Cochrane

Cochrane Handbook for Systematic Reviews of Interventions Version 6.5

Technical Supplement to Chapter 4: Searching for and selecting studies

Trusted evidence.
Informed decisions.
Better health.

Table of contents

Technical Supplement to Chapter 4: Searching for and selecting studies

Carol Lefebvre, Julie Glanville, Simon Briscoe, Robin Featherstone, Anne Littlewood, Maria-Inti Metzendorf, Anna Noel-Storr, Robin Paynter, Tamara Rader, James Thomas and L. Susan Wieland on behalf of the Cochrane Information Retrieval Methods Group

This Technical Supplement should be cited as: Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf M-I, Noel-Storr A, Paynter R, Rader T, Thomas J, Wieland LS. Technical Supplement to Chapter 4: Searching for and selecting studies [last updated September 2024]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5. Cochrane, 2024. Available from www.training.cochrane.org/handbook.

Throughout this Technical Supplement we refer to the Methodological Expectations of Cochrane Intervention Reviews (MECIR), which are methodological standards to which all Cochrane Protocols, Reviews, and Updates are expected to adhere. More information can be found on these standards at: https://community.cochrane.org/mecir-manual and, with respect to searching for and selecting studies, in Chapter 4 of the *Cochrane Handbook for Systematic Review of Interventions*.

Further evidence-based information about searching for studies for evidence syntheses can be found on the SuRe Info portal, which is updated twice per year (Isojarvi and Glanville 2021).

1 Sources to search

For discussion of CENTRAL, MEDLINE and Embase as the key database sources to search, please refer to Chapter 4, Section 4.3. For discussion of sources other than CENTRAL, MEDLINE and Embase, please see the sections below. For discussion of some of the specific issues around searching for medical devices, please refer to this recent brief method note (Cooper et al 2022).

1.1 Bibliographic databases other than CENTRAL, MEDLINE and Embase

1.1.1 The Cochrane Register of Studies

The Cochrane Register of Studies (CRS) is a bespoke Cochrane data repository and data management system, primarily used by Cochrane Information Specialists (CISs). The Specialized Registers (registers or databases of trial records on specific topics), maintained by some Cochrane Groups, are stored within the CRS. As such, it acts as a 'meta-register' of all the trials identified by Cochrane (Cochrane Information Specialist Support Team 2021b). The CRS

includes not only the Specialized Registers but also all records of studies, or reports of studies, from the included and excluded sections of Cochrane Reviews. The Cochrane Central Register of Controlled Trials (CENTRAL) is created within the CRS, drawn partly from references from Specialized Registers and partly from references to trial reports sourced from other sources, such as bibliographic databases (e.g. PubMed, Embase and CINAHL). The CRS is the only route available for publication of records in CENTRAL (Cochrane Information Specialist Support Team 2021b).

As a piece of web-based software, the CRS provides tools to manage search activities for Cochrane Reviews. CISs are able to import records from external bibliographic databases and other sources into the CRS, de-duplicate them, share them with author teams and track what has been previously retrieved via searching and screened for each review. A further benefit is that trials register records relating to randomized and quasi-randomized studies (currently from ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP)) are searchable from within the CRS. It is possible to store the full text of each bibliographic citation (and any accompanying documents, such as translations) within the CRS as an attachment but this should always be done in compliance with local copyright and database licensing agreements. Records added to the CRS that are published in CENTRAL are automatically edited in accordance with the Cochrane HarmoniSR guidance, which ensures consistency in record formatting and output (HarmoniSR Working Group 2015). Records from the CRS can be exported in several formats for uploading into Covidence, RevMan or other review management software.

The CRS captures links across references, studies and the Cochrane Reviews within which they appear. This information is drawn from CRS-D, a data repository which sits behind the CRS and includes all CENTRAL records, all included and excluded studies together with ongoing studies, studies awaiting classification and other records collected by CISs in the Specialized Registers. CRS-D has been designed to integrate with RevMan and this linking of data and information back to the reviews will ultimately help review teams find trials more efficiently. For example, CRS-D records can be linked to records in the Reviews Database that powers RevMan, so users can access additional data about the studies that appear in reviews, such as the characteristics of studies, 'Risk of bias' tables and, where possible, the extracted data from the study.

The CRS is a mixture of public records, i.e. CENTRAL records and private records for the use of Cochrane staff only. Full access to the content in CRS is available only to designated staff within Cochrane.

1.1.2 National and regional databases

In addition to MEDLINE and Embase, which are generally considered to be the key international general healthcare databases, many countries and regions produce bibliographic databases that focus on the literature produced in those regions and which often include journals and other literature not indexed elsewhere, such as African Index Medicus and LILACs (for Latin America and the Caribbean). It is highly desirable, for Cochrane Reviews of interventions, that

searches be conducted of appropriate national and regional bibliographic databases (MECIR C25). Searching these databases in some cases identifies unique studies that are not available through searching major international databases (Clark et al 1998, Brand-de Heer 2001, Clark and Castro 2001, Clark and Castro 2002, Abhijnhan et al 2007, Almerie et al 2007, Xia et al 2008, Barnabas et al 2009, Manriquez 2009, Roberts et al 2009, Waffenschmidt et al 2010, Atsawawaranunt et al 2011, Wu et al 2013, Bonfill et al 2015, Cohen et al 2015, Cooper et al 2015, Xue et al 2016). Access to many of these databases is available free of charge. Others are only available by subscription or on a 'pay-as-you-go' basis. Indexing complexity and consistency varies, as does the sophistication of the search interfaces.

For a list of general healthcare databases, see Appendix.

1.1.3 Subject-specific databases

It is highly desirable, for authors of Cochrane Reviews of interventions, to search appropriate subject specific bibliographic databases (MECIR C25). Which subject-specific databases to search in addition to CENTRAL, MEDLINE and Embase will be influenced by the topic of the review, access to specific databases and budget considerations.

Most of the main subject-specific databases such as AMED (allied and complementary medicine), CINAHL (nursing and allied health) and APA PsycInfo (psychology and psychiatry) are available only on a subscription or 'pay-as-you-go' basis. Access to databases is, therefore, likely to be limited to those databases that are available to the Cochrane Information Specialist at the Cochrane Group's editorial base or those that are available at the institutions of the review authors. Access arrangements vary according to institution. Review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist about access at their institution.

Although there is overlap in content coverage across Embase, MEDLINE and CENTRAL and subject-specific databases such as AMED, CINAHL and APA PsycInfo (Moseley et al 2009), their performance and facilities vary (Watson and Richardson 1999a, Watson and Richardson 1999b). In addition, a comparison of British Nursing Index and CINAHL shows that even in databases in a specific field such as nursing, each database covers unique journal titles (Briscoe and Cooper 2014). To find qualitative research, CINAHL and APA PsycInfo should be searched in addition to MEDLINE and Embase (Subirana et al 2005, Wright et al 2015, Rogers et al 2017). Even in cases where research indicates low benefit in searching CINAHL, it is still suggested that for subject-specific reviews it should be considered as an option (Beckles et al 2013, Dhippayom et al 2023).

There are also several studies, each based on a single review, and therefore not necessarily generalizable to all reviews in all topics, showing that searching subject specific databases identified additional relevant publications. It is unclear, however, whether these additional publications would change the conclusions of the review. For example, for a review of exercise therapy for cancer patients, searching CancerLit, CINAHL, and APA PsycInfo identified additional records which were not retrieved by MEDLINE searches but searching SPORTDiscus

identified no additional records (Stevinson and Lawlor 2004); for a review of social interventions, only four of the 69 (less than 6%) relevant studies were found by searching databases such as MEDLINE, while about half of the relevant studies were found by searching the Transport database (Ogilvie et al 2005); in an obesity review, searching the Health Management Information Consortium (HMIC) database identified about one fifth of included publications in addition to MEDLINE searches while CINAHL identified no new publications; and in a tuberculosis review, searching CINAHL identified over 5% of the included publications in addition to MEDLINE, whereas the HMIC database identified no additional publications (Levay et al 2015). A review of database sources for a food science systematic review found that the specialist agriculture and food science databases AGRICOLA and FSTA had the highest precision of all databases searched, but did not return any unique citations alongside Academic Science Premier (ASP), CAB Direct, PubMed and Web of Science (Urhan et al 2019).

For a list of subject-specific healthcare databases, see Appendix.

1.1.4 Citation indexes

(Note: see also section 1.3.6. Other sources for citation searching other than via citation indexes).

Citation indexes are bibliographic databases which index citations in addition to the standard bibliographic content. They were originally developed to identify efficiently the reference lists of scholarly authors and the number of times a reference or author is cited (Garfield 2007). Citation indexes can also be used creatively to identify studies which are similar to a source study, as it is probable that studies which cite or are cited by a source study will contain similar content.

Searching using a citation index is usually called 'citation searching' or 'citation chasing' and is further defined as 'forwards citation searching' or 'backwards citation searching' depending on which direction the citations are searched. Forwards citation searching identifies studies which cite a source study and backwards citation searching identifies studies cited by the source study. Citation indexes are mainly used for forwards citation searching, which is practically impossible to conduct manually, whereas backwards citation searching is relatively easy to conduct manually by consulting reference lists of source studies (see Section 1.3.4). Thus the focus in this section is on forwards citation searching. Citation indexes also facilitate author citation searching which is used to identify studies that are carried out and subsequently published by an author and studies that cite an author.

It is good practice to carry out forwards citation searching on reports of studies that meet the eligibility criteria of a systematic review. Thus forwards citation searching usually takes place after the results of the bibliographic database searches have been screened and a set of potentially includable studies has been identified (Briscoe et al 2020a). Because citation searching is not based on pre-specified terminology it has the potential to retrieve studies that are not retrieved by the keyword-based search strategies that are conducted in bibliographic databases and other resources. This makes citation searching particularly effective in

systematic reviews where the search terms are difficult to define, usefully extending to iterative citation searching of citations identified by citation searching (also known as 'snowballing') in some reported cases (Booth 2001, Greenhalgh and Peacock 2005, Papaioannou et al 2010, Linder et al 2015, Hirt et al 2024). Since researchers may selectively cite studies with positive results, forwards citation searching should be used only as an adjunct to other search methods in Cochrane Reviews (Urlings et al 2021, Hirt et al 2024).

There are varied findings on the efficiency of forwards citation searching, measured as the labour required to export and screen the results of searches relative to the number of unique relevant studies identified (Wright et al 2014, Hinde and Spackman 2015, Levay et al 2016, Cooper et al 2017b). Most studies, however, which compared the results of forwards citation searching with other search methods found that citation searching identified one or more unique studies which were relevant to the review question (Greenhalgh and Peacock 2005, Papaioannou et al 2010, Wright et al 2014, Hinde and Spackman 2015, Linder et al 2015). Reviews of recently published studies, such as review updates, are less likely to benefit from forwards citation searching than reviews with no historical date limit for includable studies due to the relatively limited time for recent studies to be cited. When conducting a review update, however, searchers should consider carrying out forwards citation searching on the studies included in the original review and on the original review itself.

The two main subscription citation indexes are Web of Science, which was launched in 1964 and is currently provided by Clarivate Analytics, and Scopus, which was launched in 2004 by Elsevier. Google Scholar, which was also launched in 2004, can be used for forwards but not backwards citation searching. Microsoft Academic was relaunched in 2015 (Sinha et al 2015) but closed in December 2021. It could be used for both forward and backward citation searching. A new resource, OpenAlex, was launched in early 2022. A summary of these resources is provided below. There are published comparative studies which can be consulted for a more detailed analysis (Kulkarni et al 2009, Wright et al 2014, Levay et al 2016, Cooper et al 2017a).

Web of Science

Web of Science (formerly known as Web of Knowledge), produced by Clarivate Analytics, comprises several databases. The 'Core Collection' databases cover the sciences (1900 to date), social sciences (1956 to date), and arts and humanities (1975 to date). The sciences and social sciences collections are divided into journal articles and conference proceedings, which can be searched separately. In total, as of June 2024, the Web of Science Core Collection contains approximately 2 billion cited references, 90 million records from more than 20,000 journal titles, and books and conference proceedings going back to 1900. Additional databases are available via the Web of Science platform, also on a subscription basis. Author citation searching is possible in Web of Science but it does not automatically distinguish between authors with the same name unless they have registered for a uniquely assigned Web of Science ResearcherID.

https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/web-of-science/web-of-science-core-collection/

Scopus

Scopus, produced by Elsevier, covers health sciences, life sciences, physical sciences and social sciences. As of June 2024, it contains approximately 95 million records from approximately 30,000 active serial titles and approximately 10 million conference abstracts. Records date back to 1788, with approximately 90 million post-1969 records, including references, and approximately 5 million pre-1970 records (Scopus 2024). A unique identification number is automatically assigned to each author in the database which enables it to distinguish between authors with the same names when author citation searching. Errors are still possible, however, as publications are not always assigned correctly to author ID numbers and authors are sometimes erroneously assigned more than one ID number.

https://www.elsevier.com/products/scopus

https://www.elsevier.com/products/scopus/content

Google Scholar

Google Scholar is a freely available scholarly search engine which uses automated web crawlers to identify and index scholarly references, including published studies and grey literature. Although it can only be used for forwards citation searching, this limitation has little practical significance as backwards citation searching can easily be conducted manually by checking reference lists. The precise number of journals indexed in Google Scholar is not known because it does not use a pre-specified list of journals to populate its content. There is, however, evidence that it has sufficient citation coverage to be used as an alternative to Web of Science or Scopus, if these resources are not available (Wright et al 2014, Levay et al 2016).

A disadvantage of Google Scholar's automated study identification method is that it produces more duplicate citations than Web of Science, which indexes pre-specified journal content (Haddaway et al 2015). Scopus, which uses a similar indexing method to Web of Science, is also likely to produce fewer duplicates than Google Scholar. A further disadvantage of Google Scholar is that the export features are basic; however, this can be improved by searching it via the freely available Publish or Perish software (Harzing 2006). Finally, Google Scholar limits the number of viewable results to 1,000 and does not disclose how the top 1,000 results are selected, thus compromising the transparency and reproducibility of search results (Levay et al 2016).

https://scholar.google.com/

OpenAlex

OpenAlex is a tool produced by the non-profit organization OurResearch. In its documentation, OpenAlex is described as a free and open catalogue of the world's scholarly entities, including scholarly works, authors, journals and other repositories, and institutions, and is probably the world's largest dataset of its type, with approximately 250 million works as of June 2024. OpenAlex's first beta data release was in mid-November 2021, positioning itself as a successor to Microsoft Academic, which was retired on 31 December 2021. OpenAlex supports faceted search across many fields (e.g. institution, author, geographical location, journal) and the ability to conduct forwards and backwards citation searches (https://help.openalex.org/coverage).

According to the OpenAlex website, "Using OpenAlex, you can build your own scholarly search engine, recommender service, or knowledge graph. You can help manage research by tracking citation impact, spotting promising new research areas, and identifying and promoting work from underrepresented groups. And you can do research on research itself, in areas like bibliometrics, science and technology studies, and Science of science policy."

https://openalex.org/about

Web of Science, Scopus, Google Scholar and OpenAlex all provide wide coverage of healthcare journal publications. There are, however, differences in the number of records indexed in each citation index and in the methods used to index records, and there is evidence that these differences affect the number of citations which are identified when citation searching (Kulkarni et al 2009, Wright et al 2014, Rogers et al 2016, Rogers et al 2020). It is not a requirement for Cochrane Reviews, however, to conduct forward citation searching. Review authors and information specialists should consider the time and resources available and the likelihood of identifying unique studies for the review question, when planning whether and how to conduct forwards citation searching.

See also section 1.3.6 Other sources for citation searching other than via citation indexes.

Further evidence-based information about the value of citation searching for evidence syntheses can be found in the section entitled 'Value of using different search approaches' on the SuRe Info portal , which is updated twice per year (Isojarvi and Glanville 2021).

1.1.5 Dissertations and theses databases

It is highly desirable, for authors of Cochrane Reviews of interventions, to search relevant grey literature sources such as reports, dissertations, theses, and conference abstracts (MECIR C28). Dissertations and theses are a subcategory of grey literature, which may report studies of relevance to review authors. Searching for academic research published only in the form of dissertations or theses may be important for countering possible publication bias but it can be time consuming and in some cases yield few included studies (van Driel et al 2009). In some areas of health care, searching for and retrieving studies published only as dissertations or theses has been shown to have a limited influence on the conclusions of a review (Vickers and

Smith 2000, Royle et al 2005, Schmucker et al 2017). In other areas of health care, however, it is essential to broaden the search to include trials published in more diverse sources, for example in oncology and in complementary medicine (Egger et al 2003). In a study of 129 systematic reviews from three Cochrane Groups (the Acute Respiratory Infections Group, the Infectious Diseases Group and the Developmental, Psychosocial and Learning Problems Group) there was wide variation in the retrieval and inclusion of dissertations and theses (Hartling et al 2017). It is possible that a study which would affect the conclusions would be missed if the search is not comprehensive enough to include searches for unpublished trials and those reported only in dissertations and theses (Egger et al 2003). The failure to search for unpublished trials and those published only in dissertations and theses, may lead to biased results in some reviews (Ziai et al 2017). Dissertations and theses are not normally indexed in general bibliographic databases such as MEDLINE or Embase, but there are exceptions, such as CINAHL, which indexes nursing, physiotherapy and occupational health dissertations and theses and APA PsycInfo, which indexes dissertations and theses in psychiatry and psychology.

