

How to search medical databases: part 2

John Eyers unravels more of the mysteries



In the last issue we discussed the most common way of searching databases, called ‘natural language’ searching, using keywords found in the title or abstract of papers. We now turn to the more difficult, but more effective, search method called ‘controlled language’ or ‘thesaurus’ searching. Some databases, such as Pubmed/Medline, assign index or thesaurus terms to each article, which greatly assists retrieval from databases, and if used properly, can result in fewer important papers being missed. Those conducting systematic reviews will use thesaurus searching, as well as natural language searching, as a matter of course to ensure that as many relevant papers as possible are found for review.

To illustrate this, I shall take the example of Pubmed/Medline, which has a thesaurus of index terms called MeSH (Medical Subject Headings). This is now such a powerful retrieval tool that it has been adopted by other databases such as the Cochrane Library, the evidence-based medicine database. The MeSH terms can be searched to identify the unique subject heading for a particular concept, and includes categories of terms such as diseases, organisms, anatomical terms, procedures, drugs, as well as terms such as public health, child, adult, infant, medical personnel, geographical areas, and countries as well as other disciplines related to health.

MeSH also has a system called Tree Structures in which all these terms are placed in a listing of broader and narrower terms or in a genus/species relationship to one another. An example of part of a tree listing is given in Figure 1, in which *malaria* has been selected and categorised as a *protozoan infection* which in turn is a *parasitic disease*. Further down the branch of the tree *malaria* has been subdivided into types of malaria with *malaria, falciparum* being further subdivided. This allows the searcher to move up or down the

Figure 1

Parasitic Diseases
Protozoan Infections
Malaria
Malaria, Avian
Malaria, Cerebral
Malaria, Falciparum
Blackwater Fever
Malaria Cerebral
Malaria, Vivax

branch of the tree to broaden or narrow a whole category of term. So in Figure 1, if we clicked on *protozoan infections* it would include not only *malaria*, but also *Trypanosomiasis*, *Amebiasis*, and *Trichomoniasis*, among others, without having to type them in individually as you would with natural language searching. Note that US spelling (amebiasis not amoebiasis) is used in MeSH. Another advantage to the tree structures is the ability to search across geographic areas without the need to type into the search equation each country within that area – see Figure 2 for the MeSH listing of Africa, Eastern. Within MeSH, it is also possible to identify more detailed aspects of a subject heading using subheadings – for malaria you might only be interested in *drug therapy, epidemiology, prevention and control* subheadings, and as in a natural language search the result can be limited by language, type of publication (review, clinical trial, practice guideline etc), gender (male/female), species (animal/human).

branches of the tree to broaden or narrow a whole category of term. So in Figure 1, if we clicked on *protozoan infections* it would include not only *malaria*, but also *Trypanosomiasis*, *Amebiasis*, and *Trichomoniasis*, among others, without having to type them in individually as you would with natural language searching. Note that US spelling (amebiasis not amoebiasis) is used in MeSH. Another advantage to the tree structures is the ability to search across geographic areas without the need to type into the search equation each country within that area – see Figure 2 for the MeSH listing of Africa, Eastern. Within MeSH, it is also possible to identify more detailed aspects of a subject heading using subheadings – for malaria you might only be interested in *drug therapy, epidemiology, prevention and control* subheadings, and as in a natural language search the result can be limited by language, type of publication (review, clinical trial, practice guideline etc), gender (male/female), species (animal/human).

Figure 2

Africa
Africa, South of the Sahara
Africa, Eastern
Burundi
Djibouti
Eritrea
Ethiopia
Kenya
Rwanda
Somalia
Sudan
Tanzania
Uganda

result can be limited by language, type of publication (review, clinical trial, practice guideline etc), gender (male/female), species (animal/human).

So, to search for malaria in children, we would identify the MeSH term for malaria and its narrower terms in the tree structure as in Figure 1, deciding whether to ‘explode’ the malaria tree to retrieve *vivax malaria, falciparum malaria* and *cerebral malaria*. Subheadings and other limits (as explained above) can also be applied. The MeSH term for *child* (defined as 6–12 years) has also *child, preschool* (defined as 2–5 years) under it in the tree. Further categories of *infant* (birth–23 months) and *adolescent* (13–18 years) are also available for inclusion. Once results for each of the search sets are available, they can then be combined together using the OR/AND operators as appropriate (as explained in the January issue) for the required result.

As a member of HIFA2015 (<http://www.hifa2015.org/>), John Eyers believes one of the biggest challenges for improved health in less-developed countries is ensuring that equal access to appropriate and evidence-based health information is made available to all.

John would like to have feedback from readers and in particular suggestions for further topics for discussion. Email him on: johneyers@hotmail.com.

References

1. Hinari website training modules – 4.3 Pubmed MeSH & Workbook <http://who.int/hinari/training/modulesfour/en/index.html>. Accessed 25th Feb 2010.
2. US National Library of Medicine. Online Tutorial – How to Use MeSH. <http://www.nlm.nih.gov/bsd/viewlet/mesh/searching/mesh1.html>. Accessed 25th Feb 2010.
3. Eyers JE. (1998). Searching bibliographic databases effectively. *Health Policy Plann* 1998; 13: 339–42. Free to download at <http://heapol.oxfordjournals.org/cgi/reprint/13/3/339.pdf>.
4. Holzer L, Eyers J. Bibliographic databases. *Student BMJ* 2008; 16: 366–7. Free to download at <http://archive.student.bmj.com/issues/08/10/education/366.php>.