How to carry out a literature search for a systematic review: a practical guide

Lauren Z. Atkinson & Andrea Cipriani

SUMMARY
Performing an effective literature search to obtain the best available evidence is the basis of any evidence-based discipline, in particular evidence-based medicine. However, with a vast and growing volume of published research available, searching the literature can be challenging. Even when journals are indexed in electronic databases, it can be difficult to identify all relevant studies without an effective search strategy. It is also important to search unpublished literature to reduce publication bias, which occurs from a tendency for authors and journals to preferentially publish statistically significant studies. This article is intended for clinicians and researchers who are approaching the field of evidence synthesis and would like to perform a literature search. It aims to provide advice on how to develop the search protocol and the strategy to identify the most relevant evidence for a given research or clinical question. It will also focus on how to search not only the published but also the unpublished literature using a number of online resources.

LEARNING OBJECTIVES
• Understand the purpose of conducting a literature search and its integral part of the literature review process
• Become aware of the range of sources that are available, including electronic databases of published data and trial registries to identify unpublished data
• Understand how to develop a search strategy and apply appropriate search terms to interrogate electronic databases or trial registries

DECLARATION OF INTEREST
None.

A literature search is distinguished from, but integral to, a literature review. Literature reviews are conducted for the purpose of (a) locating information on a topic or identifying gaps in the literature for areas of future study, (b) synthesising conclusions in an area of ambiguity and (c) helping clinicians and researchers inform decision-making and practice guidelines. Literature reviews can be narrative or systematic, with narrative reviews aiming to provide a descriptive overview of selected literature, without undertaking a systematic literature search. By contrast, systematic reviews use explicit and replicable methods in order to retrieve all available literature pertaining to a specific topic to answer a defined question (Higgins 2011). Systematic reviews therefore require a priori strategies to search the literature, with predefined criteria for included and excluded studies that should be reported in full detail in a review protocol.

Performing an effective literature search to obtain the best available evidence is the basis of any evidence-based discipline, in particular evidence-based medicine (Sackett 1997; McKeever 2015). However, with a vast and growing volume of published research available, searching the literature can be challenging. Even when journals are indexed in electronic databases, it can be difficult to identify all relevant studies without an effective search strategy (Hopewell 2007). In addition, unpublished data and ‘grey’ literature (informally published material such as conference abstracts) are now becoming more accessible to the public. It is important to search unpublished literature to reduce publication bias, which occurs because of a tendency for authors and journals to preferentially publish statistically significant studies (Dickersin 1993). Efforts to locate unpublished and grey literature during the search process can help to reduce bias in the results of systematic reviews (Song 2010). A paradigmatic example demonstrating the importance of capturing unpublished data is that of Turner et al (2008), who showed that using only published data in their meta-analysis led to effect sizes for antidepressants that were one-third (32%) larger than effect sizes derived from combining both published and unpublished data. Such differences in findings from published and unpublished data can have real-life implications in clinical decision-making and treatment recommendation. In another relevant publication, Whittington et al (2004) compared the risks and benefits of selective serotonin reuptake inhibitors (SSRIs) in the treatment
of depression in children. They found that published data suggested favourable risk-benefit profiles for SSRIs in this population, but the addition of unpublished data indicated that risk outweighed treatment benefits. The relative weight of drug efficacy to side-effects can be skewed if there has been a failure to search for, or include, unpublished data.

In this guide for clinicians and researchers on how to perform a literature search we use a working example about efficacy of an intervention for bipolar disorder to demonstrate the search techniques outlined. However, the overarching methods described are purposefully broad to make them accessible to all clinicians and researchers, regardless of their research or clinical question.

Defining the clinical question

The review question will guide not only the search strategy, but also the conclusions that can be drawn from the review, as these will depend on which studies or other forms of evidence are included and excluded from the literature review. A narrow question will produce a narrow and precise search, perhaps resulting in too few studies on which to base a review, or be so focused that the results are not useful in wider clinical settings. Using an overly narrow search also increases the chances of missing important studies. A broad question may produce an imprecise search, with many false-positive search results. These search results may be too heterogeneous to evaluate in one review. Therefore from the outset, choices should be made about the remit of the review, which will in turn affect the search.

