



Learning to successfully search the scientific and medical literature

Emily A. Thompson¹ · Laurissa B. Gann² · Erik N. K. Cressman³

Received: 11 February 2019 / Revised: 16 February 2019 / Accepted: 19 February 2019 / Published online: 6 March 2019
© Cell Stress Society International 2019

Abstract

Searching the literature is often overlooked and receives inadequate attention. In this article, we seek to address this issue by presenting several strategies. Here, five steps are outlined and discussed to facilitate effective literature searching.

Keywords Literature · Search · Database · Publication · Autoalert · Citation mining · Search strategy

Rapidly and thoroughly understanding published literature in a new area of research is a challenge for graduate students in the life sciences, medical students, and veterinary students, but the issue applies in any area of science. Using powerful tools with a focused strategy is of tremendous assistance in learning the literature on a topic in a timely manner. Acquiring this skill set is essential, but it is often overlooked, causing frustration, wasted time, and wasted resources, and can result in unnecessary duplication of effort. Frank Westheimer, noted chemist and National Science Medal Awardee, is credited with the statement, “A couple of months in the laboratory can frequently save a couple of hours in the library” (Crampon 1988). Some would argue that a couple of months is a significant underestimate of the time involved. This article will detail a functional and systematic approach to a literature search using tools that were unavailable in Westheimer’s day and describe how to leverage the power of online resources.

What type of question is being asked?

When beginning a literature search, it is important to consider the type of information in question because the nature of the question will dictate the search approach. For example, questions concerning a drug, disease, or general medical information can be answered by reviewing sources such as UpToDate (www.uptodate.com) or DynaMed (www.dynamed.com/home/). For clinical questions or questions regarding clinical research, PubMed Clinical Queries is a readily available resource. Further, for more encompassing literature searches, such as those beginning a new research project, a database such as PubMed or Web of Science is necessary for a comprehensive search of scientific literature. Other fields will have similar comprehensive databases but are outside the scope of this article.

Five steps are necessary to perform a comprehensive search: develop a focused question, plan the search, execute the plan, track and store results, and finally, stay current with newly published material.

✉ Erik N. K. Cressman
ECressman@mdanderson.org

Emily A. Thompson
EThompson@mdanderson.org

Laurissa B. Gann
LGann@mdanderson.org

¹ Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1472, Houston, TX, USA

² Research Medical Library, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1499, Houston, TX, USA

³ Department of Interventional Radiology, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1471, Houston, TX 77030, USA

Step one

Ask a focused question

Form a question

A sound strategy is to begin with a question that is focused, well-defined, and answerable. Diving into a literature review without clearly identifying the question will lead to lost time, and frustration. The PICO tool (as seen in Table 1) was designed to help researchers think through their proposed topic by breaking it down into concepts

Table 1 PICO tool acronym and description

Acronym	Description
P: patient, population, or problem	How would you describe the patient population or problem? What are the characteristics?
I: intervention	What intervention is under study?
C: comparison	What is the current practice? How does this compare to other alternatives?
O: outcome	What is the overall expected result?

(Richardson et al. 1995). Overlooking this step can lead to a search that is much too broad or too narrow. A question that is too broad may lead to too many search results, and a question that is too narrow will lead to the opposite. The PICO tool also assists in generating keywords or search terms that will assist authors later in their database searching.

Explore the question

After defining a question, perform an exploratory search to determine if the question is original or not. Very often, a seemingly original idea has already been explored by others and the novice researcher is only just entering the fray. At this point, it should be confirmed that the question under review is both answerable and has not already been answered by other authors. To gain a better understanding of the topic, and to ensure a completed PICO model, search DynaMed, UpToDate, Google, books, or [PubMed Clinical Queries](#) for preliminary information.

Finalize the question

As experienced researchers can often provide advice concerning question formulation, it is advisable to review the topic and proposed question with potential collaborators/co-authors. Discussing with collaborators and co-authors also ensures that all parties are in agreement with the potential research project and scope of the literature search.

Step two

Plan the literature search

Before searching, determine what search methods are available and strategize how to best use them.

Determine appropriate databases

Make a list of databases that can be used to search for information related to the topic. While there may be content overlap between databases, each will contain unique material as well. Thus, multiple databases should be investigated for a comprehensive review. Investing time in learning how to navigate these databases is worthwhile. Each database has unique searching features which will help focus a search. If unsure, contact an experienced collaborator, co-author, or medical librarian who can provide advice. [Table 2](#) identifies databases that are often recommended in the medical sciences.

Develop the search strategy

After selecting appropriate databases, develop a comprehensive search strategy for database searching. A comprehensive strategy should include a combination of keywords and subject headings or index terms.