To identify relevant studies published in dissertations or theses it is advisable to search specific dissertation sources:

- The US-based Center for Research Libraries (CRL) is an international consortium of university, college, and independent research libraries (http://catalog.crl.edu/search~S4)
- The LILACS database includes some dissertations and theses from Latin American and Caribbean countries (http://lilacs.bvsalud.org/en/)
- Open Access Theses and Dissertations (OATD) includes dissertations and theses that are free to access and read online from participating universities from around the world (https://oatd.org/). OATD has shown a high level of success at retrieving electronic dissertations and theses (88.5% retrieved) compared to Google Scholar (76% retrieved) and other search engines (Loan et al 2022).
- ProQuest Dissertations and Theses Global (PQDT) is the best-known commercial database for searching dissertations and theses. Access to PQDT is by subscription. As of June 2024, ProQuest Dissertations and Theses Global database indexes 6 million doctoral dissertations and Master's theses from around the world with full text available for approximately 3 million of these records (https://about.proquest.com/en/products-services/pqdtglobal).

Other sources of dissertations and theses include the catalogues and resources produced by national libraries and research centres, for example:

- Australian dissertations and theses are searchable via the National Library of Australia's *Trove* service (http://trove.nla.gov.au/)

- DART-Europe is a partnership of several research libraries and library consortia which provides global access to European research dissertations and theses via a portal. A list of institutions, national libraries and consortia who contribute to the portal can be found here: (https://www.dart-europe.org/)

- Deutsche Nationalbibliothek (German National Library) provides access to electronic versions of dissertations and theses since 1998 (but now also included in DART-Europe – see above) https://www.dnb.de/EN/Professionell/Services/Dissonline/dissonline_node.html).

- The Networked Digital Library of Theses and Dissertations (NDLTD) is an international organization dedicated to promoting the adoption, creation, use, dissemination, and preservation of electronic dissertations and theses (http://search.ndltd.org/).

- Swedish University Dissertations / Dissertations.se offers dissertations and theses in English, about half of which are available to download (http://www.dissertations.se/)

- Theses Canada provides access to the National Library of Canada's records of PhD and Master's dissertations and theses from Canadian universities (https://library-archives.canada.ca/eng/services/services-libraries/theses/Pages/theses-canada.aspx).

Other countries also offer access to dissertations and theses in their national languages.

Whenever possible, review authors should attempt to include all relevant studies of acceptable quality, irrespective of the type of publication, since the inclusion of these may have an impact in situations where there are few relevant studies, or where there may be vested interests in the published literature (Hartling et al 2017, Greyson et al 2019). The inclusion of unpublished trials and those published only in dissertations and theses should increase precision, generalizability and applicability of findings (Egger et al 2003). In the interest of feasibility, review authors should assess their research questions and topic area, and seek advice from content experts when selecting dissertations and theses databases to search. Review authors should consult a Cochrane Information Specialist, local library or university for information about dissertations and theses databases in their country or region.

1.1.6 Grey literature databases

As stated above, it is highly desirable, for authors of Cochrane Reviews of interventions, to search relevant grey literature sources such as reports, dissertations, theses, and conference abstracts (MECIR C28).

Grey literature was defined at GL3, the Third International Conference on Grey Literature on 13 November 1997 in Luxembourg as "that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers" (Farace and Frantzen 1997). On 6 December 2004, at GL6, the Sixth Conference in New York City, a clarification was added: grey literature is "... not controlled by

commercial publishers, i.e. where publishing is not the primary activity of the producing body …" (Farace and Frantzen 2005). In a 2017 audit of 203 systematic reviews published in high-impact factor general medical journals in 2013, 64% described an attempt to search for unpublished studies together with those published only in dissertations and theses. The audit showed that reviews published in the Cochrane Database of Systematic Reviews were significantly more likely to include a search for grey literature than those published in standard journals (Ziai et al 2017). A Cochrane Methodology Review indicated that trials published in journal articles showed an overall greater treatment effect than trials published in grey literature (Hopewell et al 2007a). Although failure to identify trials reported in conference proceedings and other grey literature might affect the results of a systematic review (Hopewell et al 2007a), a subsequent systematic review showed that this was only the case in a minority of reviews (Schmucker et al 2017). Since the impact of excluding unpublished data and data published only in grey literature is unclear, review authors should consider the time and effort spent when planning the grey-literature portion of the search.

Grey literature's diverse formats and audiences can present a significant challenge in a systematic search for evidence. Locating grey literature can often be challenging, requiring librarians and information specialists to use several databases from various providers or websites, with which they may not be familiar (Saleh et al 2014, Haddaway and Bayliss 2015). There are many characteristics of grey literature that make it difficult to search systematically. Further, there is no 'gold standard' for rigorous systematic grey literature search methods and few resources on how to conduct this type of search (Godin et al 2015, Paez 2017). One challenge of searching the grey literature is managing an abundance of material. Often, there are many sources to search but some review authors of very broad or cross-disciplinary topics may find it necessary to impose some limits on the extent of their grey literature searching by considering what is feasible within limited time and resources (Mahood et al 2014). For example, since nearly half of the citations found in reviews of new and emerging non-drug technologies are published in grey literature, searchers should consider focusing their efforts on search engines and aggregator sites to increase feasibility (Farrah and Mierzwinski-Urban 2019). Google Scholar can help locate a large volume of grey literature and specific, known studies, however, it should not be used as the only resource for systematic review searches (Haddaway et al 2015). The types of grey literature that are useful in specific reviews may depend on the research question and researchers may decide to tailor the search to the question (Levay et al 2015). For example, academic research published only in grey literature may be important for countering possible publication bias and can be targeted via specific repositories for preprints, theses and funding registries. Alternatively, if the research question is related to implementation or if the researchers are interested in material to support their implications for practice section, then organizational reports, government documents and monitoring and evaluation reports, might be important for ensuring the search is extensive and fit for purpose (Haddaway and Bayliss 2015).

The inclusion of grey literature can help to overcome the time lag between novel developments in health research and the publication of findings in scientific journals. Since the amount of

data in grey literature can be overwhelming, the use of automated data extraction tools holds potential to manage the volume of material (Schmidt et al 2024).

Careful documentation throughout the search process and reporting of search methods will demonstrate that efforts have been made to be comprehensive and will help in making the grey literature searching as reproducible as possible (Stansfield et al 2016).

The following resources can help authors plan a manageable and thorough approach to searching the grey literature for their topic.

- Canada's Drug Agency (CDA-AMC), previously known as CADTH (https://greymatters.cadth.ca/), publishes a resource entitled 'Grey Matters: a tool for searching health-related grey literature', (https://greymatters.cadth.ca/), which lists a considerable number of grey literature sources together with annotations about their content as well as search hints and tips.

- The Health-Related Grey Literature guide (http://www.greylitguides.com/health-related-grey-lit/) is a source of health-related grey literature, organized geographically. It is part of Grey Literature Guides, a directory of research guides, online courses and webinars. The website provides links to a selection of existing English language educational and training resources devoted to grey literature worldwide. It is maintained by GreyNet's Grey Literature Education and Training Committee and revised annually.

- The Health Management Information Consortium (HMIC) Database (https://www.kingsfund.org.uk/consultancy-support/library-services; https://www.wolterskluwer.com/en/solutions/ovid/hmic-database-99) contains records from the Library and Information Services department of the UK Department of Health (DH Data) and the King's Fund Information and Library Service. It includes all UK Department of Health publications including circulars and press releases. The King's Fund is an independent health charity that works to develop and improve management of health and social care services. The database is considered to be a good source of grey literature on topics such as health and community care management, organizational development, inequalities in health, user involvement, and race and health. The King's Fund Information and Library Service records can be searched free of charge via the link above. The UK Department of Health data can be searched on subscription only.

- The US National Technical Information Service (NTIS) (https://www.ntis.gov) provides access to the results of both US and non-US government-sponsored research and can provide the full text of the technical report for most of the results retrieved. NTIS is free of charge on the internet and goes back to 1964. For access to technical reports see the National Technical Reports Library at https://ntrl.ntis.gov/NTRL/; for access to the NTIS Bibliographic Database, log in is required.

- OpenGrey was a multidisciplinary European grey literature database, covering science, technology, biomedical science, economics, social science and humanities. Each record had an English title and/or English keywords. Some records included an English abstract (starting in 1997). The database included technical or research reports, doctoral dissertations, conference presentations, official publications, and other types of grey literature. Information was also provided regarding how to access the documents included in the database. Access to this database via Inist-CNRS ceased in November 2020, but a searchable archived version is available from the Data Archiving and Networked Services (DANS) Easy system (https://ssh.datastations.nl/dataset.xhtml?persistentId=doi:10.17026/dans-xtf-47w5).

- APA PsycExtra (http://www.apa.org/pubs/databases/psycextra/) is a companion database to APA PsycInfo in psychology, behavioural science and health. It includes references from newsletters, magazines, newspapers, technical and annual reports, government reports and consumer brochures. APA PsycExtra is different from APA PsycInfo (https://www.apa.org/pubs/databases/psycinfo/index) in its format, because it includes abstracts and citations plus full text for a major portion of the records. There is no coverage overlap between APA PsycExtra and APA PsycInfo.

Conference abstracts are a particularly important source of grey literature and are further covered in Section 1.3.3 Dissertations and theses are covered above in Section 1.1.5.

1.2 Ongoing studies and unpublished data sources: further considerations

This section should be read in conjunction with Chapter 4, Sections 4.3.2 to 4.3.4.

1.2.1 Trials registers and trials results registers

It is mandatory, for authors of Cochrane Reviews of interventions, to search trials registers and repositories of results, where relevant to the topic, through ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform (ICTRP) portal and other sources as appropriate (MECIR C27) (see Chapter 4, Section 4.3.3). Although ClinicalTrials.gov is included as one of the registers within the WHO ICTRP portal, it is recommended that both ClinicalTrials.gov and the ICTRP portal should be searched separately, from within their own interfaces, due to additional features in ClinicalTrials.gov (Glanville et al 2014) (see below). The extent to which this might still be the case with the new ICTRP interface released in its final version in June 2021 and the changes being made under the ClinicalTrials.gov modernization programme remains to be ascertained. Therefore, the guidance that it is not sufficient to search the ICTRP alone still stands, pending further research. It is recognized, however, that the search interfaces of many of these trials registers and portals is suboptimal, resulting in challenges for those searching them (Cooper et al 2021b).

Several initiatives have led to the development of and recommendations to search trials registers. The International Committee of Medical Journal Editors (ICMJE) requires prospective registration of studies for subsequent publication in their journals, and there is a legal

requirement that the results of certain studies must be posted within a given timeframe. It must be noted, however, that there is still no global, legal, universal requirement to register clinical trials at inception or at any other stage in the process (only funder-specific, country-specific etc.) and no global, legal, universal requirement to post/publish trial results, although the situation is improving. Several studies have shown that adherence to these requirements, where they exist, is mixed (Gill 2012, Huser and Cimino 2013a, Huser and Cimino 2013b, Jones et al 2013, Anderson et al 2015, Dal-Re et al 2016, Goldacre et al 2018, Jorgensen et al 2018, Loder et al 2018, DeVito et al 2020, Ramachandran et al 2021, Chen et al 2022, Lauer 2023) and that results posted on ClinicalTrials.gov show discordance when compared with results published in journal articles (Gandhi et al 2011, Earley et al 2013, Hannink et al 2013, Becker et al 2014, Hartung et al 2014, De Oliveira et al 2015) or both of the above (Jones and Platts-Mills 2012, Adam et al 2018, Talebi et al 2020). A recent study indicated that, despite the initiatives mentioned above, many systematic reviews in the field of critical care failed to include searches of trials registers (Greiner et al 2021). A recent analysis of ClinicalTrials.gov registration data from 2000-2020 showed that the number and percentage of registered trials reporting results had increased since the inception of ClinicalTrials.gov, peaking in 2007 (n=2,840, 36%) when the ClinicalTrials.gov results database was launched, compared to previous years ranging from 8.7 to 24.5% of trials with posted results (Gresham et al 2022). The time to report results had also improved over time, decreasing from a median of 29 months in 2007, when the result database was launched, to 12 months in 2015, and 10 months in 2020. A recent study, however, highlighted failings in the then current implementation of ClinicalTrials.gov (the 'classic version') including the lack of MeSH terms for interventions and outcomes, the fact that synonyms in searches could not be toggled off and that inclusion criteria and exclusion criteria were in the same field (Miron et al 2020). Further guidance on searching trials registers and related portals can be found in a recent article (Hunter et al 2022). There are few but conflicting studies assessing whether including study results from trials registers significantly alters the effect estimates in meta-analyses (Baudard et al 2017, Bagg et al 2020, Alqaidoom et al 2023).

ClinicalTrials.gov

In February 2000, the US National Library of Medicine (NLM) launched ClinicalTrials.gov (https://clinicaltrials.gov). ClinicalTrials.gov was created as a result of the Food and Drug Administration Modernization Act of 1997 (FDAMA). FDAMA required the US Department of Health and Human Services, through the US National Institutes of Health (NIH), to establish a register of clinical trials information for both (US) federally and privately funded trials conducted under 'investigational new drug' applications to test the effectiveness of experimental drugs for "serious or life-threatening diseases or conditions". The ClinicalTrials.gov registration requirements were expanded after the US Congress passed the FDA Amendments Act of 2007 (FDAAA). Section 801 of FDAAA (FDAAA 801) required more types of trials to be registered and additional trial registration information to be submitted. The law also required the submission of results for certain trials. This led to the expansion of ClinicalTrials.gov to include information on study participants and a summary of study

outcomes, including adverse events. Results have been made available since September 2008. Further legislation has expanded the coverage of results in ClinicalTrials.gov, which now serves as a major international register including more than 500,000 study records from over 200 countries (as of June 2024). Searches of ClinicalTrials.gov can be limited to studies which include results by selecting 'With results' under the 'Study Results' filter under More Filters on the home page of the 'modernized' ClinicalTrials.gov site, which became the primary site for ClinicalTrials.gov in June 2023 and fully replaced the 'classic' site on 25 June 2024, when the 'classic' site was finally retired. Research published in 2014 showed that the most reliable way of searching ClinicalTrials.gov was to conduct a highly sensitive 'single concept' search in the basic interface of ClinicalTrials.gov (Glanville et al 2014). This study also suggested that use of the advanced interface seemed to improve precision without loss of sensitivity and this interface might be preferred when large numbers of search results were anticipated. As mentioned above, the new ICTRP interface released in its final version in June 2021 and the changes being made under the ClinicalTrials.gov Modernization programme remain to be ascertained.

Search help for ClinicalTrials.gov (modernized site) is available from the following links:

How to Search for Clinical Studies

https://clinicaltrials.gov/find-studies/how-to-search

How to Read a Study Record

https://clinicaltrials.gov/study-basics/how-to-read-study-record

How to Read Study Results

https://clinicaltrials.gov/study-basics/how-to-read-study-results

How to Use Search Results

https://clinicaltrials.gov/find-studies/how-to-use-search-results

How to Search for Studies with Results

https://clinicaltrials.gov/find-studies/how-to-search-for-studies-with-results

Constructing Complex Search Queries

https://clinicaltrials.gov/find-studies/constructing-complex-search-queries

RSS Feeds (for all the studies on ClinicalTrials.gov or for a specific search)

https://clinicaltrials.gov/find-studies/rss

ClinicalTrials.gov Demonstration Videos

https://www.nlm.nih.gov/oet/ed/ct/demo_videos.html

The World Health Organization International Clinical Trials Registry Platform search portal (ICTRP)

In May 2007, the World Health Organization (WHO) launched the International Clinical Trials Registry Platform (ICTRP) search portal (https://trialsearch.who.int/), to search across a range of trials registers, similar to the initiative launched some years earlier by Current Controlled Trials with their '*meta*Register' (which has ceased publication). Currently (June 2024), the WHO portal searches across c. 20 registers (including ClinicalTrials.gov but note the guidance above regarding searching ClinicalTrials.gov separately through the ClinicalTrials.gov interface). Research has shown that the most reliable way of searching the ICTRP was to conduct a highly sensitive 'single concept' search in the ICTRP basic interface (Glanville et al 2014). This study suggested that use of the ICTRP advanced interface might be problematic because of reductions in sensitivity. The extent to which this might still be the case with the new ICTRP interface, released in its final version in June 2021, remains to be ascertained.

Search help for the ICTRP is available from the following link:

https://www.who.int/clinical-trials-registry-platform/the-ictrp-search-portal/search-tips

Other trials registers, trials register resources and trials results resources

In May 2021, the UK National Institute for Health Research (NIHR) Innovation Observatory launched ScanMedicine (https://scanmedicine.com/), a resource which draws records from 14 national and international trials register resources with information on drugs, devices and diagnostics together with digital applications approved by the FDA, enabling searches back to 1995 (https://www.nihr.ac.uk/news/nihr-launches-innovative-searchable-database-of-global-clinical-trials/27660)(Sadek et al 2023). As of June 2024, ScanMedicine held c. 700,000 records in the 'clinical trials' section of the database and 200,000 records in the devices section of the database. A small help file is available by hovering over the 'i' symbol at the end of the search box on the home page.

The European Clinical Research Infrastructure Network (ECRIN) produces the Clinical Research Metadata Repository (MDR), which the producers claim is "a searchable database of all registered clinical studies, together with links to the source registry pages, the results entries or results summary files, linked papers, protocols, data collection forms and many other related documents, and data sets, whenever those objects are available" (https://crmdr.ecrin.org/).