A number of frameworks can be used to break the review question into concepts. One such is the PICO (population, intervention, comparator and outcome) framework, developed to answer clinical questions such as the effectiveness of a clinical intervention (Richardson 1995). It is noteworthy that ‘outcome’ concepts of the PICO framework are less often used in a search strategy as they are less well defined in the titles and abstracts of available literature (Higgins 2011). Although PICO is widely used, it is not a suitable framework for identifying key elements of all questions in the medical field, and minor adaptations are necessary to enable the structuring of different questions. Other frameworks exist that may be more appropriate for questions about health policy and management, such as ECLIPSE (expectation, client group, location, impact, professionals, service) (Wildridge 2002) or SPICE (setting, perspective, intervention, comparison, evaluation) for service evaluation (Booth 2006). A detailed overview of frameworks is provided in Davies (2011).

Scoping search

Before conducting a comprehensive literature search, a scoping search of the literature using just one or two databases (such as PubMed or MEDLINE) can provide valuable information as to how much literature for a given review question already exists. A scoping search may reveal whether systematic reviews have already been undertaken for a review question. Caution should be taken, however, as systematic reviews that may appear to ask the same question may have differing inclusion and exclusion criteria for studies included in the review. In addition, not all systematic reviews are of the same quality. If the original search strategy is of poor quality methodologically, original data are likely to have been missed and the search should not simply be updated (compare, for example, Naughton et al (2014) and Caddy et al (2015) on ketamine for treatment-resistant depression).

Search strategy

The first step in conducting a literature search should be to develop a search strategy. The search strategy should define how relevant literature will be identified. It should identify sources to be searched (list of databases and trial registries) and keywords used in the literature (list of keywords). The search strategy should be documented as an integral part of the systematic review protocol. Just as the rest of a well-conducted systematic review, the search strategy used needs to be explicit and detailed such that it could be reproduced using the same methodology, with exactly the same results, or updated at a later time. This not only improves the reliability and accuracy of the review, but also means that if the review is replicated, the difference in reviewers should have little effect, as they will use an identical search strategy. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement was developed to standardise the reporting of systematic reviews (Moher 2009). The PRISMA statement consists of a 27-item checklist to assess the quality of each element of a systematic review (items 6, 7 and 8 relate to the quality of literature searching) and also to guide authors when reporting their findings.

Sources to search

There are a number of databases that can be searched for literature, but the identification of relevant sources is dependent on the clinical or research question (different databases have different focuses, from more biology to more social science oriented) and the type of evidence that is sought (i.e. some databases report only randomised controlled trials).
The examples of specific databases given here are by no means exhaustive, but they are popular and likely to be used for literature searching in medicine, psychiatry and psychology. Website links for these databases are given in Box 1, along with links to resources not mentioned above. Box 1 also provides a website link to a couple of video tutorials for searching electronic databases. Box 2 shows an example of the search sources chosen for a review of a pharmacological intervention of calcium channel antagonists in bipolar disorder, taken from a recent systematic review (Cipriani 2016a).

For a comprehensive search of the literature it has been suggested that two or more electronic databases should be used (Suarez-Almazor 2000). Suarez-Almazor and colleagues demonstrated that, in a search for controlled clinical trials (CCTs) for rheumatoid arthritis, osteoporosis and lower back pain, only 67% of available citations were found by both Embase and MEDLINE. Searching MEDLINE alone would have resulted in 25% of available CCTs being missed and searching Embase alone would have resulted in 15% of CCTs being missed. However, a balance between the sensitivity of a search (an attempt to retrieve all relevant literature in an extensive search) and the specificity of a search (an attempt to retrieve a more manageable number of relevant citations) is optimal. In addition, supplementing electronic database searches with unpublished literature searches (see ‘Obtaining unpublished literature’ below) is likely to reduce publication bias. The capacity of the individuals or review team is likely largely to determine the number of sources searched. In all cases, a clear rationale should be outlined in the review protocol.
for the sources chosen (the expertise of an information scientist is valuable in this process).

**Developing a search strategy**

Important methodological considerations (such as study design) may also be included in the search strategy. Dependent on the databases and supplementary sources chosen, filters can be used to search the literature by study design (see ‘Searching electronic databases’). For instance, if the search strategy is confined to one study design term only (e.g. randomised controlled trial, RCT), only the articles labelled in this way will be selected. However, it is possible that in the database some RCTs are not labelled as such, so they will not be picked up by the filtered search. Filters can help reduce the number of references retrieved by the search, but using just one term is not 100% sensitive, especially if only one database is used (i.e. MEDLINE). It is important for systematic reviewers to know how reliable such a strategy can be and treat the results with caution.

