Searching for information on outcomes: do you need to be comprehensive?

Alison J Brettle, Andrew F Long, Maria J Grant, Joanne Greenhalgh

Abstract
The concepts of evidence-based practice and clinical effectiveness are reliant on up to date, accurate, high quality, and relevant information. Although this information can be obtained from a range of sources, computerised databases such as MEDLINE offer a fast, effective means of bringing up to date information to clinicians, as well as health service and information professionals. Common problems when searching for information from databases include missing important relevant papers or retrieving too much information. Effective search strategies are therefore necessary to retrieve a manageable amount of relevant information. This paper presents a range of strategies which can be used to locate information on MEDLINE efficiently and effectively.
(*Quality in Health Care* 1998;7:163–167)

The concepts of evidence based practice and clinical effectiveness are reliant on up to date, accurate, high quality and relevant information. Information is needed on which interventions work best, under what conditions, and on which appropriate outcome criteria, indicators, and measures to use to monitor the achievement of desired outcomes.[1] This type of information can be obtained from a range of sources, but as healthcare practice is a rapidly evolving field, journal articles offer the best means of keeping up to date with current developments. In recent years computerised databases such as MEDLINE have offered a fast and effective means of accessing the information published in a proliferating number of medical and healthcare journals. Traditionally searches of such databases have been carried out by librarians and information professionals. However, cheaper computers, CD-ROM technology, easier access, and user friendly software interfaces have ensured that clinicians and other health service professionals can use databases to quickly retrieve relevant information themselves.

Although searching databases can be easily and effectively carried out by people who are not information professionals, common problems are missing relevant papers or retrieving too many. Missing potentially relevant studies is a crucial problem if a decision is to be based on the evidence from a literature search. On the other hand retrieving too many irrelevant studies can be frustrating for those with little time to synthesise a large amount of information.

Also each database searcher has different preferences and needs depending on the reason for carrying out the search. For example, a searcher carrying out a systematic review would need to find and synthesise a large amount of studies to ensure that no potentially relevant studies are excluded from the review. A searcher looking for background information, however, is more likely to need a few pertinent references to satisfy the enquiry.

Effective search strategies are therefore needed to locate relevant, high quality studies, suited to the needs of the enquirer. Strategies can be constructed to identify a general concept related to clinical effectiveness—such as, randomised controlled trials, systematic reviews, or outcome measures—and can then be combined with search terms representing the condition of interest. These are known as methodological filters (box) and have been recommended as a means of providing busy clinicians with an effective means of retrieving high quality, relevant information.[2] Methodological filters have been developed and published on a range of concepts related to evidence-based practice.[3][4] This paper describes the development and discusses the use of three filters that can be used to retrieve articles relating to outcome measurement from MEDLINE. The strategies were tested on three areas (rheumatoid arthritis, hip replacement, and incontinence) and the results compared against a handsearch of two journals. The results obtained raised several issues, and will be used as a basis for search recommendations later in the paper.

Information on outcome measurement is essential to clinicians, health service professionals, and researchers to determine what outcome measures are, how they can be used in research or clinical practice, and their appropriateness and feasibility for use in a particular area. The strategies reflect the needs of different users. For example, a short strategy which is quick and easy to run will retrieve a high proportion of relevant material, but may miss some references. At the other end of the scale is a strategy which will retrieve a lot of information and is less likely to miss references, but takes considerably longer to run. The short strategy is useful to busy clinicians and the long one to researchers who require a comprehensive search of the literature. The strategies were developed as part of the work of the United Kingdom Clearing House on Health Outcomes (subsequently referred to as the clearing house), an initiative funded by the Department of Health 1992–6.