Find a pivotal/seminal paper

Often, the best starting point is to find one seminal or pivotal paper to serve as a guide for the entire search. A pivotal paper is one that is highly cited and provides crucial information the field builds upon. In some cases, it may be a review article rather than a major scientific study. In other words, a pivotal paper is one that serves as a building block for the field. For a researcher, this paper can stand as the basis of the literature search. Start by performing a quick preliminary search in Web

Table 2 Databases common in medical sciences

Database	Coverage
PubMed	Covers biomedical and life sciences journals and is easy to search. PubMed is the most up-to-date database and mostly covers US journals.
Medline (Ovid)	Covers biomedical and medical science journals and includes special search features for a more focused search. Medline covers mostly US journals.
Embase (Ovid or Embase.com)	Covers biomedical, medical, drug, and pharmacy journals, and includes special Emtree index terms for a more focused search. Embase includes significantly more European and Asian journal titles than PubMed or Medline.
Scopus	Covers biomedical and medical sciences and offers specialized author searches and citation mining tools.
Web of Science	Covers biomedical, social sciences, art, and humanities, and offers specialized citation mining tools.

of Science, Google Scholar, or Scopus. The key feature of these databases is that they all have a sorting function that enables display of the results by the number of times a paper has been cited. Some databases may also suggest related citations. This will provide another starting point from which the literature search can continue. (1) Repeating the algorithm of who cited a paper, (2) sorting the resulting list by number of citations, and (3) selecting the most cited paper(s) and determining who cited them are powerful, leveraged approaches that will quickly produce a fairly comprehensive set of important papers and prominent investigators within a particular field.

Choose keywords

Keywords should be used alongside subject headings to capture any alternate concepts that cannot be captured with subject headings alone. It may take several months for subject headings to be applied to an article after it appears in a database, and some concepts are not easily defined with one term. For example, hematologic malignancies require multiple subject headings or keywords to search for all possible hematologic cancers (i.e. leukemia, lymphoma). When selecting keywords, remember to include synonyms (ultrasonography vs. ultrasound, elasticity imaging[MeSH] vs elastography), differences in spelling, suffixes, and generic and trade names of drugs, and be sure to spell out abbreviated words (i.e., FISH or fluorescence in situ hybridization). Learn to truncate with wild card symbols according to the syntax of any given software. An example of truncation is the asterisk used in PubMed to permit a search using any possible permutations of a root term. For instance, if you search PubMed for malignan*, PubMed will retrieve malignant, malignancy, malignancies, etc. After searching, one can review the search details box on the right-hand side of the search results page as shown in Fig. 1 to find out which permutations PubMed applied to the truncated term. Note in this example, the scroll bar indicates that there

Fig. 1 Example of PubMed expansion of the truncated search term “malignan*” using the wild card asterisk. This section is found on the right side of the search result page. Note the scroll bar to the right indicating many additional variations not shown in the window and the link to see more terms used

are many additional permutations not shown, and be aware that truncation may not be appropriate for every search as it can broaden your search significantly.

Identify subject headings

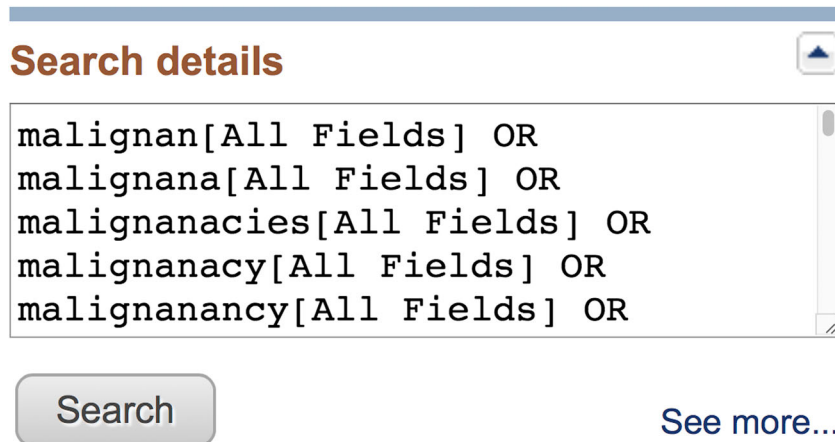
Subject headings are a set of descriptive vocabulary terms organized in a hierarchical structure within a database. Headings are important because they identify the most important concepts in the article. Medical Subject Headings (MeSH) is a tool provided by PubMed and Medline that enables readers to explore the database of terms or “mine search results” by looking at the MeSH terms associated with similar articles (see citation mining below). These can be found under the article abstract in a collapsible menu titled MeSH. Be sure to note that it may take a few months before MeSH terms are added for an article, so they may not be available for recently published items. In particular, pay close attention to MeSH terms from any pivotal papers you identify along the way. These MeSH subheadings may not have been terms that you initially used, or they may help you be more specific, and are frequently a fast yet effective way to narrow down your search to results that are truly meaningful.

Use Boolean language

Many databases and search engines capitalize on the use of Boolean language. Commonly, the use of AND, OR, or NOT can be used to combine, link, or exclude search terms. Try this when looking to expand or narrow a search.

Define inclusion and exclusion criteria

From the beginning, determine characteristics of which articles should be included in the literature search. Critically consider date ranges, languages, type of manuscript, and type of study. It can be tempting to only include recent articles (last



Search details

malignan[All Fields] OR
 malignana[All Fields] OR
 malignanacies[All Fields] OR
 malignanacy[All Fields] OR
 malignanancy[All Fields] OR

Search

[See more...](#)

10 years) or articles in a certain language, but keep in mind that a pivotal or foundational article could be excluded if outside the bounds of the search criteria. Searching the older literature can help one understand the history and scientific progression of a field. Often, seemingly new ideas will have already been explored decades earlier.

