4. Searching databases – basic guidelines

How do we locate research studies and abstracts?

The following is a very brief introduction to the basic principles of searching databases. For practitioners interested in improving their skills, a specialist course is recommended. These may be arranged through:

http://evidencenetwork.org/Training02.html

Search strategies

The best approach for anyone wishing to formulate a simple search strategy is to generate key words or concepts derived from their search question. Questions will be broken down into phrases and phrases into words. These are often referred to as 'search terms'. The search engine of the database will scan the title and/or abstract for these terms, depending on which option is selected.

An example

Search question: *How effective are work placements for people with mental health problems as a route to permanent employment?*

Key concepts in the phrase:
Subject: *people with mental health problems*
Topic: *work*
Outcome: *permanent employment*

Search terms: *work*, *mental health, employment.*

We will also want to consider using other associated terms: *training*, *vocation, job, mental illness, mentally ill.* This is particularly important if we are searching international literature, where different terminology may be used.

Databases allow their content to be searched in two ways: either by key words, which have been catalogued as such and are recognised by the programme; or by free text, where we are able to insert whatever terms we think are relevant. Every database has its own conventions for searching and time spent becoming acquainted with their specific idiosyncrasies is well worth it. Therefore, always check the help section of a new database. Nonetheless, there are a few simple search rules that most database search engines have in common.

Search engines enable us to connect terms using joining words or symbols. These are known as Boolean operators – AND, OR and NOT. For example, connecting the words with AND (**work AND mental health**) will only locate items which feature *both* terms. The term AND, therefore, narrows our search. Connecting the words with OR (**work OR mental health**) will locate items that contain *either* term, broadening our search. Searching with the term NOT will narrow the search, so **work NOT mental health** will return all items that concern work *except* those that feature mental health. However, caution should be exercised when using the NOT operator; by excluding a certain category of material, we could easily inadvertently miss some documents that have useful information on the topic we are researching. For example, **work NOT mental health** may be a helpful search phrase if we are interested in employment but not interested in mental health aspects of it such as stress or depression; however, we may inadvertently exclude relevant research on those working in the mental health sector.

Another important operator is the truncation or wildcard. This is a symbol, usually an asterisk, which substitutes for a suffix. For example, **child*** will locate records that contain the word child, but also children; **therap*** will locate therapy and therapeutic; **depress*** will find depression and depressive, and so on. Asterisks can often also be used as a prefix; ***abuse**, for example, would

locate child abuse, elder abuse and financial abuse. Some databases use the dollar sign (e.g. child$) instead of an asterisk. Many databases allow us to use 'variant spelling', which would locate different spellings of the same word, e.g. organisation, and organization. A question mark is often used for this, i.e. **organi?ation**. The help page of a database will advise on the appropriate truncation symbol.

Phrasing is also a valuable strategy. This means an entire phrase rather than individual terms are located. For example, **"community care"** would ensure that items which contain this phrase, rather than the words "community" and "care" separately, are returned. This is an important searching tool, not only for databases but also for general search engines such as Google.

Advanced techniques for database searching

For more complex searches, many databases allow us to use parentheses to structure our search. Using brackets allows us to group terms and avoid ambiguity. For example, **"social work" AND (evidence OR research)** will search for documents containing either social work and *evidence*, or social work and *research*. A more complicated search string might look like this: **("social work" AND (evidence OR research)) NOT case study**. The technique where one parenthesis is used inside another is referred to as 'nesting'. 'Proximity searching' is another way of narrowing a search. This enables us to retrieve keywords within a certain distance of each other. Different databases use different proximity operators. Web of Science[3], for instance (a database covering science, social science and health topics), uses the word 'same' – e.g. the phrase **"social work" same research**, retrieves records where the two keywords appear in the same sentence. Others use the word 'near' to achieve a similar thing. **"Social work" near evidence** will locate records where the two keywords appear no more than 10 words apart.

[3]www.isinet.com/products/citation/wos/

Refining your search

Having conducted our initial database search, we may find that we have a) too little information, or b) too much information. In either case, we may have to come up with alternative search terms related to our original ones, as in the example above of work, employment, job and so on. When we have too little information, we may have to broaden our search by using the OR operator described above. When we have too much information, we may have to limit our search using AND or NOT. Alternatively, we may wish to limit our search by language, date or study type (e.g. only systematic reviews or randomised controlled trials). Most databases provide these 'limiters'. In a well-researched field we may have the luxury of choosing between many studies, in which case we should choose the best available research. By best, we mean research which is well-conducted and which uses bias-reducing methodologies.

Limiting searches: sensitivity and specificity

So we now know that databases enable the user, with differing degrees of sophistication, to place limits on a search. The purpose of limiting a search is to locate the largest number of relevant results while minimising the irrelevant ones. Different databases vary in their ability to deliver the maximum number of 'hits' and the minimum number of 'misses'. These two dimensions are referred to as sensitivity and specificity.

Sensitivity describes the proportion of relevant items which are located by a search. While all searches need high sensitivity, this dimension alone is insufficient. We must also strive for high *specificity*, which is an expression of the ratio between 'hits' and misses'. The importance of this should not be underestimated. A search with low sensitivity will simply not return the information we need, or worse, will return incomplete information leading us to make erroneous conclusions. A search with low specificity may return many hundreds, or even thousands, of items, most of which will be irrelevant, but which will all have to be examined before being discarded – often a very lengthy and arduous task. As we are unlikely to know the total number of relevant items in the 'universe' which we are examining in advance, we measure the sensitivity and specificity of a particular database by comparing it with the combined performance of all relevant databases.

No search strategy is likely to achieve 100 per cent sensitivity and specificity. However, different databases, depending on the subject, vary in their degree of sensitivity and specificity, and it is important for practitioners to be aware of the ones that are more likely to produce high scores in their particular speciality.