A detailed User Guide is available at: https://crmdr.ecrin.org/Guide.

https://ecrin.org/clinical-research-metadata-repository

To search, go to: https://crmdr.ecrin.org/

There is a detailed MDR Wiki at: https://crmdr.ecrin.org/About

OZMOSI produces the Global Clinical Trial (GCT) website which provides access to information on c. 500,000 clinical trials. Records can be downloaded into an Excel spreadsheet.

https://www.ozmosi.com/global-clinical-trial-data/

HSRProj (Health Services Research Projects in Progress) provided information about ongoing health services research and public health projects. It contained descriptions of research in progress funded by US federal and private grants and contracts for use by policy makers, managers, clinicians and other decision makers. It provided access to information about health services research in progress before, and irrespective of whether, results were available in a published form. In June 2021, the US National Library of Medicine (NLM) announced that they would discontinue HSRProj from September 2021. HSRProj data are now archived and downloadable as below.

https://healthdata.gov/dataset/Health-Services-Research-Projects-in-Progress-HSRP/u8mi-83iu/data

https://wayback.archive-it.org/7189/20210912160016/https:/hsrproject.nlm.nih.gov/

Many countries and regions maintain trials and/or trials results registers. There are also many condition-specific trials registers, especially in the field of cancer, which are too numerous to list. Some pharmaceutical companies and device manufacturers make available information about their clinical trials through their own websites, either instead of or in addition to the information they make available through national or international registers or websites. Additionally, there are commercially produced trials registers, which are available on a subscription basis.

Clinical Trial Results (https://clinicaltrialresults.org/) is a website that hosts slide and video presentations from clinical trialists, especially in the field of cardiology but also other specialties, reporting the results of clinical trials.

Further listings of international, national, regional, subject-specific and industry trials registers, together with guidance on how to search them can be found on a website developed in 2009 and since then updated by two of the co-authors of this chapter (JG and CL) entitled Finding clinical trials, research registers and research results (https://sites.google.com/a/york.ac.uk/yhectrialsregisters/).

1.2.2 Regulatory agency sources and clinical study reports

Regulatory agencies serve as sources of trial records by producing trials registers and also as a source of clinical study reports and related documents. Both these types of regulatory information are discussed below.

The EU Clinical Trials Register (EU CTR)

The EU CTR contains protocol and results information for interventional clinical trials on medicines, conducted in the European Union (EU) and the European Economic Area (EEA), which started after 1 May 2004. It enables searching for information in the EudraCT database, used by national medicines regulators for data related to clinical trial protocols. Results data are extracted from data entered by the sponsors into EudraCT. The EU CTR has been a 'primary registry' in the ICTRP since September 2011 but in the absence of any evidence to the contrary, it is recommended that searches of the EU CTR should be carried out within the EU CTR and not solely within the ICTRP (in line with the advice above regarding searching ClinicalTrials.gov). A recent technical review, however, used a validated checklist to assess the search interface of EU CTR (compared with ClinicalTrials.gov and the ICTRP) and found low overall scores for each of the interfaces, with the EU CTR performing the worst (Cooper et al 2021b). The register currently (June 2024) contains information about approximately 60,000 clinical trials. Searches can be limited to 'Trials with results' under the 'Results Status' option. Records can be selected individually for downloading or can be downloaded one page at a time (maximum 20 records). The posting of clinical trial summary results became mandatory in 2014. There is no information on non-interventional clinical trials (e.g. observational studies), clinical trials for surgical procedures, medical devices or psychotherapeutic procedures. As of 31 January 2023, all EU/EEA initial clinical trial applications must be submitted through the European Clinical Trials Information Service (CTIS) for the EU Clinical Trials database. See also the paragraph below on the European Clinical Trials Information Service (CTIS).

https://www.clinicaltrialsregister.eu/ctr-search/search

https://www.clinicaltrialsregister.eu/about.html

The European Clinical Trials Information Service/the EU Clinical Trials database

https://euclinicaltrials.eu/search-for-clinical-trials/?lang=en

The European Union (EU) has launched a new database called the EU Clinical Trials Database as part of a newly-established EU Clinical Trials Information System (CTIS). As noted above, from 31 January 2023, all initial clinical trial applications in the EU must be submitted via the CTIS. The CTIS includes the EU Clinical Trials Database as a public, searchable database. By 31 January 2025, all ongoing trials that were approved under the EU Clinical Trials Directive will be governed by the new Regulation and will have to be transitioned to the CTIS. The website currently contains limited information on clinical trials entered since its launch on 31 January 2022 (c. 4,000 as of June 2024). It will gradually contain more information as clinical trial sponsors and EU/EEA regulators use it to initiate and oversee clinical trials in the EU and the EEA. Search tips and guidance are available from: https://euclinicaltrials.eu/search-tips-and-guidance/?lang=en.

https://euclinicaltrials.eu/about-this-website/?lang=en

https://www.ema.europa.eu/en/human-regulatory/research-development/clinical-trials/clinical-trials-information-system

For information about clinical trials initiated before 31 January 2022 and those initiated during the transition phase from 2022 to 2025, it will be necessary to continue searching the EU Clinical Trials Register (EU CTR) described above, to ensure sufficient overlap.

https://euclinicaltrials.eu/about-this-website/?lang=en#transition-period

The European Database on Medical Devices (EUDAMED) – under development

The EU is in the process of developing and introducing a database for medical devices, known as EUDAMED. It has been established as part of Regulation (EU) 2017/745 on medical devices and Regulation (EU) 2017/746 on in vitro diagnosis medical devices. It is expected that the database will be fully functioning and mandatory by the end of 2026. As of June 2024, there were c. 800 certificates in the database (including those issued, refused etc.).

https://health.ec.europa.eu/medical-devices-eudamed/overview_en

https://health.ec.europa.eu/medical-devices-eudamed/overview_en

https://ec.europa.eu/tools/eudamed/#/screen/home

Drugs@FDA and medical device information from the FDA

Drugs@FDA is hosted by the US Food and Drug Administration (FDA) and provides information about most of the drugs approved in the US since 1939. For those approved more recently (from 1998), there is often a 'Review', which is an internal review containing the scientific analyses that provided the basis for approval of the new drug (see Glossary link below). In 2012, new search options were introduced, enabling search strategies to be saved and re-run and results to be downloaded to a spreadsheet (Goldacre et al 2017). Guidance on how to use FDA drug approval documents for evidence syntheses is available(Ladanie et al 2018).

https://www.accessdata.fda.gov/scripts/cder/daf/

https://www.fda.gov/drugs/drug-approvals-and-databases/about-drugsfda

https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=glossary.page

Information on how to search Drugs@FDA is available from the Frequently Asked Questions pages at: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=faq.page

The FDA also makes information about devices, including several medical device databases, available on its website:

https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/medical-device-databases

Clinical study reports

Clinical study reports (CSRs) are reports of clinical trials, which provide detailed information on the methods and results of clinical trials submitted in support of marketing authorization applications. Cochrane funded a project under the Methods Innovation Funding programme to draft interim guidance to help Cochrane Review authors decide whether to include data from clinical study reports (CSRs) and other regulatory documents in a Cochrane Review (Hodkinson et al 2018, Jefferson et al 2018).

A Clinical Study Reports Working Group has been established in Cochrane to take this work forward and to consider how CSRs might be used in Cochrane Reviews in future. To date, only one Cochrane Review is based solely on CSRs, that is the 2014 review update on neuraminidase inhibitors for preventing and treating influenza in healthy adults and children (Jefferson et al 2014).

Although pharmaceutical companies are obliged to submit details of all clinical trials within their clinical study reports, studies have found that this is not always the case (Boesen et al 2022). A recent study compared the availability of CSRs from Canada Health, the EMA and the FDA, including a comparison of the data made available by the various agencies and the time taken to make this information available (Egilman et al 2021). A further study assessed the accessibility of CSRs from industry-sponsored clinical trials whose results were reported in the FDA-authorized drug labels for the top 30 highest-revenue medicines of 2021 and found that CSRs were available for only 69% of the clinical trials supporting regulatory approval of the 30 medicines sampled. Moreover, only 22% of the CSRs were directly downloadable from regulatory agencies (i.e. the EMA and Health Canada)(Hopkins et al 2024).

Clinical study reports from the European Medicines Agency (EMA)

In late 2010, the European Medicines Agency (EMA) began releasing CSRs (on request) under their Policy 0043. In October 2016, they led the field by beginning to release CSRs under their Policy 0070. The policy applies only to documents received since 1 January 2015. CSRs are available for approximately 300 products (as of June 2024) (https://clinicaldata.ema.europa.eu/web/cdp/background).

https://register.ema.europa.eu/identityiq/login.jsf

https://clinicaldata.ema.europa.eu/web/cdp/search

In December 2018, the Agency suspended the publication of clinical data as a result of the implementation of the third phase of the EMA's business continuity plan (i.e. Brexit and the resulting transfer of the EMA offices from London to Amsterdam) and it remained suspended due to ongoing business continuity linked to the COVID-19 pandemic (except for COVID-19 trials) until late 2023. The EMA continued publishing clinical data for COVID-19 medicines during this time, in line with its exceptional transparency measures for treatments and vaccines for COVID-19. As noted above, as of June 2024, there were approximately 300 CSR

records with publication dates from October 2016 to date. Publication has now resumed for non-COVID trials prospectively from late 2023 but there are no plans to fill the gap back to 2018.

Paludan-Müller et al (Paludan-Müller et al 2022) conducted a study assessing the content and characteristics of all clinical data packages released by the EMA under Policy 0070 and the time to their publication (i.e. 148 clinical data packages that contained data on a total of 1,005 clinical trials, of which 261 (26%) were labelled as phase 3 trials). Full CSRs were available for 913 (90.8%) of the trials. The median time to publication was 511 (IQR 411 to 574) days. Only 2 (1.4%) of the clinical data packages were published within the EMA's planned timeline.

In order to download the full CSR documents, it is necessary to register for use "for academic and other non-commercial research purposes" and to provide an email address and a place of address in the European Union, or provide details of a third party, resident or domiciled in the European Union, who will be considered to be the user.

https://clinicaldata.ema.europa.eu/web/cdp/termsofuse

Clinical study reports from the Food and Drug Administration (FDA)

The FDA does not currently routinely provide access to CSRs, only their own internal reviews, as noted above. In January 2018, they announced a voluntary pilot programme to disclose up to nine recently approved drug applications, limited to CSRs for the key 'pivotal' trials that underpin drug approval (Doshi 2018). They succeeded, however, in only covering one drug application within this pilot programme, which has now ended. A public consultation of this pilot project (which included only one CSR) was undertaken in August 2019.

https://www.fda.gov/drugs/development-approval-process-drugs/clinical-data-summary-pilot-program

The FDA subsequently announced that "...increasing international harmonization efforts to share clinical study reports is a long-term goal".

https://www.fda.gov/news-events/press-announcements/fda-continues-support-transparency-and-collaboration-drug-approval-process-clinical-data-summary

Clinical study reports from Health Canada

In April 2019 Health Canada announced that it was starting to make clinical information (i.e. CSRs) about drugs and devices publicly available on its website (https://clinical-information.canada.ca/search/ci-rc) (Lexchin et al 2019). As of June 2024, information was available for approximately 600 drug records and 100 medical device records.

Other related information from other regulatory agencies

Australia: the Therapeutic Goods Administration (Australia) (TGA) provides access to Australian Public Assessment Reports for prescription medicines (AusPARs) but not the full CSRs – approximately 1,000 records as of June 2024.

https://www.tga.gov.au/resources/auspar

Japan: The Japanese Pharmaceuticals and Medical Devices Agency (PMDA) also provides access to its own internal reviews of approved drugs and medical devices but not the original CSRs. These can be found in the Reviews section of its website at:

https://www.pmda.go.jp/english/review-services/reviews/0001.html

https://www.pmda.go.jp/english/review-services/reviews/approved-information/drugs/0001.html

For more information on drug and device information from regulatory sources see: Restoring Invisible and Abandoned Trials (RIAT) initiative website:

https://restoringtrials.org/regulatory-resources/

1.3 Journals and other non-bibliographic database sources

1.3.1 Handsearching

Handsearching involves a manual page-by-page examination of the entire contents of a journal issue or conference proceedings to identify all eligible reports of trials (for discussion of 'handsearching' full-text journals available electronically, see Section 1.3.2). In journals, reports of trials may appear in articles, abstracts, news columns, editorials, letters or other text. Handsearching has a specific meaning and it should not be used as a synonym for reference checking of known relevant records (see online Technical Supplement Section 1.3.4).

Handsearching healthcare journals and conference proceedings can be a useful adjunct to searching electronic databases for at least two reasons: 1) not all trial reports are included in electronic bibliographic databases, and 2) even when they are included, they may not contain relevant search terms in the titles or abstracts or be indexed with terms that allow them to be easily identified as trials (Dickersin et al 1994). It should be noted, however, that handsearching is not a requirement for all Cochrane Reviews and review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist with respect to whether handsearching might be valuable for their review, and if so, what to search and how (Cochrane Information Specialist Support Team 2021a). Methods of identifying which journals to handsearch and evidence around the usefulness of handsearching are summarized in an overview of published guidance (Cooper et al 2017a). Each journal year or conference proceeding that is to be handsearched should be searched thoroughly and competently by a well-trained handsearcher, ideally for all reports of trials, irrespective of topic, so that once it has been handsearched it will not need to be searched again. A Cochrane Methodology Review found that a combination of handsearching and electronic searching is necessary for full identification of relevant reports published in journals, even for those that are indexed in MEDLINE (Hopewell et al 2007b). This was especially the case for articles published before 1991 when there was no indexing term for randomized trials in

MEDLINE and for those articles that are in parts of journals (such as supplements and conference abstracts) which are not routinely indexed in databases such as MEDLINE. A review (Richards 2008) found that handsearching was valuable for finding trials reported in abstracts or letters, or in languages other than English. We note that Embase is now a good source of conference abstracts (see Section 1.3.3).

To facilitate the identification of all published trials, Cochrane has organized extensive handsearching efforts. Over 3,000 journals have been searched within Cochrane. The list of journals that have already been handsearched is no longer being updated, but a record of journals searched, with the dates covered by the search is available via the Handsearched Journals tab in the Cochrane Register of Studies Online at https://crso.cochrane.org/ (Cochrane Account login required). Citations for the handsearched records are available in CENTRAL.

With respect to handsearching conference abstracts or proceedings, this should still be considered for finding studies which may be published only as conference abstracts. Embase has approximately 5 million conference abstracts, but searches of Embase will not necessarily find all the trial records in a conference issue (Stovold and Hansen 2011, Cooper et al 2020). Coverage of specific conferences of interest can be ascertained by checking the list of conferences indexed in Embase: https://www.elsevier.com/products/embase/content?trial=true.

Cochrane groups and authors can prioritize handsearching based on where they expect to identify the most trial reports. This prioritization can be informed by searching CENTRAL, MEDLINE and Embase in a topic area and identifying which journals appear to be associated with the most retrieved citations. Preliminary evidence suggests that most of the journals with a high yield of trial reports are indexed in MEDLINE (Dickersin et al 2002) but this may reflect the fact that Cochrane contributors have concentrated early efforts on searching these journals. Therefore, journals not indexed in MEDLINE or Embase should also be considered for handsearching. Research into handsearching journals in a range of languages suggests that handsearching journals published in languages other than English is still helpful for identifying trials which have not been retrieved by database searches (Blumle and Antes 2005, Fedorowicz et al 2005, Al-Hajeri et al 2006, Nasser and Al Hajeri 2006, Chibuzor and Meremikwu 2009). The value of handsearching may vary from topic to topic. In physiotherapy and respiratory disease, studies have found handsearching yielded additional studies (Stovold and Hansen 2011, Craane et al 2012). Identifying studies of handsearching in specific disease areas may help to inform decisions around handsearching.

The Cochrane Training Manual for Handsearchers is available on the Cochrane Information Retrieval Methods Group Website: http://methods.cochrane.org/irmg/resources.

Handsearching may be facilitated by tools such as Paperfetcher or programs written in R or Python, which may automate the record collection process (Pallath and Zhang 2023).

1.3.2 Full text journals available electronically

The full text of many journals is available electronically on the internet. Access may be partially or wholly on a subscription basis or free of charge. In addition to providing a convenient method for retrieving the full article of already identified records, full-text journals can also be searched electronically, depending on the search interface, by entering relevant keywords in a similar way to searching for records in a bibliographic database. Electronic journals can also be 'handsearched' in a similar manner to that advocated for journals in print form, in that each screen or 'page' can be checked for possibly relevant studies in the same way as handsearching a print journal (see Section 1.3.1). When reporting handsearching, it is important to specify whether the full text of a journal has been searched electronically or using the print version. Some journals omit sections of the print version, for example letters, from the electronic version and some include supplementary information such as extra articles in the electronic format only.

Most academic institutions subscribe to a wide range of electronic journals and these are therefore available free of charge at the point of use to members of those institutions. Review authors should seek advice about electronic journal access from the library service at their institution. Some professional organizations provide access to a range of journals as part of their membership package. In some countries similar arrangements exist for health service employees through national licences.

