An Insider's Insight into Literature Searches

Searching the literature can take various forms, ranging from a quick scan of recent publications to a formal, systematic interrogation of all available data sources to establish the scientific consensus on a specific topic. In these days of online journal databases, the relative ease of conducting a search means that they often start informally with no thought-out search strategy or defined goal.

A long list of articles can be generated almost instantaneously, but what did you miss and how long will it take to review the data? How easily can the search strategy be repeated and adapted to obtain a more complete and refined set of references? We offer some insights from the Niche medical writing team who have been conducting literature searches for their clients since 1998.

Before you start

- Establish a formal plan for your literature search if you propose to do anything more than conduct a cursory review of the literature

- Know what you want to achieve so you avoid endless futile or repetitive searching. Set yourself an objective and identify an endpoint that qualifies whether or not you have achieved your goal

- Ensure you have access to an appropriate search engine as different databases will provide different results for the same search strategy. Establish whether you need access to more than one system to achieve your goals

Prepare to succeed

- Your searches will create outputs in the form of lists of publications. Decide what information you need to record about each reference in order to help determine its relevance, how you will store the information and how you will 'score' the overall efficacy of a search strategy

- Know something about your subject area before you decide on the operational parameters of your search; consider the coverage history in the literature, controversies, specialist journals, sub categories, etc.

- Plan to extract surgically what you need rather than boiling the oceans dry to find it

You can burn up a great deal of time and energy searching the literature. Your search's quality and value is wholly dependent on the thought and effort you put into developing your strategy. It also determines the effort and subjectivity needed to sort through the list of references it produces. As searches are often an iterative process, pre-defining your goals and strategy allows you to record your starting point and review the decision tree you finally use to select what you consider to be the most appropriate references.

Key Insights

If you work in biomedical sciences it is very likely that you have conducted an online literature search to find articles published in academic journals, data repositories, archives and/or scientific collections. You will therefore appreciate that such searches can quickly generate a great deal of information for you to review. Consequently, poorly considered searches can be time consuming and inefficient. A certain level of skill and experience is necessary for generating an effective literature search strategy. Adopting a methodological approach will serve to improve the utility of your searches [see Figure 1].

Objective – what is it for?

A good literature search starts with a well-defined research question that has been devised to achieve your goals. Poorly defined objectives can result in too many irrelevant search 'hits' requiring you to spend valuable time identifying the most relevant articles. Understand the purpose of your literature search and know what you want to achieve. Crafting an effective search strategy is essential and you'll need to consider the scope of work. A search intended to give an informed awareness of the current understanding on a specific subject needs a different approach than one for a systematic review, a meta-analysis or a Cochrane review.

Endpoints

Identify endpoints that will give you some indication of whether or not you have achieved your stated goals and establish how this will be best described. Endpoints are defined by your objective, however determining them will not necessarily be straightforward. Achieving your objective does not necessarily equate to the number of research papers you find, which might be 10 or 200.

"I don't pretend we have all the answers. But the questions are certainly worth thinking about"

Arthur C Clarke, c.2001

"Nothing has such power to broaden the mind as the ability to investigate systematically and truly all that comes under thy observation in life"
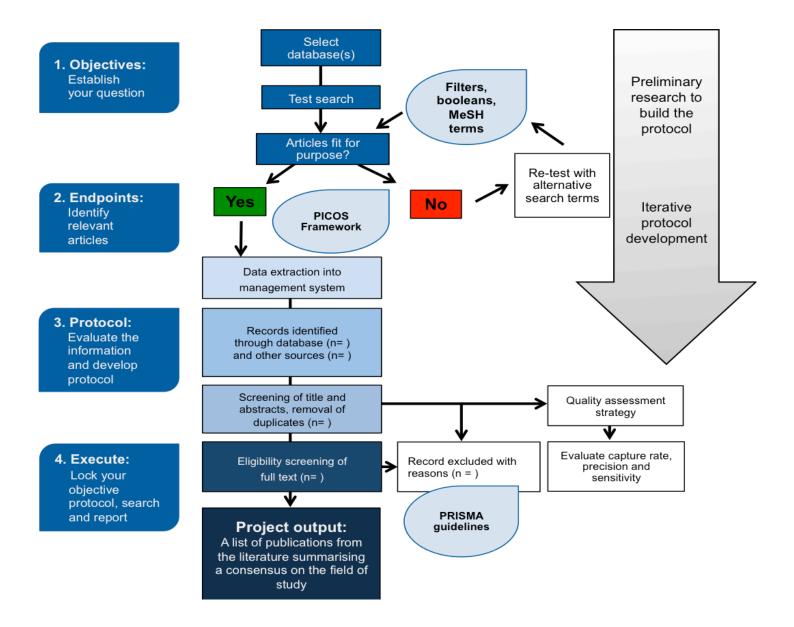
Marcus Aurelius, c.161–180

