your query has generated, and offers more advanced features than the default screen. It is also a great way to explore MeSH terms associated with your query—or to see whether your term is even listed in MeSH (enter your term in the lower query box, select MeSH terms, and click on “index”). You can also use the MeSH database to explore the vocabulary, a text link available on the left side of the screen. Make use of the “history” mode, which reviews up to 100 of the last queries. The “clipboard” allows you to see a list of the references you have checked off. Be sure to take the tutorial.

Enter search terms or concepts in the smallest units possible: use the lower query box in the “preview/index” mode to search specific fields listed in the “all fields” menu, being sure to click on the “clear” button near the upper box between queries. Combine your query results using the Boolean operators AND, OR, NOT, for example (#1 OR #2) AND (#3 OR #4)—remember to use the “history” mode to view a complete list of your queries. Use parentheses to override the default order (that is, left to right in PubMed) in the execution of these operators. Use the Boolean operator OR to combine synonymous concepts before combining additional concepts with AND. Remember to type the Boolean operators in upper case. Avoid using NOT: it may eliminate relevant “hits” from your search.

Once you’ve found relevant hits, check them off and save them in the clipboard—and keep searching. Up to 500 hits will remain in the clipboard for up to one hour of inactivity. You can also display a retrieved reference in Medline format (when viewing a record, choose Medline from the “display” menu, then click on the “display” button), to see what MeSH terms (MH in the Medline display format) were used to index it. This may give you new ideas on MeSH terms to search.

Finally, make sure to review what field PubMed has searched in your query. To do so, click on the “details” tab. For example, if you enter “injury prevention” (without quotations) into the query box, click on “preview”, and then on the “details” tab: you will discover that the system has only searched the [journal] field! If you would like to know more about why this happened, go to PubMed Help. One last tip: to save your search history, click on the URL button in the “details” mode.

To sum up: knowing how to search is more than half the battle. The other half is knowing what to search, that is, which databases. In our field, there are fortunately—or unfortunately—many options. These options reflect the multidisciplinary nature of injury prevention and may require ranging as far afield as education, psychology, or political science. The best rule is practice makes perfect . . . do it yourself, do it often, and don’t depend on research assistants alone!

Injury Prevention 2003;9:103–104

Author’s affiliation
G Gore, McGill University, Montreal, Canada

Correspondence to: Ms Gore;
genevieve.gore@mail.mcgill.ca

REFERENCES