Several international initiatives provide free or low-cost online access to full-text journals over the internet. The Health InterNetwork Access to Research Initiative (HINARI) programme, set up by the World Health Organization (WHO) together with major publishers and now part of the Research4Life programme (R4L), provides access to a wide range of resources, including journals, for healthcare professionals in local, not-for-profit institutions in more than 120 countries, areas and territories. The International Network for the Availability of Scientific Publications (INASP) also provides access to a wide range of journals (and databases). Electronic Information for Libraries (EIFL) is a similar initiative based on library consortia to support affordable licensing of journals and other sources in over 50 developing and transition countries in Africa, Asia, Europe and Latin America.

A local electronic or print copy of any possibly relevant article as well as its online supplementary material found in a subscription journal should be taken and filed (within copyright legislation), as the subscription to that journal may cease. The same applies to electronic journals available free of charge, as the circumstances around availability of specific journals might change. We have not been able to identify any research evidence regarding searching full-text journals available electronically. Review authors are not routinely expected to search full-text journals available electronically for their reviews, but they should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, as to whether, in their particular case, this might be beneficial.

1.3.3 Conference abstracts and proceedings

It is highly desirable, for authors of all Cochrane Reviews of interventions, to search relevant databases of conference abstracts (MECIR C28). Some bibliographic databases do include conference abstracts. While MEDLINE does index some conference abstracts and proceedings (to identify these, use Congress, Meeting Abstracts and Overall as the Publication Types), Embase is a better source with about 5 million conference abstracts from about 15,000 conferences (as of June 2024).

As noted above, Elsevier provides a list of conferences it indexes in Embase. As a result of Cochrane's Embase project (see Section 2.1.2), conference abstracts that are indexed in Embase and are reports of RCTs are now being included in CENTRAL. Other conference abstracts resources include the Web of Science Conference Proceedings Citation Index (https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/web-of-science/web-of-science-core-collection/conference-proceedings-citation-index/) and Northern Light Life Science Conference Abstracts (https://northernlight.com/life-sciences-conference-abstracts/). Additionally, many conference proceedings are published as journal supplements or as proceedings on the website of the conference or the affiliated organization. A Cochrane Methodology Review found that trials with positive results tended to be published in approximately four to five years whereas trials with null or negative results were published after about six to eight years (Hopewell et al 2007c) and not all conference presentations are published or indexed (Slobogean et al 2009). Over one-half of trials reported in conference abstracts never reach full publication (Diezel et al 1999, Scherer et al 2018) and those that are eventually published in full have been shown to have results that are systematically different from those that are never published in full (Scherer et al 2018). In addition, conference abstracts/proceedings are a good source to track disagreements between the original abstract and the full report of studies (known as reporting bias) (Chokkalingam et al 1998, Pitkin et al 1999, Saric et al 2019). Trials with positive findings are more likely to be published than those which do not have positive findings (known as publication bias) (Salami and Alkayed 2013, Treanor et al 2020, Tumin et al 2020). It is, therefore, important to try to identify possibly relevant studies reported in conference abstracts through specialist database sources and by searching those abstracts that are made available on the Internet, on CD-ROM/DVD or in print form. Controversies around the usefulness of searching for conference abstracts include: whether the reported information is dependable; poor CONSORT abstract reporting quality; and the high screening burden with likely few includable studies (van Driel et al 2009, Schmucker et al 2017, Scherer and Saldanha 2019, Hackenbroich et al 2022).

1.3.4 Other reviews, guidelines and reference lists as sources of studies

It is highly desirable, for authors of Cochrane Reviews of interventions, to search within previous reviews on the same topic (MECIR C29) and it is mandatory, for authors of Cochrane Reviews of interventions, to check reference lists of included studies and any relevant systematic reviews identified (MECIR C30). Reviews can provide relevant studies and references, and may also provide information about the search strategy used, which may

inform the current review (Hunt and McKibbon 1997, Glanville and Lefebvre 2000). Copies of previously published reviews on, or relevant to, the topic of interest should be obtained and checked for references to the included (and excluded) studies. Various sources for identifying previously published reviews are described below.

As well as the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Library used to include the Database of Abstracts of Reviews of Effects (DARE) and the Health Technology Assessment Database (HTA Database), produced by the Centre for Reviews and Dissemination (CRD) at the University of York in the UK. Both databases provided information on published reviews of the effects of health care (Petticrew et al 1999). Searches of MEDLINE, Embase, CINAHL, APA PsycInfo and PubMed to identify candidate records for these two databases were continued until the end of 2014 and bibliographic records were published on DARE until 31 March 2015. CRD will maintain secure archive versions of DARE until at least the end of December 2024 (https://www.crd.york.ac.uk/CRDWeb). CRD continued to maintain and add records to the HTA database until 31 March 2018. In July 2019, the HTA database records were exported from the CRD platform and imported into the new platform that was developed by the International Network of Agencies for Health Technology Assessment (INAHTA). The rebuild of the new platform was launched in June 2020. The International HTA database (https://www.inahta.org/hta-database/) provides free access to bibliographic information of approximately 20,000 ongoing and completed/published health technology assessments commissioned or undertaken by HTA organizations internationally (as of June 2024).

Since 1 April 2015 the UK National Institute for Health and Care Research (NIHR) has produced summaries of single NIHR research studies (Alerts) and summaries of several NIHR research studies within a specific theme or health and care topic (Collections and Themed Reviews). Details can be found at https://evidence.nihr.ac.uk/.

KSR Evidence, a subscription database, aims to include all systematic reviews and meta-analyses published since 2015 (https://ksrevidence.com/). KSR Evidence was developed by Kleijnen Systematic Reviews Ltd (KSR) (www.systematic-reviews.com). KSR produces and disseminates systematic reviews, cost-effectiveness analyses and health technology assessments of research evidence in health care. The database also includes an advanced search option, suitable for information specialists.

CRD provides an international register of prospectively registered systematic reviews in health and social care called PROSPERO (Page et al 2018), which (as of June 2024) contained over 270,000 records (www.crd.york.ac.uk/prospero/). Key features from the review protocol are recorded and maintained as a permanent record. PROSPERO aims to provide a comprehensive listing of systematic reviews registered at inception to help avoid duplication and reduce opportunity for reporting bias by enabling comparison of the completed review with what was planned in the protocol. PROSPERO, therefore, provides access to ongoing reviews as well as completed and/or published reviews. It should be noted, however, that PROSPERO does not accept the registration of protocols for scoping reviews.

Epistemonikos is a web-based bibliographic service which provides access to approximately 500,000 systematic reviews (as of June 2024) together with other records such as broad syntheses of reviews and structured summaries, and their included primary studies (http://www.epistemonikos.org/en). The aim of Epistemonikos is to provide rapid access to systematic reviews in health. Epistemonikos uses the eligibility criteria specified by the review authors to include primary studies in the database.

The Systematic Review Data Repository (SRDR) and the Systematic Review Data Repository Plus (SRDR+) were both open and searchable archives of systematic reviews and their data (Saldanha et al 2019). As of December 2021, it was announced that the SRDR resource would be decommissioned on 7 January 2022 and SRDR+ would be the only actively updated resource in future (https://srdrplus.ahrq.gov/).

Health Systems Evidence, from McMaster, is a repository of evidence syntheses about governance, financial and delivery arrangements within health systems, and about implementation strategies that can support change in health systems. The types of syntheses include evidence briefs for policy, overviews of systematic reviews, systematic reviews, protocols, and registered titles. The audience is policy makers/researchers.

Health Evidence (https://www.healthevidence.org/), also from McMaster, provides access to c. 10,000 quality-rated systematic reviews evaluating the effectiveness and cost-effectiveness of public health interventions, including cost data, relevant to public health (as of June 2024).

Specific evidence-based search services such as Trip (previously known as Turning Research into Practice (TRIP)) (https://www.tripdatabase.com/) can also be used to identify reviews and guidelines (Brassey 2007). For the range of systematic review sources searched by Trip see www.tripdatabase.com/about. Access is offered at two levels: Trip is free of charge and Trip Pro is available on subscription.

MEDLINE, Embase and other bibliographic databases, such as CINAHL (Wright et al 2015), can also be used to identify review articles and guidelines. For the 2019 release of the Medical Subject Headings (MeSH), Systematic Review was introduced as a Publication Type term. NLM announced: “We added the publication type ‘Systematic Review’ retrospectively to appropriate existing MEDLINE citations. With this re-indexing, you can retrieve all MEDLINE citations for systematic reviews and identify systematic reviews with high precision.”

https://www.nlm.nih.gov/pubs/techbull/ma19/brief/ma19_systematic_review.html

Embase has a thesaurus (Emtree) term ‘Systematic Review’, which was introduced in 2003. For records prior to 2003, the Emtree terms ‘review’ or ‘evidence-based medicine’ could be used.

Several filters to identify reviews and overviews of systematic reviews in MEDLINE (Boynton et al 1998, Shojania and Bero 2001, Montori et al 2005, Wilczynski and Haynes 2009, Lee et al 2012, Lunny et al 2015, Salvador-Oliván et al 2021) and Embase (White et al 2001, Wilczynski et al

2007, Lee et al 2012) have been developed, tested and published over the years. Until late 2018, the PubMed Systematic Reviews filter under the Clinical Queries link was very broad in its scope and retrieved many references that were not systematic reviews. The strategy was defined by NLM as follows: "This strategy is intended to retrieve citations identified as systematic reviews, meta-analyses, reviews of clinical trials, evidence-based medicine, consensus development conferences, guidelines, and citations to articles from journals specializing in review studies of value to clinicians. This filter can be used in a search as systematic [sb]." An archived version of this search filter is available from the InterTASC Information Specialists' Sub-Group's Search Filter Resource at:

https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/sr-filter?authuser=0

This search filter was replaced by NLM in late 2018 with a much more precise filter and is defined by NLM as follows: "This strategy is intended to retrieve citations to systematic reviews in PubMed and encompasses: citations assigned the 'Systematic Review' publication type during MEDLINE indexing; citations that have not yet completed MEDLINE indexing; and non-MEDLINE citations. This filter can be used in a PubMed search as systematic [sb]."

Example: exercise hypertension AND systematic [sb]

This filter is also available in PubMed on the Filters sidebar under 'Article types' and on the Clinical Queries screen. The full search filter is available at:

https://www.nlm.nih.gov/bsd/pubmed_subsets/sysreviews_strategy.html

The sensitive Clinical Queries Filters for therapy, diagnosis, prognosis, and aetiology perform well in retrieving not only primary studies but also systematic reviews in PubMed. In a test of the Clinical Queries Filters by the McMaster Health Information Research Unit (HIRU), Wilczynski and colleagues reported that performance could be improved by combining the Clinical Queries Filters with the HIRU systematic review filter using the Boolean operator 'OR' (Wilczynski et al 2011). As well as filters for study design, some filters are available for special populations, and these might be combined with systematic review filters (Boluyt et al 2008).

Research has been conducted to help researchers choose the filter appropriate to their needs (Lee et al 2012, Rathbone et al 2016). Filters and current reviews of filter performance to identify systematic reviews can be found on the InterTASC Information Specialists' Sub-Group's Search Filter Resource website (https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/systematic-reviews) (Glanville et al 2006). For further information on search filters see Section 3.6 and subsections.

National and regional drug approval and reimbursement agencies may also be useful sources of reviews:

- The Agency for Healthcare Research and Quality (AHRQ) publishes systematic reviews and meta-analyses. Evidence reports, comparative effectiveness reviews, technical briefs,

Technology Assessment Program reports, and US Preventive Services Task Force evidence syntheses are available under the Evidence-based Practice Centers (EPC) Program of the Agency for Healthcare Research and Quality. Access to the evidence reports is provided at: https://www.ahrq.gov/research/findings/evidence-based-reports/search.html.

- Canada's Drug Agency (CDA-AMC), previously known as CADTH (https://www.cadth.ca), is an independent, not-for-profit organization responsible for providing healthcare decision-makers with evidence reports to help make informed decisions about the optimal use of drugs, diagnostic tests, and medical, dental, and surgical devices and procedures. CADTH's Reimbursement Review Reports, Health Technology Assessments, Technology Reviews and Therapeutic Reviews are published in full text on their website and include the full search strategy for the clinical evidence used in that review. The transition from CADTH to Canada's Drug Agency is ongoing (as of June 2024) and it will become clearer in due course which evidence syntheses they will publish in future.

- The National Institute for Health and Care Excellence (NICE) (https://www.nice.org.uk) publishes guidance that includes recommendations on the use of new and existing medicines and other treatments within the National Health Service (NHS) in England and Wales. These reviews can be about medicines, medical devices, diagnostic tests, surgical procedures, or health promotion activities. Each guidance and appraisal document is based on a review of the evidence and reports the searches used.

Clinical guidelines, based on reviews of evidence, may also provide useful information about the search strategies used in their development: see the Appendix for examples of sources of clinical guidelines. Guidelines can also be identified by searching MEDLINE where guidelines should be indexed under the Publication Type term 'Practice Guideline', which was introduced in 1991. Embase has a thesaurus term 'Practice Guideline', which was introduced in 1994.

The ECRI Guidelines Trust (https://guidelines.ecri.org/) provides access to a free web-based repository of objective, evidence-based clinical practice guideline content. It includes evidence-based guidance developed by nationally and internationally recognized medical organizations and medical specialty societies. Guidelines are summarized and appraised against the US Institute of Medicine (IOM) Standards for Trustworthiness. The Guidelines Trust provides the following guideline-related content:

- *Guideline Snapshots and Guideline Profiles.*

- *TRUST (**T**ransparency and **R**igor **U**sing **S**tandards of **T**rustworthiness) Scorecards*: ratings of how well guidelines fulfil the IOM Standards for Trustworthiness.

The Agency for Healthcare Research and Quality (AHRQ)'s National Guideline Clearinghouse existed as a public resource for summaries of evidence-based clinical practice guidelines but ceased production in July 2018 with the latest guidelines being accepted for inclusion until March 2018. The resource offered systematic comparisons of selected guidelines that

addressed similar topic areas. For further information as to what will replace this resource and progress towards this see: https://www.ahrq.gov/gam/updates/index.html.

Evidence summaries such as online/electronic textbooks, point-of-care tools and clinical decision support resources are a type of synthesized medical evidence. Examples of these tools include BMJ Best Practice, ClinicalKey, Dynamed/DynaMed Plus and UpToDate in addition to Cochrane's own point-of-care tool Cochrane Clinical Answers, available within the Cochrane Library. Although they are designed to be used in clinical practice, they offer evidence for diagnosis and treatment of specific conditions and are regularly updated with links to and reference lists to reports of relevant studies which can help in identifying studies, reviews, and overviews. Most evidence summaries for use in clinical practice are available via subscription to commercial vendors.

As noted above, it is mandatory, for authors of Cochrane Reviews of interventions, to check reference lists of included studies and any relevant systematic reviews identified (MECIR C30). Checking reference lists within eligible studies supplements other searching approaches and may reveal new studies, or confirm that the topic has been thoroughly searched (Greenhalgh and Peacock 2005, Horsley et al 2011). Examples of situations where checking reference lists might be particularly beneficial are:

- when the review is of a new technology
- when there have been innovations to an existing technique or surgical approach
- where the terminology for a condition or intervention has evolved over time
- where the intervention is one which crosses subject disciplines, for example, between health and other fields such as education, psychology or social work. Researchers may use different terminology to describe an intervention depending on their field (O'Mara-Eves et al 2014) and
- where a relevant realist review has been identified, as realist review searches are likely to be iterative and rely on a range of supplementary searching techniques (Booth et al 2020, Duddy and Roberts 2022).

It is not possible to give overall guidance as to which of the above sources should be searched in the case of all reviews to identify other reviews, guidelines and reference lists as sources of studies. This will vary from review to review. Review authors should discuss this with a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist.

1.3.5 General web searching (including search engines/Google Scholar, etc.)

Searching the World Wide Web (hereafter, web) involves using resources which are not specifically designed to host and facilitate the identification of studies. This includes general search engines such as Google Search and the websites of organizations that are topically relevant for review topics, such as charities, research funders, manufacturers and medical societies. These resources often have basic search interfaces and host a wide range of content, which poses challenges when conducting systematic searching (Stansfield et al 2016). Despite these challenges web searching has the potential to identify studies that are eligible for inclusion in a review, including 'unique' studies that are not identified by other search methods (Eysenbach et al 2001, Ogilvie et al 2005, Stansfield et al 2014, Godin et al 2015, Bramer et al 2017a, Coleman et al 2020, Briscoe et al 2023). It is good practice to carry out web searching for review topics where studies are published in journals that are not indexed in bibliographic databases or where grey literature is an important source of data (Ogilvie et al 2005, Stansfield et al 2014, Godin et al 2015). Grey literature is literature "which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers" (see Section 1.1.6) (Farace and Frantzen 1997, Farace and Frantzen 2005).

It is good practice to base the search terms used for web searching on the search terms used for searching bibliographic databases (Eysenbach et al 2001). A simplified approach, however, might be required due to the basic search interfaces of web resources. For example, web resources are unlikely to support multi-line search strategy development or nested use of Boolean operators, and single-line searching is often limited by a maximum number of alphanumeric characters. As such, it might be necessary to rewrite a search using fewer search terms or to conduct several searches of the same resource using different combinations of search terms (Eysenbach et al 2001, Stansfield et al 2016, Briscoe et al 2020b). In addition to using search terms, web searching involves following links to webpages and websites. This is less structured than searching using pre-specified search terms and the searcher will need to use their discretion to decide when to start and stop searching (Stansfield et al 2016). Wherever possible, a similar approach to searching should be used for different web resources to ensure consistency and searches should be documented in full and reported in the review (see Chapter 4, Section 4.5).