Health Care Practice Research and Development Unit, University of Salford, UK
Alison J Brettle, research fellow (information)
Andrew F Long, director

Centre for Health Information Quality, Highcroft, Romsey Road, Winchester, UK
Maria J Grant, information officer

Nuffield Institute for Health, Clarendon Road, Leeds, UK
Joanne Greenhalgh, MRC training fellow in health service research

Correspondence to:
Alison Brettle, Health Care Practice Research and Development Unit, University of Salford, 5th Floor, Humphrey Booth House, The Crescent, Salford M5 4QA, UK. Telephone 0044 161 295 2823; fax 0044 113 295 2818; email A.Brettle@ health-faculty.salford.ac.uk

Accepted for publication 26 June 1998

Terminology

MESH (medical subject headings)=list of terms arranged in a hierarchical structure used to index records on MEDLINE

Methodological filter=search strategy constructed to identify a concept such as outcome measures or randomised controlled trials and then combined with a separate strategy of terms representing the condition of interest. The filter can be used on any number of occasions, changing the condition area as appropriate.

OVID=Name of software used to search MEDLINE.

Precision=The proportion of citations in a given search that are relevant to the search question. It is used to assess the results of search strategies.

Recall=The proportion of relevant citations in a given search among the total number of relevant citations. It is used to assess the results of search strategies.

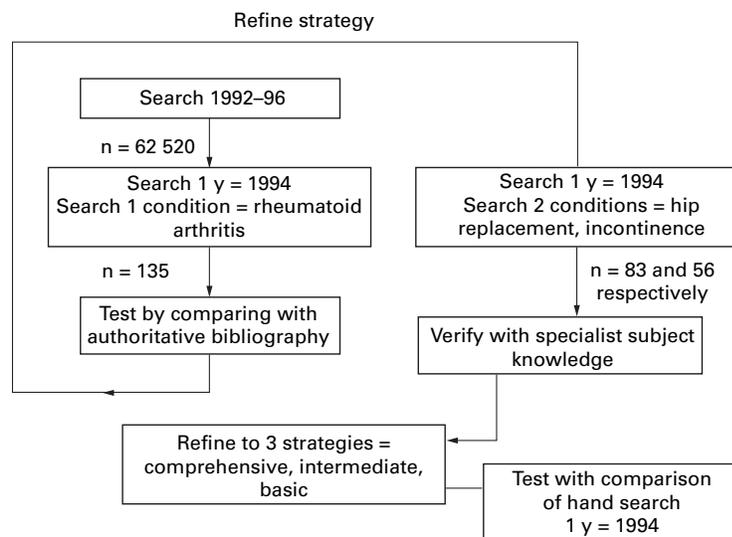### Developing the search strategies

The three strategies were developed in an iterative manner with OVID software searching between 1992 and 1996 (figure). The concepts of recall (the proportion of relevant citations in a given search relative to the total number of relevant citations) and precision (the proportion of citations in a given search that are relevant to the search question)[5] have been used to assess the results of database searches (box). These are sometimes equated to the concepts of sensitivity and specificity, respectively.

Effective search strategies usually make use of "free text" terms of text words as well as thesaurus terms (sometimes known as keywords or in the case of MEDLINE, MESH terms (box)). Initially the thesaurus function and automatic mapping provision were used to identify relevant MESH terms. These were included in a strategy and free text terms added to increase the recall of the search results. The

free text terms included a list of synonyms of relevant terminology identified by the authors. This general strategy yielded a high number of articles (n=62520) when the whole database was searched from 1992 onwards (the earliest year at that time). The search was rerun restricting the year to 1994 (to ensure all articles for a particular year had been included) and to a single topic area of rheumatoid arthritis (a subject the clearing house was exploring at that time). This yielded a more manageable number of articles (135) which enabled testing of the strategy. The abstracts were used to assess the relevance of the results which were compared with a printed bibliography.[6] Several items were missing. These were found with an author search which identified further terms for inclusion in the strategy. The search was then rerun.