Figure 1. Conducting and reporting literature searches: an overview.
MeSH=Medical Subject Headings; PICOS=patient population, interventions or exposure, comparison and outcome or endpoint and study design; PRISMA=Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Cochrane reviews

Cochrane reviews summarise primary research in humans. They investigate the effects of interventions for prevention, treatment and rehabilitation and are published in the Cochrane Library. Reviews are prepared and maintained by the Cochrane Collaboration, consisting of over 31,000 specialists. Their reviews are renowned for the highest standard of evidence-based healthcare. There are six types of Cochrane review [1,2]:

- Intervention reviews – provide assessments of interventions in healthcare

- Diagnostic test accuracy reviews – evaluate diagnostic tests for a particular disease

- Methodology reviews – address how reviews and trials are conducted

- Qualitative reviews – assess evidence other than effectiveness

- Prognosis reviews – look at the probable course or outcome of a condition

- Overviews of Systematic Reviews (OoRs) – a new type of review that compiles data from several systematic reviews into one accessible document

Defining and refining search terms

When defining your search terms, carefully consider the key concepts of your topic. Think about the characteristics that define your patient population and/or their conditions. For example, if you are looking at older patients with diabetes you may want to search 'diabetes' and 'elderly'. After incorporating these key terms into the search, you can begin to address the complexities of the synonyms, alternative spellings and acronyms:

- Decide on a date range. It may be possible that terminology has changed over the years if you have included a broad date range

- Using filters can reduce the search hits returned and increase the degree of control you have over the search. For example, you can look specifically for reviews or meta-analyses and apply filters to include only those articles published in the last 5 years or exclude 'human' if you are looking at non-clinical studies

- The truncation feature can be used to search for different variants of a root or stem word. This option can be selected by putting an asterisk (*) at the end of a stem word. For example, searching for the term 'cardio*' will result in listings of all citations that include words that can be derived from the 'cardio' stem i.e., cardiogenic, cardiology, cardiogram, etc [3]

- Decide if you want to include common and/or scientific names, abbreviations and any synonyms in your search strategy. Increase the likelihood of retrieving relevant results by including several alternative words for elderly, such as 'senior', 'geriatric', 'older' and/or 'aged'. Remember, the more terms you combine, the more results are likely to be returned

- Using booleans enables the inclusion and exclusion of multiple search terms. Booleans need to be capitalised in your search.

    'AND' – narrows the search to hits containing both terms
    'OR' – broadens the search to hits containing either term
    'NOT' – removes terms that you don't want to include

You can combine booleans, for example:
['diabetes'] AND ['elderly' OR 'senior' OR 'geriatric' OR 'older' OR 'aged']

- Use quotation marks when you want to search for phrases or word combinations where all words appear immediately next to each other in a specific order

- Do you need to search for both British and American English variant spellings? Some search engines only use the specific word you enter, so if you want to ensure that you search for articles by American and English authors you will need to consider both language variants

What are MeSH headings?

Medical Subject Headings (MeSH) make up the National Library of Medicine's controlled vocabulary thesaurus and they are used for indexing Medline articles.

The hierarchical structure of MeSH enables users to search at varying degrees of specificity [4]. Over 27,000 descriptors are included in the MeSH dictionary and these files are updated daily.

The MeSH browser is a vocabulary look-up tool, which allows terms to be inserted into boxes. Using this technique is easier than using multiple brackets for several search terms and phrases in complicated searches.

If at the start of your search you are struggling to identify appropriate MeSH headings, you can obtain suggestions using 'MeSH on Demand'. This tool will process text to identify any contained MeSH terms. You could use text from a relevant review.

"The hardest thing of all is to find a black cat in a dark room, especially if there is no cat"

Confucius

Improve the relevancy of your results with PICOS...

Formulating a well-focused question is critical to obtaining a relevant articles that relate to your clinical question. The Cochrane Collaboration promotes use of PICOS when developing your question, it stands for patient population, interventions or exposure, comparison and outcome or endpoint and study design. The PICOS framework splits search terms into five components, which can be combined to formulate a search entry relevant to the research question [5, 6].

P   Patient population. What are the defining characteristics of the participants? What condition, age, gender, setting of care are you interested in?

I   Intervention (exposures). This may be a treatment (so consider the dose, frequency and duration), diagnostic tool, prevention or educational intervention.

C   Comparison. Is your intervention being compared to a placebo, standard or no treatment?

O   Outcome. Consider if your research question needs to address an improvement in condition, an effective diagnosis, pain reduction or improved quality of life.