Web resources are unlikely to have a function for exporting results to reference management software, in which case the searcher may decide to screen the results 'on screen' while searching. Alternatively, screenshots can be taken and screened at a later time (Stansfield et al 2016). This process can be facilitated by software such as Evernote or OneNote. Because website content can be deleted or edited by the website editor at any time, a permanent record of any relevant studies should be retained.

Web searching should use a combination of search engines and websites to ensure a wide range of sources are identified and searched in depth.

Search engines

Due to the scale and diversity of content on the web, searching using a search engine is likely to retrieve an unmanageable number of results (Mahood et al 2014). Results are usually ranked according to relevance as determined by a search engine's algorithm, so it might be useful to limit the screening process to a pre-specified number of results, e.g. limits ranging from 100 to 500 results have been reported in Cochrane Reviews (Briscoe 2018). Alternatively, an ad hoc decision to stop screening can be made when the search results become less relevant (Stansfield et al 2016). It is good practice to use a more comprehensive approach when screening Google Scholar results, which are limited to 1,000, to ensure that all relevant studies, including grey literature, are identified (Haddaway et al 2015). Some search engines allow the user to limit searches to a specified domain name or file type, or to web pages where the search terms appear in the title. These options might improve the precision of a search though they might also reduce its sensitivity. The reported number of results identified by Google Scholar is usually an estimate which varies over time, and the viewable results might be lower than reported (Bramer 2016). Similarly, recent studies show that the viewable number of results in Google Search is typically much lower than the estimated number reported by the search engine (Briscoe and Rogers 2021, Briscoe et al 2023). Search engines often combine search terms using the 'AND' Boolean operator by default. Some search engines support additional search operators and features such as 'OR', 'NOT', wildcards and phrase searching using quotation marks.

There are many freely available search engines, each of which offers a different approach to searching the web. Because each search engine uses a different algorithm to retrieve and rank its results, the results will differ depending on the search engine that is used (Dogpile.com 2007). Thus it might be worth experimenting with or combining use of different search engines to retrieve a wider selection of results. There are freely available meta-search engines which search a combination of search engines, though they are often limited with regard to which search engines can be combined. Some search engines tailor the search results to a user's search history and location, so the search results might differ between users, thus limiting reproducibility (Cooper et al 2021a). It has been suggested that clearing a web browser's cache and cookies before searching should reduce the personalization of results(Curkovic and Kosec 2018). Cooper et al, however, found that even when a user was logged out of Google accounts and cookies/caches were cleared before each search, there were still variations in search returns depending on the geographical location of the searcher (Cooper et al 2021a). To improve the transparency of reporting, the searcher location (e.g. Oxford, UK) should be reported alongside the website URL, date of search, and search syntax used.

A selection of freely available search engines and meta-search engines is shown in Box 1.a.These are examples of different types of search engine rather than a list of recommended search engines. No specific search engines are recommended for a Cochrane Review.

Box 1.a Search engines

Dogpile http://www.dogpile.com/

Dogpile is a meta-search engine which in a study from 2007 is reported to search Google Search, Yahoo!, Ask and Bing (Dogpile.com 2007). A more up to date list of search engines used by Dogpile has not been identified, although the About page on the website states that "Dogpile returns all the best results from leading search engines including Google and Yahoo!".

DuckDuckGo https://duckduckgo.com/

DuckDuckGo protects the privacy of its users by not recording their IP addresses and search histories. A potential advantage for systematic review authors is that DuckDuckGo does not use search histories to personalize its search results, which might make it better at ranking less frequently visited but useful pages higher in the results.

Google Scholar https://scholar.google.com/

Google Scholar is a specialized version of Google Search which limits results to scholarly literature, including published studies and grey literature. It cannot be used instead of searching bibliographic databases due to its basic search interface and a block on viewing more than 1,000 records per search (Boeker et al 2013a, Bramer et al 2016a). It can, however, be a useful resource when used alongside bibliographic databases for identifying studies including those reported in grey literature not indexed in bibliographic databases or not retrieved by the bibliographic database search strategies (Haddaway et al 2015, Bramer et al 2017a). The option to search the full text of studies can contribute to the identification of unique studies when using similar or the same search terms as used in bibliographic databases (Bramer et al 2017a). References can be exported to reference management software, though the number of references that can be exported at a time is limited to 20 (Bramer et al 2013). However, Google Scholar can be searched via the freely available Publish or Perish software, which also facilitates bulk exportation of results to reference management software (Harzing 2006).

Google Search https://www.google.com/

Google Search is the most widely used search engine worldwide. An advantage of its popularity is that there is an abundance of online material on how to make the most of its advanced search features. The Verbatim feature in the Google Search Tools menu can be used to ensure search results contain the precise search terms used (e.g. will not retrieve "nursing" if searching for "nurse") and to switch off the personalization of search results based on websites which the user has previously visited. Personalization can also be deactivated via the settings menu. Google Search can be set to display 100 results per page, which can help to ascertain a more accurate estimate of the total number of results which

are viewable for screening than the number estimated by the search engine (Briscoe and Rogers 2021). Studies show that the viewable number of results is typically much lower than the estimated number (Briscoe and Rogers 2021, Briscoe et al 2023).

Not all content on websites is indexed by search engines, so it is important to consider accessing and searching any potentially useful websites which are identified in the search results (Devine and Egger-Sider 2013).

Websites

The selection of websites to search will be determined by the review topic. It is good practice to investigate whether the websites of relevant pharmaceutical companies and medical device manufacturers host trials registers which should be searched for studies. The websites of medicines regulatory bodies such as the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) should be searched for regulatory documentation (see Section 1.2 and subsections). It might also be useful to search the websites of professional societies, national and regional health departments, and health related non-governmental organizations and charities for studies not indexed in bibliographic databases and grey literature (Ogilvie et al 2005, Godin et al 2015, Briscoe et al 2020b).

Searching websites will usually yield a lower number of results than search engines, so it should be possible to screen all the results rather than a pre-specified number.

1.3.6 Other sources

In addition to citation searching via citation indexes (see Section 1.1.4), other resources have become available over recent years to achieve both forwards and backwards citation searching. Some of these tools also offer proprietary algorithms to find related or similar articles, and also to achieve co-citation analysis. The following list is a selection of useful tools that search a large volume of records and offer analyses ranging from simple forwards and backwards analysis to more complex document relationship algorithms:

- citationchaser (https://estech.shinyapps.io/citationchaser/) offers rapid backwards and forwards citation chasing and results can be downloaded in RIS format (Haddaway et al 2021, Haddaway et al 2022). citationchaser uses Lens.org (PubMed, PubMed Central, CrossRef, Microsoft Academic Graph and CORE).
- ResearchRabbit (https://www.researchrabbit.ai/) offers backwards and forwards citation searching as well as similar works. It claims to search more than 90% of the documents that are found in databases such as Scopus and Web of Science. It offers integration with Zotero for managing results as well as exports in BibTeX, RIS and CSV formats.
- Inciteful (https://inciteful.xyz/) offers forward and backwards citation searching as well as a range of other analyses of the interconnections of publications. It achieves citation

analysis using OpenAlex, Semantic Scholar, Crossref and OpenCitations. It has a Zotero plugin as well as exporting records in BibTeX, RIS, Mendeley and Zotero formats.
- Litmaps (https://docs.litmaps.com/en/) offers forwards, backwards and a range of other citation analyses. It accesses data from Crossref, Semantic Scholar and OpenAlex. Export formats include BibTeX, CVS and RIS.

A list of citation analysis tools can be found here: https://start.me/p/Rn5m5a/citation-analysis-tools.

1.4 Summary points

- Cochrane Review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, on sources to search.
- Authors of non-Cochrane reviews should seek advice from a medical/healthcare librarian or information specialist, with experience of conducting searches for studies for systematic reviews.
- The key database sources which should be searched are the Cochrane Group's Specialized Register where such Specialized Registers exist (internally, e.g. via the Cochrane Register of Studies, or externally via CENTRAL), CENTRAL, MEDLINE and Embase (if access to Embase is available to either the review authors or the Cochrane Group).
- Appropriate national, regional and subject specific bibliographic databases should be searched according to the topic of the review.
- Relevant grey literature sources such as those containing reports, dissertations/theses and conference abstracts should be searched.
- Searches should be conducted to locate previous reviews on the same topic, to identify additional studies included in (and excluded from) those reviews.
- Reference lists of included studies should be checked to identify additional studies.
- Trials registers and repositories of results, where relevant to the topic, should be searched through both ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP) portal and other sources as appropriate.
- Regulatory agency sources and clinical study reports should also be considered as sources for study data.
- Citation searching should be considered as an additional means of identifying relevant studies.

2 Planning the search process

2.1 Cochrane-wide search initiatives and the Cochrane Centralized Search Service (CSS)

In 2015, building on the processes established for the Embase project to identify records from Embase and MEDLINE (see Section 2.1.2), Cochrane began a pilot initiative with the objective of adding to the number of sources to be searched and screened 'centrally', known as the Cochrane Centralized Search Service (CSS). The CSS initiative has since been expanded to cover six resources. They are MEDLINE/PubMed (see Section 2.1.1), Embase (see Section 2.1.2), ClinicalTrials.gov (see Section 2.1.3.2), the WHO International Clinical Trials Registry Platform (ICTRP) (see Section 2.1.3.3), KoreaMed (see Section 2.1.3.4) and CINAHL Plus (see Section 2.1.3.5). All sources are searched or queried via an API (Application Programming Interface) each week, with the exception of ClinicalTrials.gov, which is queried daily and CINAHL Plus which is queried monthly. For each source an appropriately sensitive search approach to identifying possible RCTs has been developed and implemented (see Table 2.1.a for an overview, and for further details see the 'How CENTRAL is created' file in the Cochrane Library: https://www.cochranelibrary.com/central/central-creation). For both Embase and CINAHL Plus, a methodological search filter has been developed (see Sections 3.6.2 and 3.6.3 respectively).

Each of the CSS sources had 'backlogs' to deal with in parallel to setting up prospective routines to identify newly indexed reports of RCTs. The backlogs for all sources (Embase, MEDLINE/PubMed, ClinicalTrials.gov, ICTRP, KoreaMed and CINAHL Plus) have been cleared. This was achieved by using a combination of machine learning in the form of the RCT Classifier (Thomas et al 2021) and crowdsourcing via the Cochrane Crowd (https://crowd.cochrane.org/). The CSS aims to provide systematic review authors and others with an even baseline of access, via CENTRAL, to the relevant evidence needed to produce systematic reviews and other evidence products. It is unlikely it will ever completely replace the need for some multi-source, bespoke, review-based searches, especially for cross-disciplinary or complex reviews, but it is hoped that it will substantially improve access to RCT evidence and reduce the amount of multi-source searching currently needed. A recent, retrospective analysis showed that 97.5% of RCTs published in 2017 and 2018 that had been included in Cochrane Reviews had been identified by the CSS (Noel-Storr et al 2020).

Information specialists should consider numerous factors when deciding which sources to include in their searches. These include being aware of the time taken for records to appear in CENTRAL from source databases such as MEDLINE and Embase, understanding that across the years different processes and searches have been used to populate CENTRAL, and recognizing that for trials register records not all fields of content available for those records in their source databases are included in CENTRAL. Each of these factors is discussed in more detail in an analysis of the CSS (Noel-Storr et al 2020, Noel-Storr and Wisniewski 2024).

Table 2.1.a is designed to be a quick reference to sources that feed or have recently fed into CENTRAL: Figure 2.a illustrates the contents of CENTRAL.

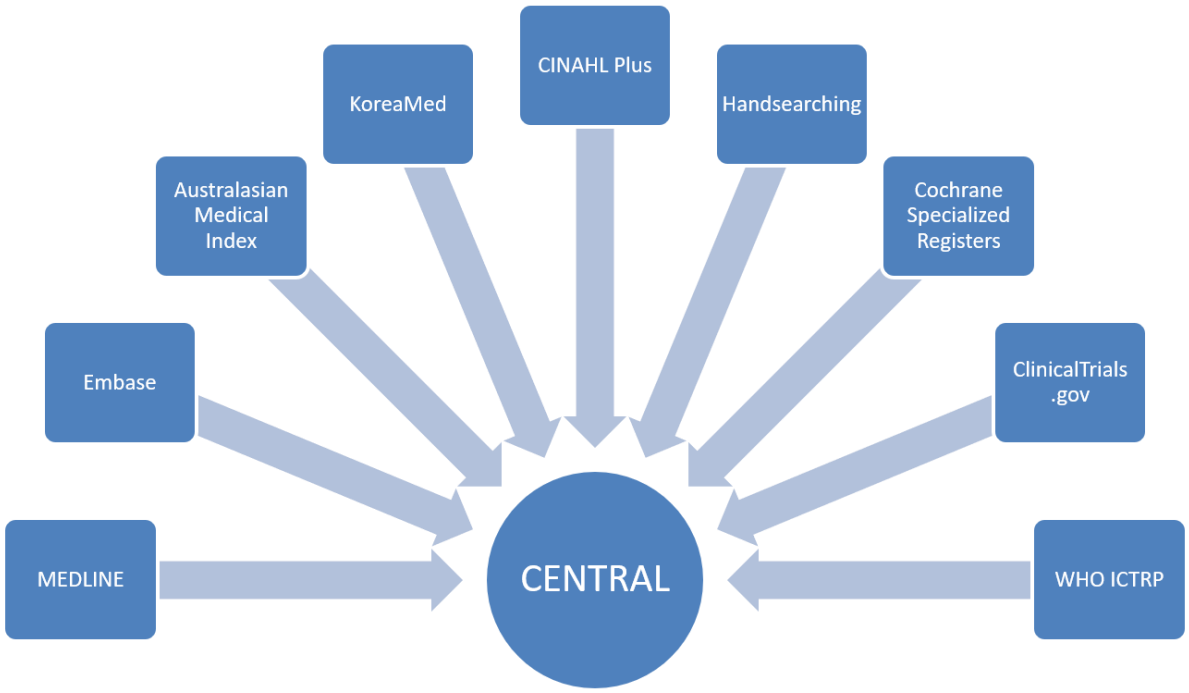
Table 2.1.a Sources searched as part of the Cochrane Centralized Search Service (CSS)

Source (provider)	Workflow description (current/most recent)	Harvested from external source
PubMed* (US National Library of Medicine)	Direct feed of records into CENTRAL based on index terms (MeSH) search: "randomized controlled trial" [Publication Type] OR "controlled clinical trial" [Publication Type]	Weekly API call
Embase* (Elsevier)	Sensitive search of Embase.com via the Embase.com API using the Cochrane search filter for identifying RCTs in Embase (see Section 3.6.2). Records with both the Emtree term: Randomized Controlled Trial exp and that receive a high RCT Classifier score are submitted directly to CENTRAL (i.e. both the above criteria must be met). Records that receive below threshold score by RCT Classifier are discarded. Remaining records are manually assessed by the Cochrane Crowd.	Weekly API call
ClinicalTrials.gov (US National Library of Medicine)	All records are downloaded and run through the RCT Classifier. Records scoring below threshold score are discarded. Remaining records are manually assessed by the Cochrane Crowd.	Daily API call
ICTRP (World Health Organization)	All records are downloaded. Records with: {(randomised OR randomized) NOT (randomised: no OR randomized: no)} in the study design or study type fields, together with (randomised OR randomized) found in any other field of the record, are directly submitted to CENTRAL, i.e. both the above criteria must be met. Remaining records are manually assessed by the Cochrane Crowd.	Weekly API call
KoreaMed** (Korean Association of Medical Journal Editors)	All records were downloaded. Records that received below threshold score by RCT Classifier were discarded. Remaining records were manually assessed by the Cochrane Crowd.	Monthly API call was on 15th of each month (now discontinued)
CINAHL Plus (EBSCOhost)	Sensitive search of CINAHL Plus via API using the Cochrane search filter for identifying RCTs in CINAHL (see Section 3.6.3). Records that receive below threshold score by RCT Classifier are discarded. Remaining records are manually assessed by the Cochrane Crowd.	Monthly API call on 15th of each month

*The search of Embase includes MEDLINE/PubMed records.

**As of May 2021, KoreaMed is no longer searched as part of the Cochrane Centralized Search Service (CSS).

Figure 2.a Illustration of the contents of CENTRAL



2.1.1 What is in the Cochrane Central Register of Controlled Trials (CENTRAL) from MEDLINE?

CENTRAL contains all records from MEDLINE indexed with the Publication Type term 'Randomized Controlled Trial' or 'Controlled Clinical Trial' except those that are indexed solely as animal studies (not also as human studies). For further details see the 'How CENTRAL is created' file in the Cochrane Library:

https://www.cochranelibrary.com/central/central-creation

A substantial proportion of the MEDLINE records coded 'Randomized Controlled Trial' or 'Controlled Clinical Trial' in the Publication Type field have been coded as a result of early work within Cochrane (Dickersin et al 2002). Handsearch results from Cochrane Groups, for journals indexed in MEDLINE, were sent to the US National Library of Medicine (NLM), where the MEDLINE records were re-tagged with the Publication Types 'Randomized Controlled Trial' or 'Controlled Clinical Trial' as appropriate. In addition, the US Cochrane Center (formerly the New England Cochrane Center, Providence Office and the Baltimore Cochrane Center and now Cochrane US) and the UK Cochrane Centre (thereafter Cochrane UK) conducted a search of MEDLINE from 1966 to 2004 to identify reports of randomized trials, identifiable from the MEDLINE titles and/or abstracts, not already indexed as such, using the first two phases of the original Cochrane Highly Sensitive Search Strategy first published in 1994 (Dickersin et al 1994) and thereafter updated and included in subsequent editions of this *Handbook* (see Section

3.6.1). The free-text terms used were: clinical trial; (singl$ OR doubl$ OR trebl$ OR tripl$) AND (mask$ OR blind$); placebo$; random$. The $ sign indicates the use of a truncation symbol. The subject heading terms (MeSH) used were ('exploded' where possible to include narrower, more specific terms): randomized controlled trials; random allocation; double-blind method; single-blind method; clinical trials; placebos. The following subject heading term (MeSH) was used 'unexploded': research design. The Publication Type terms used were: randomized controlled trial; controlled clinical trial; clinical trial.