To test whether the search retrieved relevant items for a particular year, two journals were handsearched and the results compared with those of the database search. The two journals selected for handsearching were *The Lancet* (to represent United Kingdom studies) and *Medical Care* (a journal known to the authors to regularly include articles relating to outcome measurement and to represent studies from the United States). Further tests were made by searching for named measures. The measures were entered as text words and the items retrieved were compared with the results of the search. To verify whether the search was relevant for other medical areas, it was rerun with suitable MESH headings for hip replacement and incontinence (further areas of interest to the clearing house team at that time). The results of all search stratigies were checked by research and information staff within the clearing house team to verify that the results were relevant and that items known from experience were included.

The resulting full search strategy was detailed and long (29 items) (table 1). Although it maximised recall (see results) and was suitable when reviewing the full range of measures available for a particular medical area, it took a long time to run and retrieved many references. Thus it was not necessarily suitable for end users who did not need to retrieve every possible article relating to outcome measurement in a particular area or did not have time to synthesise the results. It was therefore decided to refine the search strategy to produce further strategies which were more precise (at the expense of recall). Shorter strategies were generated by reviewing the search results to identify the MESH headings and combinations of text words used most often. Terms were tested in different combinations against the results of the original strategy in the three areas. Two further strategies were derived: one which uses mainly MESH headings and is therefore precise—but at the expense of recall; and the other which aims to provide a balance between these two approaches. The strategies were named basic (most precise), intermediate, and comprehensive (least precise, but with highest recall).

Process of development of the search strategy.

Table 1 Presentation of three search strategies for retrieving information about outcome measurement from MEDLINE

Basic	Intermediate	Comprehensive
1 health status indicators/	1 health status indicators/	1 health status indicators/
2 outcome and process assessment (health care)/	2 outcome and process assessment (health care)/	2 outcome and process assessment (health care)/
3 outcome assessment (health care)/	3 outcome assessment (health care)/	3 outcome assessment (health care)/
4 quality of life/	4 quality of life/	4 quality of life/
5 outcome measure$.tw.	5 outcome measure$.tw.	**5 health status/**
6 health outcome$.tw.	6 health outcome$.tw.	**6 severity of illness index/**
7 1 or 2 or 3 or 4 or 5 or 6	**7 quality of life.tw.**	**7 self assessment [psychology]**
8 your search term(s) (subject specific)	**8 measure$.tw.**	8 outcome measure$.tw.
9 7 and 8	**9 assess$.tw.**	9 health outcome$.tw.
	10 (score$ or scoring).tw.	10 quality of life.tw.
	11 index.tw.	**11 health status.tw.**
	12 indices.tw.	**12 (endpoint$ or end point$ or end-point$).tw.**
	13 scale$.tw.	**13 (self-report$ or self report$).tw.**
	14 monitor$.tw.	**14 functional outcome$.tw.**
	15 8 or 9 or 10 or 11 or 12 or 13 or 14	**15 outcome$.ti.**
	16 outcome$.tw.	16 or/1–15
	17 (outcome$ adj3 (measure$ or assess$ or (score$ or scoring) or index or indices or scale$ or monitor$)).tw.	17 outcome$.tw.
	18 or/1–7	
	19 17 or 18	18 measure$.tw.
	20 your search term(s) (subject specific)	19 assess$.tw.
	21 19 and 20	20 (score$ or scoring).tw.
		21 index.tw.
		22 indices.tw.
		23 scale$.tw.
		24 monitor$.tw.
		25 or /18–24
		26 17 and 25
		27 16 or 26
		28 your search term(s) (subject specific)
		29 27 and 28

ti=word in title; tw=text word (word in title or abstract); /=MESH heading; $=truncation symbol; adj (number)=ajacent by (x number of words).
Software: the searches have been developed for use with OVID software. An explanation of abbreviations and search terms is given above. If the search is run on different software, such as Silverplatter, translation of the search strategy will be necessary. Help from an information professional should be sought before doing this.

Presentation of the strategies

The three search strategies are presented in table 1. The detail of the comprehensive strategy is self evident and makes use of a range of synonyms and combination of text words and MESH terms to describe outcome measurement. In comparison the basic strategy uses more obvious MESH terms and text words—for example, outcome measurement. The text has been emboldened to show the terms added at each stage.