S   Study design. Are you focussing on randomised trials, or will you include case reports and observational studies?

Write a protocol...

The utility of your output is dependent on the reproducibility of your search strategy and so it is important to minimise and potential subjectivity. Managing a multi-factorial process benefits from a system that clearly defines as many of the variables as possible. This is best addressed in the form of an objective protocol (you can get our search methodology template from the 'Resources' page on the Niche website) that outlines the brief, your proposed search strategy and criteria for review. Record the precise search terms, details of any filters and search engine(s), so that the search is reproducible.

Describe how outputs of the literature search are to be captured or stored and what information on each citation you will keep. The protocol should also describe how you will review and score each citation. Manually reviewing the titles and/or abstracts ensures that all the results adhere to the search criteria and all literature pertaining to the topic are collected. Employing stringent methods for selecting studies will limit bias, which in turn improves the reliability and accuracy of your conclusions. In your report, detail the process of study selection giving the number of studies screened, reasons for exclusion and the final numbers of articles included.

The PRISMA checklist for reporting guidelines

At least 2,500 literature new reviews are added to MEDLINE annually. A recent study of these entries found that many important aspects of the methodology for literature searches were not reported. For example, half of the 300 articles that this study reviewed failed to mention the terms 'systematic review' or 'meta-analysis' in the title or abstract and only two thirds reported the range of years over which the search was performed [7].

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were developed to help authors standardise the reporting of literature searches to achieve complete and transparent reporting and consist of a 27 point checklist of items to include. Items in the checklist include a rationale for the literature search, numbers of studies screened for eligibility and limitations that were encountered. The checklist is detailed in the PRISMA statement and website and is a helpful resource for authors to follow, ensuring clarity and transparency of the reporting of systematic reviews [8].

"Though this be madness, yet there is method in it." – Hamlet

William Shakespeare

Iterative protocol development, approval and finalising search strategies

You need to construct a search strategy that provides a comprehensive summary of the current research status in the field. To provide any degree of certainty that this is being achieved it is necessary to build iterative search adaptation into your methodology whereby you incorporate a 'check' of your outputs at each new iteration step. Be prepared to search and re-search the literature, using alternative search terms and keep a record of the method you followed for each step. There is no reason not to modify your protocol to see if one particular strategy gives a better result. However you will need to have a means of comparing the quality of your findings. For example, your change in approach is likely to affect the ratio of relevant references retrieved to non-relevant references. Just remember to keep a detailed record of the changes you make to your protocol.

Why write a protocol? With a protocol you can ask your colleagues to review your methodology and rationale. This will help to craft your search strategy by identifying any missing search terms or endpoints. It could save you time when sifting through the search results to eliminate indifferent or non-relevant articles. Without a protocol it will be more difficult to elicit and capture valuable feedback or to backtrack if you find that you have reached a dead end. In developing your search strategy, you may need to include a broad set of search terms in order to capture all the appropriate articles you need. For instance, there might be 10 different common and scientific names for a given condition, 10 more therapies you wish to include and several animal models in which relevant work may have been conducted.

Where to look…

It is possible to conduct online searches (on any topic pertaining to science) through academic search engines and commercial databases [9]. They provide a centralised platform and allow the researchers to acquire information on the cited literature within seconds. While there are many academic search engines available, some incorporate data from more trusted resources than others. They may provide information on a range of topics from engineering and technology to biology and natural science. Although these sources provide a one-stop solution to all research-related needs and literature searches, access is usually limited to academic institutional licenses.

Here we provide a list of academic databases and directories that are considered to be among the most trusted search engines for data on scientific research.

Online sources and commercial databases provide a personal and customised way to search research materials on any given topic. However, there is currently no easy way to compare the outcomes of two different databases – you need to find ways to do that yourself (see Reference Management Software Systems [10]).

Online search tools:

- BioOne
- CiteSeer
- Embase
- GetCITED
- Google Scholar
- MedWorm
- Microsoft Academic Research
- Portland Perpustakaan
- Pubmed
- OvidSP
- Science Direct
- Scopus
- Springer Links

"No idea is so outlandish that it should not be considered with a searching but at the same time with a steady eye."

Winston Churchill

An interview with one of our medical writers

Q What are you trying to achieve with a literature search?

A Although I am looking for data my main goal is efficiency. I might first do a simple search through PubMed to establish the extent of the likely output. This allows me to quickly assess what the search outputs are likely to be and tweak the search strategy to better achieve my goal. I aim to get the most precise record of the current understanding of the field of search; that equates to only the articles I need to see. Searching and re-searching the literature takes no time at all. The process of identifying relevant articles from among the many candidate citations is what takes the most effort and resource. Searching through hundreds of references can take days. Each iteration of your search should become more effective in that it results in a greater proportion of articles relevant to the objective. It can be useful to check your search terms with medical subject headings (MeSH), to cross reference whether or not your search terms are the same as those indexer uses [4].