A test was carried out using the terms in phase three of the 1994 Cochrane Highly Sensitive Search Strategy but the precision of those terms, having already searched on all the terms in phases one and two as listed above, was considered to be too low to warrant using these terms for the above project (Lefebvre and Clarke 2001). It was, however, recognized that some of these terms might be useful when combined with subject terms to identify studies for some specific reviews (Eisinga et al 2007).

The above search was limited to humans. The following years were completed by the US Cochrane Center (1966 to 1984; 1998 to 2004) and by the UK Cochrane Centre (1985 to 1997). The results for these years were forwarded to the NLM and re-tagged in MEDLINE and are thus included in CENTRAL. More recent MEDLINE records, which are now included in Embase, are being searched as part of the Cochrane Centralized Search Service (CSS) (see Section 2.1.2).

CENTRAL includes from MEDLINE not only reports of trials that meet the more restrictive Cochrane definition for a quasi-randomized trial (indexed in MEDLINE as 'Controlled Clinical Trial') (Box 2.a) but also trial reports that meet the less restrictive NLM definition (Box 2.b) which includes historical comparisons. There is currently no method of distinguishing, either in CENTRAL or in MEDLINE, which of these records meet the more restrictive Cochrane definition, as they are all indexed with the Publication Type term 'Controlled Clinical Trial'.

Box 2.a Cochrane definitions and criteria for randomized controlled trials (RCTs) and quasi-randomized trials

Records identified for inclusion in CENTRAL should meet the eligibility criteria devised and agreed in November 1992, which were first published, in 1994, in the first version of this *Handbook* (Oxman et al 1994). According to these eligibility criteria:

A trial is eligible if, on the basis of the best available information (usually from one or more published reports), it is judged that:

- the individuals (or other units) followed in the trial were definitely or possibly assigned prospectively to one of two (or more) alternative forms of health care using:
    - random allocation; or

- some quasi-random method of allocation (such as alternation, date of birth, or case record number).

Trials eligible for inclusion are classified according to the reader's degree of certainty that random allocation was used to form the comparison groups in the trial. If the author(s) state explicitly (usually by some variant of the term 'random' to describe the allocation procedure used) that the groups compared in the trial were established by random allocation, then the trial is classified as a RCT (randomized controlled trial). If the author(s) do not state explicitly that the trial was randomized, but randomization cannot be ruled out, the report is classified as a CCT (controlled clinical trial). The classification CCT is also applied to quasi-randomized studies, where the method of allocation is known but is not considered strictly random, and also trials that are possibly quasi-randomized. Examples of quasi-random methods of assignment include alternation, date of birth, and medical record number.

The classification as RCT or CCT is based solely on what the author has written, not on the reader's interpretation; thus, it is not meant to reflect an assessment of the true nature or quality of the allocation procedure. For example, although 'double-blind' trials are nearly always randomized, many trial reports fail to mention random allocation explicitly and should therefore be classified as CCT.

Relevant reports are reports published in any year, of studies comparing at least two forms of health care (healthcare treatment, healthcare education, diagnostic tests or techniques, a preventive intervention, etc.) where the study is on either living humans or parts of their body or human parts that will be replaced in living humans (e.g. donor kidneys). Studies on cadavers, extracted teeth, cell lines, etc. are not relevant. *Searchers should identify all controlled trials meeting these criteria regardless of relevance to the entity with which they are affiliated.*

The highest possible proportion of all reports of controlled trials of health care should be included in CENTRAL. Thus, those searching the literature to identify trials should give reports the benefit of any doubts. Review authors will decide whether to include a particular report in a review.

In 2013, a Cochrane working group was formed to review the record type eligibility for CENTRAL and to ensure consistency of practice and guidance for the Embase project and handsearcher training. This group focused on types of report rather than types of study. The group determined that reports of protocols for randomized or quasi-randomized trials, along with letters, replies, errata, and retractions relating to RCTs or quasi-RCTs are all to be included in CENTRAL.

Box 2.b US National Library of Medicine 2024 definitions (Scope Notes) for the Publication Type terms 'Randomized Controlled Trial' and 'Controlled Clinical Trial'

Randomized Controlled Trial

A work that reports on a clinical trial that involves at least one test treatment and one control treatment, concurrent enrollment and follow-up of the test- and control-treated groups, and in which the treatments to be administered are selected by a random process, such as the use of a random-numbers table

Controlled Clinical Trial

A work that reports on a clinical trial involving one or more test treatments, at least one control treatment, specified outcome measures for evaluating the studied intervention, and a bias-free method for assigning patients to the test treatment. The treatment may be drugs, devices, or procedures studied for diagnostic, therapeutic, or prophylactic effectiveness. Control measures include placebos, active medicine, no-treatment, dosage forms and regimens, historical comparisons, etc. When randomization using mathematical techniques, such as the use of a random numbers table, is employed to assign patients to test or control treatments, the trial is characterized as a RANDOMIZED CONTROLLED TRIAL

https://www.ncbi.nlm.nih.gov/mesh/68016449

https://www.ncbi.nlm.nih.gov/mesh/68018848

MEDLINE records are also currently being added into CENTRAL from Embase. Since 2010, Elsevier has included MEDLINE records in Embase (see further details in Section 2.2.2 on specific issues when searching MEDLINE and Embase).

2.1.2 What is in the Cochrane Central Register of Controlled Trials (CENTRAL) from Embase?

A retrospective search conducted by the UK Cochrane Centre (thereafter Cochrane UK) for reports of trials in Embase was undertaken for the years 1974 to 2010. For the years 1974 to 1979, the free-text terms: random$; factorial$; crossover$; cross-over$; and placebo$ were used. For the years 1980 to 2008, the following free-text terms: random$; factorial$; crossover$; cross-over$; cross over$; placebo$; doubl$ adj blind$; singl$ adj blind$; assign$; allocat$; volunteer$; and the following index terms, known as Emtree terms: crossover-procedure; double-blind procedure; randomized controlled trial; single-blind procedure were used. For 2009, the following free-text terms: random$; crossover$; cross-over$; cross over$; placebo$; doubl$ adj blind$; singl$ adj blind$; allocat$; and the following index terms, known as Emtree terms: crossover-procedure; double-blind procedure; randomized controlled trial; single-blind procedure were used. In addition, the following terms were searched limited to the title only:

trial; comparison. For 2010, the following free-text terms were searched limited to the title, abstract and original title fields only: crossover$, cross over$, placebo$, doubl$ adj blind$, allocat$, random$; and limited to the title only: trial; and the following index terms were searched: crossover-procedure; double-blind procedure; single-blind procedure and randomized controlled trial. (Note: cross over$ includes cross-over$ in Ovid syntax).

The searches across all years of this project (1974 to 2010) yielded a total of approximately 100,000 reports of trials not indexed, at the time of the search, as randomized controlled trial or controlled clinical trial in MEDLINE. All of these reports are now published in CENTRAL (Lefebvre et al 2008). The final submission of reports under this project, of trials identified in *journal article records* added to Embase in 2010, was published in CENTRAL in February 2012. This project then formally ended, with a newly funded project starting in 2013.

In March 2013, Cochrane launched a further Embase project to provide ongoing screening of records from Embase to identify additional reports of trials. This project was co-ordinated by Metaxis Ltd., the Cochrane Dementia and Cognitive Improvement Group and York Health Economics Consortium. Initially, a search covering January 2011 to December 2013, inclusive, was run, from which approximately 30,000 unique Embase records were identified and published in CENTRAL, January 2014 (Issue 1). All these records were identified from a search in Embase (via Ovid) using the Emtree headings Randomized Controlled Trial (RCT) or Controlled Clinical Trial (CCT). It is estimated that this search, using only these two headings, identified two-thirds of records eligible for inclusion in CENTRAL from the 2011 to 2013 period.

The remaining records were identified using the search strategy developed by the UK Cochrane Centre (thereafter Cochrane UK), described above, with records indexed as either RCT or CCT removed, as those records had already been identified and added to CENTRAL. A small team of expert screeners screened the results retrieved and identified a further 20,000 records eligible for CENTRAL.

In parallel to the work described above, a new search filter to identify potential reports of randomized trials in Embase was developed in 2013 and initiated in January 2014 (Glanville et al 2019a). It was developed following an examination of 1,000 relevant reports (reference standard) of randomized trials, and was tested on a second set of 1,000 records. The filter was tiered. The first tier identified records with the most relevant Emtree headings RANDOMIZED CONTROLLED TRIAL or CONTROLLED CLINICAL STUDY. The second tier comprised search terms likely to find records from the reference standard which did not contain those two Emtree headings. The revised filter was used from January 2015. It was initially run as two searches with records containing Emtree terms RANDOMIZED CONTROLLED TRIAL or CONTROLLED CLINICAL STUDY being directly fed into CENTRAL. The remaining records retrieved by the new filter were sent for manual screening via the Cochrane Crowd. Minor revisions to the filter were made in 2017 and 2021. These revisions were aimed at reducing the number of non-RCTs being fed directly into CENTRAL and reducing the number of animal RCTs

identified. For details of the current process and filter used, see https://www.cochranelibrary.com/central/central-creation.

Currently, records are screened using a crowdsourcing model, accessible from the Cochrane Crowd platform (https://crowd.cochrane.org/). Here, Cochrane contributors and members of the general public can contribute to screening records after completing a brief training exercise. As of April 2024 over 2 million Embase records had been collectively screened, and over 200,000 additional reports of trials had been identified and added to CENTRAL.

In 2009, Elsevier began adding conference records to Embase, and to date (April 2024) has added about 5.1 million conference abstracts from about 15,000 conferences (https://beta.elsevier.com/products/embase/content?trial=true). The addition of conference abstracts in Embase created a sizable backlog of records for the Cochrane Centralized Search Service (CSS). The Embase screening project searched and downloaded all records (not just conference abstracts) added to Embase between 2010 and 2013 inclusive. The search strategy used for the conference 'backlog' was the most recent version in use by the UK Cochrane Centre at that time. This was so that screening of this backlog could get underway quickly whilst the new search filter was being developed. All reports of RCTs identified from the screening of these records had been published in CENTRAL by the end of 2014.

Introducing machine learning into the workflow

In January 2016 the machine learning RCT Classifier was used for the first time on records identified from Embase via the monthly sensitive search described above. Records that received a likelihood score below a pre-specified cut-off-point were deemed to be not RCTs and no further action was taken on them. Those records that scored on or above the cut-off-point were then sent to the Cochrane Crowd for manual assessment. This remained the workflow for Embase records from the start of 2016 until February 2023. From February 2023 onwards, records with both the Emtree term: Randomized Controlled Trial exp *and* that receive a high RCT Classifier score are submitted directly to CENTRAL (i.e. *both* the above criteria must be met); records that have been indexed with the Emtree term: Randomized Controlled Trial exp but receive a low RCT Classifier score are sent to Cochrane Crowd for manual screening. Work to evaluate the potential and the performance of the RCT Classifier can be found in (Wallace et al 2017, Marshall et al 2018, Thomas et al 2021). In terms of the application of the RCT classifier to the central feed of Embase records, approximately 50% of records score below the currently used cut-off-point representing a significant reduction in manual screening required by the Cochrane Crowd. (See Chapter 4, Section 4.6.6.2 for further information about using machine learning to classify reports of RCTs).

2.1.3 What is in the Cochrane Central Register of Controlled Trials (CENTRAL) from other non-Cochrane sources and handsearching?

2.1.3.1 Introduction

Many Cochrane Groups and Fields have undertaken searching of the specialist healthcare literature (both journals and databases) in their areas of interest. More than 3,000 journals have been handsearched. Identified trial reports that were not relevant to a Cochrane Group's scope and thus were not appropriate for their Specialized Register (see Section 2.1.4) were published in CENTRAL as handsearch results. Handsearch records can be identified in CENTRAL as they are assigned the tag HS-HANDSRCH in addition to a source code indicating the Centre, Field or Review Group that submitted the record (see https://www.cochranelibrary.com/central/central-creation).

The Australasian Cochrane Centre (now Cochrane Australia) co-ordinated a search of the National Library of Australia's Australasian Medical Index from 1966 (McDonald 2002). This search was updated to include records added up to December 2009, when the database ceased to be updated. All records identified have been added to CENTRAL.

The Chinese Cochrane Center (now Cochrane China), with support from the Australasian Cochrane Centre (now Cochrane Australia), the UK Cochrane Centre (thereafter Cochrane UK) and Cochrane centrally co-ordinated a search of the Chinese Biomedical Literature Database (CBM) from 1978 to 2008 and identified approximately 30,000 reports of trials. These records have not been added to CENTRAL.

2.1.3.2 Records from ClinicalTrials.gov

From August 2017, eligible ClinicalTrials.gov (https://clinicaltrials.gov/) records are being identified and systematically added to CENTRAL through Cochrane's Centralized Search Service (CSS).

Process description

All ClinicalTrials.gov records go through Cochrane's RCT Classifier developed specifically for ClinicalTrials.gov records. The ClinicalTrials.gov RCT Classifier provides likelihood scores for each record being either a randomized or quasi-randomized trial report. Records with an 80% or greater likelihood score are submitted directly to CENTRAL. Records with a 10% or less likelihood score are rejected without any further action. Records with a likelihood score between 11 and 79% inclusive are sent to the Cochrane Crowd to be manually screened. Performance evaluations, conducted internally by the Centralized Search Service Team (CSS), show over 99% accuracy at the thresholds described above.

Field mappings

The ClinicalTrials.gov records contain many fields, but not all fields are included in CENTRAL. The fields that are displayed in CENTRAL are the Public and Scientific titles, the URL to the register record, the brief summary of the trial, MeSH, and the "date first received" (i.e. the date the record was first processed in ClinicalTrials.gov). The following data fields from

ClinicalTrials.gov have not been republished in CENTRAL: Recruitment status, Study results, Condition, Intervention, Sponsor, Gender, Age, Phase, Enrolment, Funded by, Study type, Study design, Other IDs, Start date, Completion date, Last updated, Last verified, Acronym, Primary completion date, Outcome measures.

2.1.3.3 Records from the WHO's International Clinical Trials Registry Platform (ICTRP)

The World Health Organization's International Clinical Trials Registry Platform (ICTRP) (https://trialsearch.who.int/) is a meta-register containing trials data from 20 national and international registers. Since July 2018, eligible trial register records from ICTRP are being identified and systematically added to CENTRAL through Cochrane's Centralized Search Service (CSS). As with ClinicalTrials.gov, only ICTRP records for RCTs or quasi-RCTs are being added to CENTRAL; other study designs are not included.

Process description

The prospective workflow for identifying reports of RCTs and quasi-RCTs from ICTRP uses both a 'direct feed' search (for records that are extremely likely to be describing a randomized trial) and a process of manual screening via the Cochrane Crowd. The search query used for the direct feed is: {(randomised OR randomized) NOT (randomised: no OR randomized: no)} in the study design or study type fields, together with (randomised OR randomized) in any other field of the record, i.e. both the above criteria must be met. All other newly added ICTRP records are sent to Cochrane Crowd for screening. Note that 'no' in the ICTRP entry above refers to the picklist value selected by those registering their trial in ICTRP to indicate that the trial is not a randomized controlled trial. Records where the picklist value was 'no' in answer to this question about study design were excluded from the set of records directly fed into CENTRAL. Instead they were manually screened.

Field mappings

Not all fields for ICTRP records are included in CENTRAL. The fields that are included are Public and Scientific titles, the URL for the register record on ICTRP, the Key inclusion and exclusion criteria (which are mapped to the abstract field), the date of registration (mapped to the year field), and the Study ID and the Source register.

2.1.3.4 Records from KoreaMed

KoreaMed (https://www.koreamed.org) is a database provided by the Korean Association of Medical Journal Editors that contains citations to articles published in Korean medical, dental, nursing and nutrition-related journals. This database was routinely searched and records systematically added to CENTRAL through Cochrane's Centralized Search Service (CSS) until 1 May 2021.

Process description

Inception to December 2013

A project led by Cochrane Australia, in partnership with KoreaMed, sought to identify all unique reports of randomized trials across all dates within the database. As part of this work a search strategy was developed and run in KoreaMed. The search strategy was:

placebo*[ALL] OR randomi*[ALL] OR randomly[ALL] OR trial*[ALL] OR ((singl* OR doubl* OR tripl* OR trebl*) AND (blind OR mask)) OR "randomized controlled trial"[PT] OR "clinical trial"[PT] OR "double blind method"[MH] OR "single blind method"[MH]

That work identified approximately 3,000 unique reports of randomized trials, which were published in CENTRAL in April 2015.