Recommendations for using the strategies

Each enquirer carrying out a search has different preferences and information needs. This suggests the need for a subjective criterion of manageability as well as the traditional ones of recall and precision as a way to assess search strategies and as a basis for search recommendations. Some—for example, research staff—prefer to retrieve many references to maximise recall and then browse through manually,

whereas others—for example, busy clinicians—want a few pertinent references, preferring to maximise precision. The criterion of manageability relates specifically to the preferred balance between the amount of material retrieved and the intended use of the information. It is thus proposed as a working definition that a successful strategy is one that retrieves a manageable number of relevant references, providing a user specified balance of recall and precision. The range of search strategies presented here offers potential users different ways to achieve this balance.

The results of the tests (tables 2 and 3) show that higher recall is balanced by lower precision and vice versa (as would be expected). For recall (table 2), the comprehensive strategy is optimal, retrieving as many articles as possible. When compared with the hand search of two journals, recall was 100%. The search results also included all the articles retrieved when searching with terms to describe named measures. However many of the articles were irrelevant (52%, table 3).

The comprehensive strategy should therefore be used when full and detailed information is required about the range of outcome measures which have been used in a particular area. It must also be used when it is suspected that there has been little published work on outcomes measurement in an area, as it retrieves articles which have been insufficiently indexed. In contrast, the basic strategy misses a considerable number of articles (53%), but is best for precision, with most articles retrieved being relevant to the search question (69%). It should be used when interest lies in gauging the number of publications available for a topic area, or when a few articles are required, but it is not necessary to retrieve information about

Table 2 Recall for the three search strategies (%)

Strategy	Rheumatoid arthritis	Hip replacement	Incontinence	Mean
Basic	55	60	56	57
Intermediate	66	72	80	73
Comprehensive	100*	100*	100*	100*

*Accurate for two journals.
Recall for the basic and intermediate strategies are calculated against the results of the comprehensive strategy.

Table 3 Precision for the three search strategies (%)

Strategy	Rheumatoid arthritis	Hip replacement	Incontinence	Mean
Basic	73	65	69	69
Intermediate	67	74	55	65
Comprehensive	48	52	45	48

every outcome measure which has been used in a topic area. The intermediate strategy provides a balance between the basic and comprehensive strategies in both recall and precision. It is useful for finding information about the range of measures which may be used in a particular area, but where it is not essential that absolutely every paper is retrieved.

If it is not essential that all potentially relevant papers are retrieved, the strategies can be used in an incremental manner. For example, the user could begin with the basic strategy, if this retrieves sufficient information to answer the search question, then no further searching may be necessary. If however the question has not been answered or the user suspects that some papers have not been retrieved, the additional steps in the intermediate strategy could be added, and then if necessary the steps from the comprehensive strategy.

Points to bear in mind when using the strategies

Although searching electronic databases such as MEDLINE is a fast and efficient way of retrieving current information, it is a largely subjective process. Search strategies can never be perfect and will never retrieve absolutely every relevant reference. Databases are constantly developing. New records are added on a regular basis and concepts and terminology change over time. Therefore search strategies and methodological filters must also change over time to reflect these differences.

In this study, the strategies were tested on three medical areas, and although this suggests that they are useful and relevant for other areas, the same values for recall and precision cannot be guaranteed. For some areas, there may be further relevant outcome related terminology which could be added to the filter. Also the gold standard test of a comparison of the strategy against a hand search of journal articles was only carried out against two journals, not a wide range of journals, due to limited resources. The actual sensitivity of the search cannot be accurately calculated.