**Searching electronic databases**

*Identifying search terms*

Standardised search terms are thesaurus and indexing terms that are used by electronic databases as a convenient way to categorise articles, allowing for efficient searching. Individual database records may be assigned several different standardised search terms that describe the same or similar concepts (e.g. bipolar disorder, bipolar depression, manic-depressive psychosis, mania). This has the advantage that even if the original article did not use the standardised term, when the article is catalogued in a database it is allocated that term (Guainia 2010). For example, an older paper might refer to ‘manic depression’, but would be categorised under the term ‘bipolar disorder’ when catalogued in MEDLINE. These standardised search terms are called MeSH (medical subject headings) in MEDLINE and PubMed, and Entree in Embase, and are organised in a hierarchical structure (Fig. 1). In both MEDLINE and Embase an ‘explode’ command enables the database to search for a requested term, as well as specific related terms. Both narrow and broader search terms can be viewed and selected to be included in the search if appropriate to a topic. The Yale MeSH Analyzer tool (mesh.med.yale.edu) can be used to help identify potential terms and phrases to include in a search. It is also useful to understand why relevant articles may be missing from an initial search, as it produces a comparison grid of MeSH terms used to index each article (see Box 1 for a tutorial video link).

In addition, MEDLINE also distinguishes between MeSH headings (MH) and publication type (PT) terms. Publication terms are less about the content of an article than about its type, specifying for example a review article, meta-analysis or RCT.
Both MeSH and Emtree have their own peculiarities, with variations in thesaurus and indexing terms. In addition, not all concepts are assigned standardised search terms, and not all databases use this method of indexing the literature. It is advisable to check the guidelines of selected databases before undertaking a search. In the absence of a MeSH heading for a particular term, free-text terms could be used.

Free-text terms are used in natural language and are not part of a database’s controlled vocabulary. Free-text terms can be used in addition to standardised search terms in order to identify as many relevant records as possible (Higgins 2011). Using free-text terms allows the reviewer to search using variations in language or spelling (e.g. hypomania* or mania* or manic* – see truncation and wildcard functions below and Fig. 2). A disadvantage of free-text terms is that they are only searched for in the title and abstracts of database records, and not in the full texts, meaning that when a free-text word is used only in the body of an article, it will not be retrieved in the search. Additionally, a number of specific considerations should be taken into account when selecting and using free-text terms:

- synonyms, related terms and alternative phrases (e.g. mood instability, affective instability, mood liability or emotion dysregulation)
- abbreviations or acronyms in medical and scientific research (e.g. magnetic resonance imaging or MRI)
- lay and medical terminology (e.g. high blood pressure or hypertension)
- brand and generic drug names (e.g. Prozac or fluoxetine)
- variants in spelling (e.g. UK English and American English: behaviour or behavior; paediatric or pediatric).

Truncation and wildcard functions can be used in most databases to capture variations in language:

- truncation allows the stem of a word that may have variant endings to be searched: for example, a search for depress* uses truncation to retrieve articles that mention both depression and depression.

**FIG 2** Example of a search strategy about bipolar disorder using MEDLINE (Cipriani 2016a). The strategy follows the PICO framework and includes MeSH terms, free-text keywords and a number of other techniques, such as truncation, that have been outlined in this article. Numbers in bold give the number of citations retrieved by each search.
and depressive; truncation symbols may vary by database, but common symbols include: *, ! and #.

- wild cards substitute one letter within a word to retrieve alternative spellings: for example, ‘wom?n’ would retrieve the terms ‘woman’ and ‘women’.

**Combining search terms**

Search terms should be combined in the search strategy using Boolean operators. Boolean operators allow standardised search terms and free-text terms to be combined. There are three main Boolean operators – AND, OR and NOT (Fig. 3).

- **OR** – this operator is used to broaden a search, finding articles that contain at least one of the search terms within a concept. Sets of terms can be created for each concept, for example the population of interest: (bipolar disorder OR bipolar depression). Parentheses are used to build up search terms, with words within parentheses treated as a unit.

- **AND** – this can be used to join sets of concepts together, narrowing the retrieved literature to articles that contain all concepts, for example the population or condition of interest and the intervention to be evaluated: (bipolar disorder OR bipolar depression) AND calcium channel blockers. However, if at least one term from each set of concepts is not identified from the title or abstract of an article, this article will not be identified by the search strategy. It is worth mentioning here that some databases can run the search also across the full texts. For example, ScienceDirect and most publishing houses allow this kind of search, which is much more comprehensive than abstract or title searches only.

- **NOT** – this operator, used less often, can focus a search strategy so that it does not retrieve specific literature, for example human studies NOT animal studies. However, in certain cases the NOT operator can be too restrictive, for example if excluding male gender from a population, using ‘NOT male’ would also mean that any articles about both males and females are not obtained by the search.