January 2014 to July 2017

Between January 2014 and up to and including June 2017, all records that were added to KoreaMed within that time frame were manually screened by the Centralized Search Service (CSS) team, with approximately 1,000 records submitted to CENTRAL during this time.

August 2017 onwards

From August 2017, a new process was implemented. All KoreaMed records went through the Cochrane RCT Classifier and the Cochrane Crowd (https://crowd.cochrane.org/). Records that received a likelihood score (as described above for ClinicalTrials.gov records) of 10% or less were automatically rejected; records that received a score of 11% or above were sent to Cochrane Crowd for manual screening.

2.1.3.5 Records from CINAHL Plus

In November 2018 a memorandum of understanding was signed between Cochrane, Wiley and CINAHL Plus provider EBSCO (https://www.ebsco.com/products/research-databases/cinahl-database) to enable publication of unique CINAHL Plus records in CENTRAL.

Process description

Since May 2020, CINAHL Plus references to RCTs and quasi-RCTs have been identified and added to CENTRAL through Cochrane's Centralized Search Service (CSS). The CINAHL Plus RCT filter search was developed by Julie Glanville, York Health Economics Consortium, as shown below:

(MH randomized controlled trials OR MH double-blind studies OR MH single-blind studies OR MH random assignment OR MH pretest-posttest design OR MH cluster sample OR TI (randomised OR randomized) OR AB (random*) OR TI (trial) OR (MH (sample size) AND AB (assigned OR allocated OR control)) OR MH (placebos) OR PT (randomized controlled trial) OR AB (CONTROL W5 GROUP) OR MH (CROSSOVER DESIGN) OR MH (COMPARATIVE STUDIES) OR

AB (CLUSTER W3 RCT)) NOT ((MH ANIMALS+ NOT MH HUMAN) OR (MH (ANIMAL STUDIES) NOT MH (HUMAN)) OR (TI (ANIMAL MODEL) NOT MH (HUMAN)))

The CINAHL Plus RCT filter search was validated and published in February 2019 by Glanville et al (Glanville et al 2019b). The filter was adapted as an API direct feed by Metaxis in October 2019 and results were screened for inclusion in CENTRAL by Cochrane's RCT Classifier and by the Cochrane Crowd. Records that receive a score below the threshold for the RCT Classifier are discarded; the remaining records are sent to the Cochrane Crowd for manual assessment.

2.1.4 What is in the Cochrane Central Register of Controlled Trials (CENTRAL) from Specialized Registers of Cochrane Groups and Fields?

Many Cochrane Groups and some Fields have previously developed and maintained a Specialized Register, which aimed to contain all relevant studies in their area of interest. These individual registers, together with other relevant records from other sources, are stored together as a single Cochrane Register of Studies (CRS), public records of which can be accessed by any Cochrane member logged into their Cochrane Account via the Cochrane Register of Studies Online (CRSO) (https://crso.cochrane.org/). (Note: this web address can only be accessed when logged in as above.) These public records are also published in CENTRAL in the Cochrane Library. The purpose of the Specialized Register was to assemble a repository of reports of trials relating to the scope of a Cochrane Group or Field, to provide a reliable pool of trials for review authors that is easily retrievable, and to share this content with users of the Cochrane Library, via CENTRAL (Cochrane Information Specialist Support Team 2021c). Many of these Specialized Registers were reference-based registers, where each record represents a report of a clinical trial. Where there are multiple reports of a clinical trial, as is typical, there will be multiple records for that trial. Such registers are very similar to a bibliographic database (Wieland et al 2013). Some Cochrane Groups/Fields managed a study-based register, where the reports related to each clinical trial or study have been linked together, and identified by a study name (Shokraneh and Adams 2017). In this case, there should only be one record for each clinical trial or study, with all the reports of that clinical trial or study linked to the study record. In some of these groups, the Cochrane Information Specialist also extracted metadata about studies such as the study participants, the research question, interventions, outcomes, and study designs (Shokraneh and Adams 2017). The role of the Specialized Register has been superseded by Cochrane's (Centralized Search Service (CSS).

Specialized Registers primarily contained reports of randomized and quasi-randomized trials, however, some Cochrane Groups added other types of reports to their register, such as controlled before-and-after studies and interrupted time series (Cochrane Information Specialist Support Team 2021c). Whether or not these were added to the Specialized Register will depend on the scope of the Cochrane Group. These publication types can be published in CENTRAL. Cochrane Groups may also have added other reports to their register that may be useful to review authors (such as systematic reviews or background articles), but these would not be published in CENTRAL (Falzon and Trudeau 2007).

It is mandatory, for all Cochrane Reviews of interventions, to search the Cochrane Group's Specialized Register (internally, e.g. via the Cochrane Register of Studies, or externally via CENTRAL (MECIR C24)), where such a Specialized Register exists. However, most Cochrane Groups are no longer maintaining a Specialized Register. Relevant studies are now identified via Cochrane's CSS, and these records can be found by searching CENTRAL.

Records from the Specialized Registers previously maintained by Cochrane Groups are also available in CENTRAL. To identify records in CENTRAL from a specific Centre, CRG or Field, it is possible to search on a Specialized Register or Handsearch code (such as SR-STROKE for records from the Cochrane Stroke Group). A list of all the Specialized Register and Handsearch codes can be found in an Appendix in the 'How CENTRAL is created' file in the Cochrane Library entitled: CENTRAL codes for records submitted from Cochrane Review Groups, Geographic Groups, Fields, and Networks:

https://www.cochranelibrary.com/central/central-creation.

2.2 Searching CENTRAL, MEDLINE, Embase, ClinicalTrials.gov and the ICTRP: specific issues

For discussion of some of the specific issues around searching for medical devices, please refer to this recent brief method note (Cooper et al 2022).

2.2.1 Searching the Cochrane Central Register of Controlled Trials (CENTRAL): specific issues

CENTRAL, accessible via the Cochrane Library or from the Cochrane Register of Studies Online (CRSO), comprises records from a wide range of sources (see Section 2.1 and subsections). The consistency and formatting of these records therefore varies. In 2013, Cochrane ran a CENTRAL "clean-up" project. The aims of this project were to clean and harmonize as many fields as possible in existing records, and to formalize standards for Cochrane Information Specialists and/or automatically apply solutions in the CRS to help prevent inconsistencies in the future.

Additionally in 2013, Cochrane formed a working group called HarmoniSR (HarmoniSR Working Group 2015). The scope of this group was initially focused on the formatting of ClinicalTrials.gov records as citations for consistent use within Cochrane Reviews and publication within CENTRAL. The scope of the group, however, expanded during 2014 onwards to include the formatting of all main record types. Despite these ongoing efforts, legitimate differences between records remain, for example, records sourced from MEDLINE will contain Medical Subject Headings (MeSH), whilst 'native Embase' records identified from Embase will most likely contain Emtree terms.

Not all records in CENTRAL have an abstract. Optimal searches will, therefore, be those that contain both Medical Subject Headings (MeSH) and free-text terms. The 800,000 records sourced from PubMed are also best retrieved by a combination of MeSH (as the Cochrane Library has a MeSH search interface) together with free-text terms. The other records, including

the 700,000 records sourced from Embase, are best retrieved using free-text searches across all fields, as there is no Emtree search interface built into the Cochrane Library. Many of the records that are not sourced from PubMed or Embase (about 850,000 in CENTRAL in April 2024) have neither an abstract or a traditional bibliographic abstract in the case of trial registry records) nor any indexing terms. To retrieve these records, it may be necessary to carry out a very broad search consisting of a wide range of free-text terms, which may be considered too broad to run across all the records in CENTRAL. To isolate the records in CENTRAL to those that are not sourced from PubMed or Embase, it is necessary to use the syntax below in 'Search manager': * NOT (Embase:an OR PubMed:an) (i.e. asterisk space NOT etc).

It is highly desirable that authors of Cochrane Reviews of interventions use specially designed and tested search filters where appropriate but filters should not be used in pre-filtered databases e.g. do not use a randomized trial filter in CENTRAL (MECIR C34) or attempt to apply a limit to 'human' studies in CENTRAL. All records in CENTRAL should be reports of trials in humans even though this may not be apparent from the record itself, especially for those records with no abstract.

As of June 2024, CENTRAL contains approximately 470,000 trial register records. These records have been identified as RCTs through Cochrane's Centralized Search Service and cover trial register records from ClinicalTrials.gov and the registers included in the meta-register ICTRP. An evaluation is underway to assess the comprehensiveness of CENTRAL for trial register records. A 2020 analysis concluded that CENTRAL was not yet comprehensive enough in terms of trial register report coverage (Banno et al 2020). The Cochrane Evidence Pipeline Team has worked to improve the centralized feeds and an updated analysis is due in 2024.

2.2.2 Searching MEDLINE and Embase: specific issues

Irrespective of the fact that both MEDLINE and Embase have been searched systematically for reports of trials for certain years and that these reports of trials have been included in CENTRAL, as described in Sections 2.1.1 and 2.1.2, supplementary searches of both MEDLINE and Embase are recommended (as detailed below). Any such searches, however, should be undertaken in the knowledge of what searching has already been conducted to avoid duplication of effort.

Searching MEDLINE

There can be a delay of up to one month between records being indexed as trials in MEDLINE and appearing indexed as trials in CENTRAL. This is due to the Cochrane Library monthly updating/publication cycle for CENTRAL. As a cautious approach, therefore, the most recent two months of MEDLINE should be searched, at least for records indexed as either 'Randomized Controlled Trial' or 'Controlled Clinical Trial' in the Publication Type, to identify those records recently indexed as RCTs or CCTs in MEDLINE. For further details on the search process for MEDLINE see: https://www.cochranelibrary.com/central/central-creation.

Additionally, the most recent year to be searched under the project to identify reports of trials in MEDLINE and send them back to the US National Library of Medicine for re-tagging was 2004, so records added to MEDLINE between 2005 and 2010 inclusive should be searched using one of the Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE (see Section 3.6.1). A project is planned to identify potentially missing reports from CENTRAL from this period (2005 to 2010). The project will be designed and set up as a discrete Cochrane Crowd task. (Records added to MEDLINE from 2011 onwards will have been searched as part of the Embase project described in Section 2.1.2).

Finally, for extra sensitivity, or where the use of a randomized trial filter is not appropriate, review authors should search MEDLINE for all years using appropriate free-text and thesaurus terms relevant to their review topic without any trial filter.

The MEDLINE re-tagging project described in Section 2.1.1 assessed whether the records identified were reports of trials on the basis of the title and abstract only. Any supplementary search of MEDLINE that is followed up by accessing the full text of the articles will identify additional reports of trials, most likely through the methods sections, that were not identified through the titles or abstracts alone. It is not expected, however, that accessing the full text of all articles will be routinely undertaken. For guidance on running separate search strategies in the MEDLINE-indexed versions of MEDLINE and the versions of MEDLINE containing 'in-process' and other non-indexed records please refer to Section 3.6.1.

Any reports of trials identified by the review author should be submitted to the Cochrane Information Specialist who can ensure that they are added to CENTRAL. Any errors, in respect of records indexed as trials in MEDLINE that on the basis of the full article are definitely not reports of trials according to the definitions used by the US National Library of Medicine (NLM) (see Box 2.b), should also be reported to the Cochrane Information Specialist, so they can be referred to the NLM and corrected.

For general information about searching, which is relevant to searching MEDLINE, see Section 3 and subsections.

Searching Embase

Since 2011, the Emtree term 'randomized controlled trial' has been used by Elsevier only to index records that are reports of trials, not also for records that are about trials (as was previously the case). This change in indexing practice has made the use of the term much more precise in identifying possibly relevant studies in Embase. Users can use 'randomized controlled trial (topic)' [exact Ovid syntax: "randomized controlled trial (topic)"/] to help find records about RCTs. As well as the new Cochrane Embase filter (see Section 3.6.2) other search filters for searching for trials in Embase are available on the InterTASC Information Specialists' Sub-Group website (https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/rcts).

Additionally, for extra sensitivity, or where the use of a randomized trial 'filter' is not appropriate, review authors should search Embase for all years using appropriate free-text and thesaurus terms relevant to their review topic without any trial filter, as described under similar circumstances for MEDLINE above.

It should be remembered that Cochrane's Centralized Search Service (CSS) processes are based on assessing the vast majority of records identified as potential reports of trials on the basis of the title and abstract only. A small subset of records that have been classified Unsure by 'Resolver' level screeners in the Cochrane Crowd do go to full-text assessment. To date this has accounted for less than 1% of all records screened. Therefore, any supplementary search of Embase that is followed up by accessing the full text of the articles is likely to identify additional reports of trials, probably through the methods sections, that were not identified through the titles or abstracts alone.

There is a delay of some weeks between records being indexed in Embase and appearing in CENTRAL. The most recent months of Embase should, therefore, be searched. For more details on the Embase records workflow, go to: https://www.cochranelibrary.com/central/central-creation. Also see Table 2.1.a.

In 2011, Elsevier began to include all MEDLINE content in Embase. Before then, there had always been a sizable but not complete overlap in content between the two sources. Currently (as of June 2024), Embase provides access to approximately 12,000 currently published journals. This includes approximately 3,000 journals that are not covered by MEDLINE (https://www.elsevier.com/products/embase/content?trial=true). A search of MEDLINE, either through PubMed or through another third-party interface that includes PubMed records, is, however, still necessary. There are records in MEDLINE which have the status: PubMed-not-MEDLINE. Records with this status are "Citations that will not receive MEDLINE indexing because they are for articles in non-MEDLINE journals, or they are for articles in MEDLINE journals but the articles are out of scope, or they are from issues published prior to the date the journal was selected for indexing, or citations to articles from journals that deposit their full text articles in PMC but have not yet been recommended for indexing in MEDLINE" (https://pubmed.ncbi.nlm.nih.gov/help/#citation-status-subsets). In addition, a study found that records from MEDLINE were not always retrieved when searched through Embase due to MeSH not being available in Embase (Bramer et al 2017a). Although it is, therefore, technically possible to search across all MEDLINE records in Embase (note, not all PubMed records), it is recommended that both databases should be searched separately.

As noted above, in 2009 Elsevier began indexing conference abstracts for Embase and about 5 million conference abstracts from about 15,000 conferences (as of June 2024) are now indexed in Embase. Elsevier provides a list of conferences they index for Embase, as mentioned above: https://www.elsevier.com/products/embase/content?trial=true. Conference abstracts can be a rich source of RCT evidence. Within Embase, these records have been indexed using automated indexing procedures, and in most cases the index terms applied automatically are

about subject topics or content rather than study type. In addition, many conference abstracts have been retrospectively added to Embase, some of which have been assigned an entry date prior to the publication date of the conference abstract itself. The Embase project has made, and continues to make, efforts to identify conference records added retrospectively. It should be noted, however, that the project may not yet have identified all relevant conference publications.

2.2.3 Searching ClinicalTrials.gov and WHO ICTRP: specific issues

As noted above, records from trials registers are included in CENTRAL. They are added as part of Cochrane's Centralized Search Service (CSS) (see Section 2.1). Two trials register sources are searched centrally for CENTRAL: ClinicalTrials.gov and the meta-register WHO ICTRP. To identify trials register records in CENTRAL, users can search for: Trial registry record: pt.

Records that are sourced for CENTRAL from trials registers undergo formatting to enable their publication in CENTRAL. This involves field mapping to ensure that the most useful information in the source register record is available for searching in the CENTRAL record, however, not all metadata from the source records is available in CENTRAL. The list of fields available in CENTRAL for ClinicalTrials.gov and ICTRP records is displayed along with their CENTRAL field mapping in the How CENTRAL is created file at: https://www.cochranelibrary.com/central/central-creation. In particular, records sourced for CENTRAL from ClinicalTrials.gov do not contain the trial results in CENTRAL, even if the original record within ClinicalTrials.gov contains posted results. Nor is there any indication in the CENTRAL record as to whether the original record within ClinicalTrials.gov contains posted results and/or associated publications. Similar issues apply to records sourced from the ICTRP. Direct searches at source of ClinicalTrials.gov and the ICTRP will, therefore, identify not only records not identifiable in CENTRAL but will also provide useful information about any posted results and/or related publications.

MeSH is available for ClinicalTrials.gov records. MeSH terms, however, have been added at source (i.e. by the creators of the source record, not by the US NLM) and are often minimal. Searches should therefore not be based on MeSH alone but on a combination of both free text and MeSH terms.

There can be a delay of up to one month between records being indexed as trials in CINAHL and appearing indexed as trials in CENTRAL. This is due to the Cochrane Library monthly updating/publication cycle for CENTRAL. As a cautious approach, therefore, the most recent two months of CINAHL should be searched, at least for records indexed as either 'Randomized Controlled Trial' or 'Controlled Clinical Trial' in the Publication Type, to identify those records recently indexed as RCTs or CCTs in CINAHL. For further details on the search process for CINAHL see: https://www.cochranelibrary.com/central/central-creation.

Additionally, for extra sensitivity, or where the use of a randomized trial 'filter' is not appropriate, review authors should search CINAHL for all years using appropriate free-text and

thesaurus terms relevant to their review topic without any trial filter, as described under similar circumstances for MEDLINE above.