When many references are retrieved or even an acceptable number of relevant references are retrieved, it is easy to think that all relevant material on the subject area in question has been located. This is not always the case. Unless the searcher is an expert in the topic area, it is impossible to know what material has been missed. This suggests the importance of collaboration between an information specialist and the enquirer or subject specialist, as was done in this study, to ensure that appropriate references are retrieved. This applies as much to searches for outcome measures as to searches for randomised controlled trials,[3] evidence on effectiveness generally, or in a particular area such as rheumatoid arthritis,[7] or complementary treatments.[8]

The study highlighted limitations of searching MEDLINE for information on outcome measurement and more generally clinical effectiveness. In some cases there are no abstracts for articles, making it difficult to decide on the relevance of an article and

increasing the reliance on correct indexing of articles to aid retrieval. In other cases, material has been incorrectly indexed making it impossible to retrieve some articles. Studies may also have been developed for a purpose other than outcome measurement—and as such papers describing their development and psychometric testing, may not include the relevant MESH terms. Furthermore, the (common) names of measures are not always listed in the abstract, use being made of a general term—such as quality of life or health status measure—and different labels are given for the same thing (for example, SF-36 or SF36 or MOS short form). Finally, much of the information held on MEDLINE relates to the use of outcome measures within research studies and may not be relevant to routine practice, an essential area of interest to many clinicians, audit facilitators, and purchasers.

It is unlikely that the limitations identified in searching for outcome measures on MEDLINE are unique to this database. Initial work on CINAHL[9] showed similar limitations, together with an additional difficulty over the quality and lack of abstracts, preventing rigorous testing of strategies. Further work is needed in this area and could also be applied to other databases—such as EMBASE—and other non-medical databases—such as PSYCHLIT and SOCIO-FILE. Our own work is currently being extended to examine searching on a range of databases in the field of social care.

The problem of incorrect indexing has been previously highlighted relative to finding randomised controlled trials for systematic reviews.[3] In this case the United States National Library of Medicine (the producers of MEDLINE) have taken steps to rectify the problems. Similar work is needed for the indexing of articles related to outcome measures. Alongside this, journals could encourage authors of articles to list all the outcome measures used in the study in abstracts of articles and to use the common and recognised names for the methods of measuring outcomes used.

Conclusion

When improving healthcare quality, decisions should be made on the basis of sound research evidence while using appropriate outcome criteria and measures (effects). Finding this evidence is problematic. Methodological filters are good for clinicians, health service professionals, and information professionals to ensure that they search databases effectively and efficiently. However, no search strategy is perfect and those carrying out the search should ensure that they are trained in search techniques and know the database sufficiently well to know its limitations and how to get the best out of it. One method of improving searches is for a clinician to search alongside an information professional. This way clinical and technical knowledge can be combined to achieve effective results. Searchers need to ensure that they have clarified their search question before starting the search and be aware of what the information retrieved will be used for. By doing this the user can obtain relevant

information efficiently and effectively by deciding whether to take a more basic or comprehensive approach.

1 Long AF, Fairfield G. Confusion of levels in monitoring outcomes and/or process. *Lancet* 1996;**347**:1572.
2 Haynes RB, Wilczynski N, McKibbon A, *et al.* Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994;**1**:447–58.
3 Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;**309**:1286–91.
4 Harrison J. Designing a strategy to identify and retrieve articles on evidence-based health care using MEDLINE. *Health Libraries Review* 1997;**14**:33–42.
5 McKibbon KA, Walker-Dilks CJ. The quality and impact of MEDLINE searches performed by end users. *Health Libraries Review* 1995;**12**:91–200.
6 Berzon RA, Simeon GP, Simpson RL, *et al.* Quality of life bibliography and indexes: 1993 update. *Qual Life Res* 1995;**4**:53–73.
7 Brettle A. Searching the literature. In: Long AF, Scott D, eds. *Measuring outcomes in rheumatoid arthritis*. London: Royal College of Physicians, 1996.
8 Long AF, Brettle A, Mercer G. *Searching the literature on the efficacy and effectiveness of complementary therapies*. Leeds: University of Leeds, Nuffield Institute for Health, Yorkshire Collaborating Centre for Health Services Research 1996.
9 Brettle A, Grant MJ. Searching the literature for information on outcome measurement: a guide to searching MEDLINE and CINAHL. *Outcomes Briefing* 1996; **8**:44–8.