The conventions of each database should be checked before undertaking a literature search, as functions and operators may differ slightly between them (Cipriani 2016b). This is particularly relevant when using limits and filters. Figure 2 shows an example search strategy incorporating many of the concepts described above. The search strategy is taken from Cipriani et al (2016a), but simplified to include only one intervention.

**Search filters**

A number of filters exist to focus a search, including language, date and study design or study focus filters. Language filters can restrict retrieval of articles to the English language, although if language is not an inclusion criterion it should not be restricted, to avoid language bias. Date filters can be used to restrict the search to literature from a specified period, for example if an intervention was only made available after a certain date. In addition, if good systematic reviews exist that are likely to capture all relevant literature (as advised by an information specialist), date restrictions can be used to search additional literature published after the date of that included in the systematic review. In the same way, date filters can be used to update a literature search since the last time it was conducted. Reviewing the literature should be a timely process (new and potentially relevant evidence is produced constantly) and updating the search is an important step, especially if collecting evidence to inform clinical decision-making, as publications in the field of medicine are increasing at an impressive rate (Barber 2016). The filters chosen will depend on the research question and nature of evidence that is sought through the literature search and the guidelines of the individual database that is used.
Supplementary search techniques

Google Scholar

Google Scholar allows basic Boolean operators to be used in strings of search terms. However, the search engine does not use standardised search terms that have been tagged as in traditional databases and therefore variations of keywords should always be searched. There are advantages and disadvantages to using a web search engine such as Google Scholar. Google Scholar searches the full text of an article for keywords and also searches a wider range of sources, such as conference proceedings and books, that are not found in traditional databases, making it a good resource to search for grey literature (Haddaway 2015). In addition, Google Scholar finds articles cited by other relevant articles produced in the search. However, variable retrieval of content (due to regular updating of Google algorithms and the individual’s search history and location) means that search results are not necessarily reproducible and are therefore not in keeping with replicable search methods required by systematic reviews. Google Scholar alone has not been shown to retrieve more literature than other traditional databases discussed in this article and therefore should be used in addition to other sources (Bramer 2016).

Citation searching

Once the search strategy has identified relevant literature, the reference lists in these sources can be searched. This is called citation searching or backward searching, and it can be used to see where particular research topics led others. This method is particularly useful if the search identifies systematic reviews or meta-analyses of a similar topic.

Obtaining unpublished literature

Conference abstracts

Conference abstracts are considered ‘grey literature’, i.e. literature that is not formally published in journals or books (Alberani 1990). Scherer and colleagues found that only 52.6% of all conference abstracts go on to full publication of results, and factors associated with publication were studies that had RCT designs and the reporting of positive or significant results (Scherer 2007). Therefore, failure to search relevant grey literature might miss certain data and bias the results of a review. Although conference abstracts are not indexed in most major electronic databases, they are available in databases such as BIOSIS Previews (Box 1). However, as with many unpublished studies, these data did not undergo the peer review process that is often a tool for assessing and possibly improving the quality of the publication.

Searching trial registers and pharmaceutical websites

For reviews of trial interventions, a number of trial registers exist. ClinicalTrials.gov (clinicaltrials.gov) provides access to information on public and privately conducted clinical trials in humans. Results for both published and unpublished studies can be found for many trials on the register, in addition to information about studies that are ongoing. Searching each trial register requires a slightly different search strategy, but many of the basic principles described above still apply. Basic searches on ClinicalTrials.gov include searching by condition, specific drugs or interventions and these can be linked using Boolean operators: for example, (bipolar disorder OR manic depressive disorder) AND lithium. As mentioned above, parentheses can be used to build up search terms. More advanced searches allow one to specify further search fields such as the status of studies, study type and age of participants. The US Food and Drug Administration (FDA) hosts a database providing information about FDA-approved drugs, therapeutic products and devices (www.fda.gov). The database (with open access to anyone, not only in the USA) can be searched by the drug name, its active ingredient or its approval application number and, for most drugs approved in the past 20 years or so, a review of clinical trial results (some of which remain unpublished) used as evidence in the approval process is available. The European Medicines Agency (EMA) hosts a similar register for medicines developed for use in the European Union (www.ema.europa.eu). An internet search will show that many other national and international trial registers exist that, depending on the review question, may be relevant search sources. The World Health Organization International Clinical Trials Registry Platform (WHO ICTRP; www.who.int/ictrp) provides access to a central database bringing a number of these national and international trial registers together. It can be searched in much the same way as ClinicalTrials.gov.