Q How do you manage your search returns, do you use reference managing software?

A How one manages the many citations that are identified during your search will depend on the expected output. I tend to organise the information from each search result in a spreadsheet. There is also a function in PubMed that allows you to export into an Excel spread sheet all the available information about each article returned in a search. This is a very useful feature. For example, if you need more information you can export all the information, including the abstract.

Creating libraries within a reference management system is very helpful if you want to create bibliographies from a large number of references. There are several free and commercially available reference management systems such as Reference Manager and Endnote. These systems can take some time to set up and to incorporate your reference library of search results, but if you are planning to work in one field for some time these can be very helpful.

Reference Management Software Systems [10]

Free to use	Aigaion, Bebop, BibBase, BibDeck, BibSonomy, Bibus, CitULike, colwiz, Docear, JabRef, KBibTex, Mendeley, Pybliographer, Quiqqa*, ReadCube, refbase, RefDB, RefMe, Referencer, Wikindx, Zotero
For Purchase	Biblioscope, Bookends, Citavi, Endnote, Papers, Reference Manager, SciRef, Sente, WizFolio
Subscription	F1000Workspace, Paperpile, RefWorks

*also a premium subscription service

Q Why choose any one particular search engine?

A Your choice of search engine depends on the scope of the work and the field you're investigating. Certain subjects or therapy areas are likely to generate more data than others depending on which search engine you are using. To get an objective/comprehensive result you may need to use several search engines [9].

- If you use only one search engine you may end up inadvertently omitting important reports and/or data.

- If you are using more than one search engine it is highly likely that you will end up with duplicate articles within your list that will need to be removed. Some methods of removing duplicate articles are more efficient than others.

The method you finally decide to follow may be more influenced by practical, resource and/or budgetary factors than need of completeness of your final dataset. Similarly, poor results using only one database may spur you on to access a second, or even third, data source.

And finally...

It is possible to develop mechanisms or processes by which you can assess whether or not your search strategy provides an acceptable capture rate by testing the sensitivity and precision of the search. Start by obtaining a recent review to see what percentage of the cited references appear in your list. You'll also need to gauge what an acceptable capture rate is. If you are looking at randomised clinical trials, sensitivity might be defined by the total number of known trials identified by the search and precision is the proportion of publications retrieved that are actually randomised clinical trials [11]. Reasons for reduced sensitivity might include limited use of MeSH terms, absent publication types, missing methodological terms and truncation of terms.

The value of any single article you find is based not only on its intrinsic properties but also how it fits into and expands on previous work. You can identify what you feel are the most important articles by organising them to suit your requirements, such as selecting references published in journals with high impact factors or according to categories of clinical outcomes. If you rank articles manually you run the risk of introducing unacceptable subjectivity. You can minimise this by providing a formal list of inclusion/exclusion criteria and applying these to your search results using independent reviewers blinded to any other information about your candidate manuscripts. Using multiple reviewers introduces an additional level of objectivity.

References

1. https://en.wikipedia.org/wiki/Systematic_review
2. http://community.cochrane.org/cochrane-reviews
3. http://www.nlm.nih.gov/pubs/techbull/jf06/jf06_skillkit.html
4. http://en.wikipedia.org/wiki/Medical_Subject_Headings
5. Smith D, Reaney M, Speight J. Conducting Literature Reviews to Support the Use of Patient-Reported Outcomes (PRO) Measures in Clinical Trials – The Benefits of a Systematic Search Strategy. International Society for Pharmacoeconomics and outcomes research. News Article. July 2009
6. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 2007;7:16
7. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007; 4(3):e78
8. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009; 151: W65–94
9. https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines (accessed 20 June, 2016)
10. https://en.wikipedia.org/wiki/Comparison_of_reference_management_software (accessed 20 June, 2016)
11. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. Systematic reviews. In Chalmers I and Altman DG, editors. Systematic Reviews. London: BMJ Publishing group; 1995. p17–36
12. http://http://www.niche.org.uk/resource_centre.html

Next Steps

When done correctly, literature searching is invaluable for providing insights into research and developing evidence-based guidelines and recommendations. We created this Insider's Insight into Literature Searches to share some helpful pointers. We hope you found it useful. We can also share with you a template for your search protocol [12], which is a useful way to start collating the results of your new literature search. If you would like advice on conducting your literature search please contact me at the email address below.

Dr Justin Cook
Head of Medical Writing
Justin.Cook@niche.org.uk

Get in touch

niche
science&technology
+44 (0)20 8332 2588
www.niche.org.uk