Confirm with collaborators

Converse with co-authors and discuss search strategy and inclusion criteria.

Step three

Perform the search

Translate the search to databases

Each database has a unique structure so be prepared to modify search terms to match the searching features of each database. Follow the search strategy outlined above as closely as possible and be aware that variations will exist between databases.

Track search methods

Keep track of search methods used throughout the literature search. This is an important part of research reproducibility and will be useful when writing a literature review. Note the databases searched, inclusion and/or exclusion criteria, and the search terms used. Consider also recording inclusion and/or exclusion criteria agreed upon after discussion with co-authors. Often, a spreadsheet or similar document is helpful to organize this information. Since a search is often a dynamic process with many branches, tracking helps to avoid unnecessary duplication.

Search beyond databases

Not every important reference is in a database. Alternatively, an important reference may reside in a database already searched, but was not identified with a particular search strategy because of how the reference was indexed. Depending on the topic, it may be useful to try citation mining, cited reference searching, or looking at the gray literature.

Citation mining Citation mining is a search strategy that entails using a previously found pivotal citation to discover further relevant citations. This method involves determining who has cited an article since it was first published and exploring those people or research groups. This is a powerful way to find related articles, search more recently published literature, and analyze the impact of the original publication. Web of

Science, Scopus, and Google Scholar offer this functionality. Repeating this process as described above can quickly provide a broad overview of important contributions and contributors.

Gray literature Gray literature refers to materials that are not published in journals or searchable in a database. This includes guidelines, clinical trials, open data, reports, conference proceedings, and other government or organizational documents. Gray literature resources will vary based on the topic.

Step four

Saving, storing, and sharing

Literature searches can be time consuming and may take days, weeks, or even months to complete. It is important to stay organized by saving searches and search results along the way.

Saving the search

Most databases include a feature that allows searches to be saved. For example, creating a free NCBI account in PubMed is available to the public and enables users to save previous searches. It is a good idea to save various iterations of the chosen search strategy so results can be easily reproduced. To save a search, log in to an NCBI account, search PubMed, choose Create Alert underneath the search box, and customize the search. Here, options include simply saving the search or email updates.

Save references

Reference manager software is a powerful tool for organizing, storing, and sharing references. It will also save time when creating a bibliography. A number of tools are available for authors, some of which are available at no cost. Check with a librarian to see what resources are available at your institution.

Free tools

Mendeley (desktop, online) and Zotero (online only)—note that there are typically limited amounts of storage in the free versions.

Commercially available tools

EndNote (desktop, online) and RefWorks (online only)—check with your librarian as one or more may already be available via your institution.

Step five

Staying current

Stay current

Databases are constantly changing, with new material being added daily. It is imperative to maintain familiarity with the published literature even after completing a literature search. It may take months or even years to complete a manuscript, and relevant articles will likely be published in the meantime. Many authors will save their search strategies in different databases and re-run the search immediately before submission to make sure they have not missed recent publications (see Save the Search Strategies above). Keeping up with literature on a daily or weekly basis may include setting aside some time each day or each week to read through alerts and/or relevant articles on social media. It is not necessary to redo the entire literature search, but it is important to keep up with recent publications. Save articles and citations using tools like BrowZine, Refworks, Mendeley, and EndNote. Here are a few ways to keep up-to-date:

Email alerts Databases like PubMed and Scopus allow users to create customized alerts by saving a search strategy. The strategy may combine multiple terms or perhaps simply track the output of an author (a major contributor to the field for example). One can then customize the frequency and volume of alerts to receive a weekly or monthly email regarding new publications. Similarly, Google has a function to set up news alerts about the latest press releases from major companies and organizations.

Follow major journals Most major journals offer a feature for setting up email alerts through their website. However, setting up alerts for each individual journal can quickly overwhelm any email inbox. Try using tools like [Read by QxMD](#) or

[BrowZine](#) to curate a list of journals to follow. However, be aware of confirmation bias. Following only specific journals will limit the breadth of literature exposure.

Social media Social media is a powerful tool for collaboration, education, and learning. Twitter, Facebook, Research Gate, and other social media platforms are being used by many researchers, doctors, and academics to keep up with the latest news and research. If choosing to use these platforms, be sure to follow a variety of news, professional journals, and organization resources to avoid creating a platform for confirmation bias. Keep in mind that benefit can be gained from reading about opposing views and controversial subjects. Additionally, be sure to weigh the risks and balances of using social media before diving in. It is always good to question validity, but this becomes especially important when searching non peer-reviewed material.

Overall, the literature search is a daunting task that is best tackled systematically. Following the five steps outlined in this article will provide an effective approach to starting and completing a literature search.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Crampon JE (1988) Murphy, Parkinson, and Peter - laws for libraries. *Libr J* 113:37–41
- Richardson WS, Wilson MC, Nishikawa J, Hayward RS (1995) The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 123:A12–A13