A number of pharmaceutical companies now share data from company-sponsored clinical trials. GlaxoSmithKline (GSK) is transparent in the sharing of its data from clinical studies and hosts its own clinical study register (www.gsk-clinicalstudymonitor.com). Eli-Lilly provides clinical trial results both on its website (www.lillyclinicaltrialguide.com) and in external registries. However, other pharmaceutical companies, such as Wyeth and Roche, divert users to clinical trial results in external registries. These registries include both published and previously unpublished studies. Searching techniques differ for each company and hand-searching through documents is often required to identify studies.
Communication with authors

Direct communication with authors of published papers could produce both additional data omitted from published studies and other unpublished studies. Contact details are usually available for the corresponding author of each paper. Although high-quality reviews do make efforts to obtain and include unpublished data, this does have potential disadvantages: the data may be incomplete and are likely not to have been peer-reviewed. It is also important to note that, although reviewers should make every effort to find unpublished data in an effort to minimise publication bias, there is still likely to remain a degree of this bias in the studies selected for a systematic review.

Conclusions

Developing a literature search strategy is a key part of the systematic review process, and the conclusions reached in a systematic review will depend on the quality of the evidence retrieved by the literature search. Sources should therefore be selected to minimise the possibility of bias, and supplementary search techniques should be used in addition to electronic database searching to ensure that an extensive review of the literature has been carried out. It is worth reminding that developing a search strategy should be an iterative and flexible process (Higgins 2011), and only by conducting a search oneself will one learn about the vast literature available and how best to capture it.

Acknowledgements

We thank Sarah Stockton for her help in drafting this article. Andrea Cipriani is supported by the NIHR Oxford cognitive health Clinical Research Facility.

References


<table>
<thead>
<tr>
<th>MCQs</th>
<th>Select the single best option for each question stem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td>A systematic literature review is:</td>
</tr>
<tr>
<td>a</td>
<td>an explicit and replicable method used to retrieve all available literature pertaining to a specific topic to answer a defined question</td>
</tr>
<tr>
<td>b</td>
<td>a descriptive overview of selected literature</td>
</tr>
<tr>
<td>c</td>
<td>an initial impression of a topic which is understood more fully as a research study is conducted</td>
</tr>
<tr>
<td>d</td>
<td>a method of gathering opinions of all clinicians or researchers in a given field</td>
</tr>
<tr>
<td>e</td>
<td>a step-by-step process of identifying the earliest published literature through to the latest published literature.</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>A search protocol:</td>
</tr>
<tr>
<td>a</td>
<td>does not need to be specified in advance of a literature search</td>
</tr>
<tr>
<td>b</td>
<td>does not need to be reported in a systematic literature review</td>
</tr>
<tr>
<td>c</td>
<td>defines which sources of literature are to be searched, but not how a search is to be carried out</td>
</tr>
<tr>
<td>d</td>
<td>defines how relevant literature will be identified and provides a basis for the search strategy</td>
</tr>
<tr>
<td>e</td>
<td>provides a timeline for searching each electronic database or unpublished literature source.</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>To identify studies from a topic in general medicine, it would be most appropriate to search:</td>
</tr>
<tr>
<td>a</td>
<td>the Cochrane Central Register of Controlled Trials (CENTRAL)</td>
</tr>
<tr>
<td>b</td>
<td>MEDLINE</td>
</tr>
<tr>
<td>c</td>
<td>PsycINFO</td>
</tr>
<tr>
<td>d</td>
<td>the Cumulative Index to Nursing and Allied Health Literature (CINAHL)</td>
</tr>
<tr>
<td>e</td>
<td>the British Nursing Index.</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>Literature about available treatments for bipolar disorder would be retrieved using the search terms:</td>
</tr>
<tr>
<td>a</td>
<td>bipolar disorder OR treatment</td>
</tr>
<tr>
<td>b</td>
<td>bipolar* OR treatment</td>
</tr>
<tr>
<td>c</td>
<td>bipolar disorder AND treatment</td>
</tr>
<tr>
<td>d</td>
<td>bipolar disorder NOT treatment</td>
</tr>
<tr>
<td>e</td>
<td>(bipolar disorder) OR (treatment).</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>Supplementing electronic database searches with unpublished literature searches is likely to reduce the possibility that a systematic review will have:</td>
</tr>
<tr>
<td>a</td>
<td>publication bias</td>
</tr>
<tr>
<td>b</td>
<td>funding bias</td>
</tr>
<tr>
<td>c</td>
<td>language bias</td>
</tr>
<tr>
<td>d</td>
<td>outcome reporting bias</td>
</tr>
<tr>
<td>e</td>
<td>selection bias.</td>
</tr>
</tbody>
</table>