It should also be noted that records published in CENTRAL from CINAHL will not have abstracts unless those records have also been identified in PubMed/MEDLINE or Embase. Therefore care should be taken to ensure that the CENTRAL search contains appropriate CINAHL subject heading terms, which will be searched as free text terms in CENTRAL.

2.2.4 Searching CINAHL: specific issues

The Cumulative Index of Nursing and Allied Health Literature (CINAHL) contains over 8 million records indexed from 5,500 journals dating back to the 1930s. CINAHL indexes journals in nursing, physiotherapy, occupational health and other allied health areas. It also contains dissertations and theses.

CINAHL uses a thesaurus, called CINAHL subject headings, which follow the same structure as NLM's MeSH, but with additional terms specific to nursing and allied health. CINAHL subject headings are updated on an annual basis though articles published between 1937-1961 do not have subject headings assigned. These records are known as 'pre-CINAHL' records, as are records that are in the process of being indexed.

Major and minor subheadings are used to denote the relevance of the term to the article being indexed. All research study types and research methodology terms indicating the type of research undertaken (e.g. qualitative study, cohort study, clinical trial, odds ratio etc.) are assigned as minor headings. Subheadings can be searched independently of the Subject Headings (i.e., as free floating subheadings) using a two letter code. More detail on CINAHL subject headings can be found here: https://connect.ebsco.com/s/article/Advanced-Searching-with-CINAHL-Subject-Headings?language=en_US. A sensitive search of CINAHL should ideally include both CINAHL subject headings and free text terms.

Proximity searching is enabled in CINAHL using two proximity operators: 'N' for Near, and 'W' for Within. The Near operator will return records where the words are in any order; using the Within operator retains the words in the order that they were entered into the search string.

Phrase searching within CINAHL requires the use of quotation marks as do field tags which must also be entered in upper case.

2.3 Summary points

- Cochrane Review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, on the search process.

- Authors of non-Cochrane reviews should seek advice from a medical/healthcare librarian or information specialist, with experience of conducting searches for studies for systematic reviews.
- The key databases to be searched are the Cochrane Group's Specialized Register (internally, e.g. via the Cochrane Register of Studies, or externally via CENTRAL), where such Specialized Registers exist, CENTRAL, MEDLINE and Embase (if access is available to either the review author or the Cochrane Group).
- Approximately 1,000,000 of the 2,000,000 records in CENTRAL are from MEDLINE or Embase, so care should be taken when searching MEDLINE and Embase to avoid unnecessary duplication of effort.
- Supplementary searches of Embase and MEDLINE should be carried out as outlined in Section 2.2.2.
- Supplementary searches of ClinicalTrials.gov and WHO ICTRP should be carried out as outlined in Section 2.2.3.
- Supplementary searches of CINAHL should be carried out as outlined in Section 2.2.4.
- Additional studies can be identified in MEDLINE, Embase and CINAHL by searching across the years already searched for CENTRAL, by obtaining the full article and by reading, in particular, the methods section, however, it is not expected that accessing the full text of all articles will be routinely undertaken. The same issues apply to other sources searched for CENTRAL.
- Additional studies can also be identified in ClinicalTrials.gov and the WHO ICTRP Portal by searching across the years already searched for CENTRAL and by viewing the full record.

3 Designing search strategies: further considerations

This section should be read in conjunction with Chapter 4, Section 4.4. For discussion of some of the specific issues around searching for medical devices, please refer to this recent brief method note (Cooper et al 2022).

3.1 Service providers and search interfaces

Access to MEDLINE, Embase and other general and subject-specific databases is offered by several commercial service providers, via a range of search interfaces. In addition, the US National Library of Medicine, provider of MEDLINE, and Elsevier, provider of Embase, offer access to their own versions of their databases: MEDLINE through PubMed, which is available free of charge on the internet, and Embase through Elsevier directly, which is known as Embase.com and is available on subscription only. Each interface offers certain functionalities and unique features (Bethel and Rogers 2014) but more importantly the search syntax varies

across the interfaces. For example, to search for the Publication Type term 'Randomized Controlled Trial' in MEDLINE via different search interfaces it is necessary to enter the term as:

- PT (Randomized Controlled Trial OR Equivalence Trial OR Pragmatic Clinical Trial) (in MEDLINE on EBSCOhost)
- exp Randomized Controlled Trial/ (in MEDLINE on Ovid)
- DTYPE (Randomized Controlled Trial OR Equivalence Trial OR Pragmatic Clinical Trial) (in MEDLINE on ProQuest) and
- Randomized Controlled Trial[pt] (in PubMed).

Note: neither EBSCOhost nor ProQuest supports 'exploding' within the publication type field, so, for these two service providers, it is necessary to include all the narrower terms in the search. PubMed terms 'explode' automatically when there are narrower terms available.

Although the interfaces may offer access to the same database, running the same strategy in the same database but through different interfaces may result in different search results (Schoonbaert 1996, Younger and Boddy 2009, Boeker et al 2013b, Craven et al 2014, Burns et al 2019, Burns et al 2021). Refer to the search help sections of the guidance provided by the relevant service providers and for the specific databases for further information on specific syntax etc.

In addition to accessing bibliographic records, many service providers offer links to full-text versions of articles on other publishers' websites, such as the PubMed 'full text links' feature. Developments in the publishing industry also allow users to add the DOI number, where available, after the text 'https://doi.org/' to retrieve the permanent location of an article on the internet.

3.2 Controlled vocabulary and text words

MEDLINE and Embase (and many other databases) can be searched using a combination of two retrieval approaches. One is based on text words (terms occurring in the title, abstract or other relevant fields) in a record. The other is based on standardized subject terms assigned to the record either by indexers (specialists who appraise the article/reference and describe it by assigning terms from a specific thesaurus or controlled vocabulary) or automatically using automated indexing, with some degree of human curation. Standardized subject terms are useful because they provide a complementary way of retrieving records that may use different text words to describe the same concept and because they can provide information beyond that which is contained in the words in the title and abstract. Therefore, each concept of a robust search strategy should consist of text words together with subject terms, if the latter are available in the respective database.

It is mandatory, for Cochrane Reviews of interventions, to identify appropriate controlled vocabulary (e.g. MeSH, Emtree, including 'exploded' terms) (see below for definition of 'exploded' terms) (MECIR C33). When searching for studies for a systematic review, however, the extent to which subject terms are applied to references should be viewed with caution. Authors may not describe their methods or objectives well and indexers are not always experts in the subject areas or methodological aspects of the records that they are indexing. In those cases where subject terms are applied as a result of automated/machine indexing, this may not be as accurate as human indexing. In addition, the available indexing terms might not correspond to the terms the searcher wishes to use. It is, therefore, mandatory, for Cochrane Reviews of interventions, to identify appropriate free-text terms (considering, for example, spelling variants, synonyms, acronyms, truncation and proximity operators (MECIR C33). This is especially important, as the indexing process in databases takes time (ranging from a few days to several months until a reference is fully indexed). Therefore, very recent references might not yet be indexed and will consequently not be retrieved when using controlled vocabulary alone. Consideration should be given to searching indexed records and non-indexed/in-process records separately in databases such as MEDLINE and Embase which include both indexed and non-indexed content.

The approaches for identifying text words and controlled vocabulary to combine appropriately within a search strategy are presented in the following two sections and can generally be described as being subjective. Text mining is an emerging approach to identify terms in a more objective way, based on a set of relevant records on the topic (see Section 3.2.3 on text mining for term selection). Another objective method is based on similarity calculations derived from one or several known relevant articles. In MEDLINE, having identified a key article, additional relevant articles can be located by using the 'Find Similar' option in Ovid or the 'Similar articles' option in PubMed. The value of using a complementary search approach such as this feature, which is independent of the searcher's expertise, has been described (Sampson et al 2016).

3.2.1 Identifying relevant controlled vocabulary

In order to identify as many relevant records as possible, searches should include subject terms selected from the controlled vocabulary or thesaurus ('exploded' where appropriate - see below for definition of 'exploded' terms). The controlled vocabulary search terms for MEDLINE (Medical Subject Headings, known as MeSH) and Embase (Emtree) are not identical, and neither is the approach to indexing. For example, the pharmaceutical or pharmacological aspects of an Embase record are generally indexed in greater depth than the equivalent MEDLINE record, and in recent years Elsevier has increased the number of index terms assigned to each Embase record. Searches of Embase may, therefore, retrieve additional articles that were not retrieved by a MEDLINE search, even if the records were present in both databases. The converse also applies in that MEDLINE records available in Embase, which are not also indexed by Elsevier for Embase, are indexed differently in Embase than they were originally in MEDLINE, as the MeSH terms are replaced in Embase by Emtree terms. Thus, search strategies need to be customized for each database and should ideally be run in the original database whenever possible.

Most database interfaces offer a browsing option to show the preferred subject headings. For example, interfaces to MEDLINE will usually permit browsing the Medical Subject Headings (MeSH) so that the term definition (Scope Note) and its synonyms and related terms can be identified and assessed for relevance. Additional controlled vocabulary terms should be identified using the search tools provided with the database, such as the 'Permuted Index' or 'Map Term' under 'Search Tools' in Ovid or the 'MeSH Database' option in PubMed. As well as searching the controlled vocabulary lists, it is also common practice to identify subject headings from known relevant records. A tool which can help displaying and comparing the subject terms assigned to MEDLINE records is the 'Yale MeSH Analyzer' (https://mesh.med.yale.edu/) (Hocking 2017).

Many database thesauri offer the facility to 'explode' subject terms to include more specific terms automatically in the search. For example, a MEDLINE search using the MeSH term BRAIN INJURIES, if exploded, will automatically search not only for the term BRAIN INJURIES but also for the more specific term SHAKEN BABY SYNDROME. As articles in MEDLINE on the subject of shaken baby syndrome should only be indexed with the more specific term SHAKEN BABY SYNDROME and not/not also with the more general term BRAIN INJURIES, it is important that MeSH terms are 'exploded' wherever appropriate, in order not to miss relevant records. It is equally important, however, that MeSH terms are not 'exploded' where this is inappropriate, in order not to add irrelevant records unnecessarily. The same principle applies to Emtree when searching Embase and also to several other databases. For further advice on this topic, review authors should consult a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist.

A second option which can be applied to subject terms, is restricting the term to 'Major Topic' (in Ovid this feature is called 'focus'). When this feature is used, articles are only retrieved where the subject term has been assessed during indexing as reflecting one of the article's major topics. This is, therefore, a precision-maximizing feature and is not generally recommended in the context of searching for studies for systematic reviews, as it compromises sensitivity.

It is particularly important in MEDLINE to distinguish between Publication Type terms and other related MeSH terms. For example, a report of a randomized trial should be indexed in MEDLINE with the Publication Type term 'Randomized Controlled Trial' whereas an article about randomized controlled trials should be indexed with the MeSH term RANDOMIZED CONTROLLED TRIALS AS TOPIC (note the word TRIALS in the latter is plural). The same applies to other indexing terms for other trials, reviews and meta-analyses. It should be noted that this distinction was also introduced into Embase for records added from 2011 onwards. The Emtree term 'randomized controlled trial' is used to describe the publication type of the record, whereas the Emtree term 'randomized controlled trial (topic)' is used for records that discuss randomized trials, but are not original reports of randomized trials. Prior to 2011, the Emtree term 'randomized controlled trial' was used to index both the publication type of the record and for records that discussed randomized trials as a topic.

Review authors should assume that earlier articles are even harder to identify than recent articles. For example, abstracts are not included in MEDLINE for most articles published before 1976 and, therefore, text word searches will only apply to titles. In addition, few MEDLINE indexing terms relating to study design were available before the 1990s, so text word searches relating to study design are necessary to retrieve older records.

3.2.2 Identifying relevant text words

Relevant text words (i.e. free-text terms) can be identified by checking the terms used in the title, abstract and other relevant fields (e.g. author keywords) of a sample of relevant references. It is important to be aware of the fact that natural language allows concepts to be expressed in different words. It is essential, therefore, to look up synonyms for each concept describing the review topic. Medical dictionaries can be used to clarify definitions and identify synonyms. The MeSH database also offers both definitions (Scope Notes) and a listing of synonyms and related terms for each MeSH term ('Entry terms'), which lists different terms being used for a concept. Likewise, Elsevier's Emtree thesaurus for Embase also lists synonyms for each term. Synonyms of pharmaceutical substances can be effectively searched via PubChem (https://pubchem.ncbi.nlm.nih.gov/). A third approach for identifying text words consists of checking search strategies from other systematic reviews on a similar topic.

3.2.3 Text mining, machine learning and artificial intelligence for term selection and strategy building

Text mining, machine learning and artificial intelligence are increasingly used in the conduct of systematic reviews generally and have been the subject of a number of helpful reviews (O'Mara-Eves et al 2015, Paynter et al 2016, Stansfield et al 2017, Kohl et al 2018, Adam et al 2022, Adam and Paynter 2023, Schmidt et al 2023). There are many tools to help with term identification, to identify topics within sets of search results, to identify studies similar to known relevant studies (from which to identify further terms) and to translate strategies.

Text mining encompasses a range of statistical approaches to textual analysis including simple frequency analysis of words and phrases within records, visual presentations of the inter-relationships between concepts in a literature (corpus) and the development of complex interrogation rules to identify relevant records from a corpus of records (O'Mara-Eves et al 2015, Paynter et al 2016, EUnetHTA JA3WP6B2-2 Authoring Team 2019, Haddaway et al 2020). The value of text mining can lie in its ability to process large volumes of records objectively, to assist with concept identification and to interrogate large numbers of records from many databases using a single search process. At present, there are no free to access tools that can fully automate the design and running of searches across several databases and export results into one file, but parts of the strategy design process can benefit from text mining techniques combined with traditional searching approaches.

Text mining software can be used to identify potential keywords, phrases and subject terms from within a set of relevant records. Various software packages are listed in the Systematic

Review Toolbox (http://systematicreviewtools.com/). This service is currently unavailable (June 2024) whilst the developers seek and transition to a new hosting solution.

Software tools such as PubMed PubReMiner (https://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi) analyse the results of searches conducted in PubMed and present the words within records in order of frequency. This can aid the identification of terms, synonyms and abbreviations to test out in strategies. For databases other than MEDLINE (PubMed) frequency analysis software such as Voyant Tools (https://voyant-tools.org/) will provide similar frequency analyses or bibliographic reference software such as EndNote (https://endnote.com/) can be used with any database records (McGowan 2021). In EndNote, frequency analysis can be achieved by using the Term Lists and the Subject Bibliography option (detailed guidance at https://sites.google.com/york.ac.uk/yhectrainingpages/home/endnote-for-frequency-analysis). Frequency analysis is also available in some systematic review software such as Nested Knowledge (https://wiki.nested-knowledge.com/doku.php?id=wiki:autolit:search:exploration) and in some database interfaces such as Embase.com.

A tool to assist with identifying relevant MeSH headings is available on the MeSH on Demand website (https://meshb.nlm.nih.gov/MeSHonDemand): it is possible, for example, to paste in a Cochrane protocol and receive suggestions of MeSH terms that relate to the topics within the text.

Software tools to assist in identifying phrases and words within proximity to each other are also available in Voyant Tools, TerMine (http://www.nactem.ac.uk/software/termine/) and many other packages.

A recent analysis of selected tools recommends Text Analyzer, and the Yale MeSH Analyzer as “extremely useful” tools for term identification, and rates Anne O’Tate, PubMed PubReMiner and Voyant as “useful” tools (O'Keefe et al 2023). This analysis favours ease of use over functionality, so users seeking expert functionality may wish to explore other tools which are rated lower in this article.

Procedures to develop search strategies routinely using text mining approaches are available (Hausner et al 2012, Hausner et al 2015, EUnetHTA JA3WP6B2-2 Authoring Team 2019) and comparisons of text mining and traditional search approaches have been published (Paynter et al 2021, Scells et al 2021, Hausner et al 2022, Paynter et al 2022).

Text mining has also been used to develop methodological search filters, including the Cochrane Highly Sensitive Search Strategies for MEDLINE and Embase (Glanville et al 2006, Glanville et al 2019a) and a filter to identify overviews of systematic reviews in MEDLINE (Lunny et al 2015). Researchers are also exploring machine learning approaches to converting searches in one database to search in very different databases, such as converting PubMed searches to interrogate records in ClinicalTrials.gov (Lanera et al 2018).

Text mining may be particularly helpful when developing strategies for complex topics. Software such as VOSviewer (https://www.vosviewer.com/) can accept large numbers of records, analyse the co-occurrence of terms within records and show relationships between themes in a body of records visually. This can help with identifying, grouping and combining concepts when building strategies for complex topics (Balan et al 2014, EUnetHTA JA3WP6B2-2 Authoring Team 2019, Arruda et al 2022, Sheikh et al 2023). Thematic analysis of sets of results is also available in tools including Carrot2 (https://search.carrot2.org/#/search/web) and systematic review software such as EPPI-Reviewer and Nested Knowledge.

Text mining and machine learning tools available free of charge on the internet can also assist with identifying additional relevant studies. Tools such as Medline Ranker (http://cbdm-01.zdv.uni-mainz.de/~jfontain/), and LitSuggest (https://www.ncbi.nlm.nih.gov/research/litsuggest/) can rank search results in order of similarity to known relevant records specified by the searcher (Fontaine et al 2009, Allot et al 2019, Simon et al 2019, Zhang et al 2022). Information specialists report using unsupervised machine learning with the DoCTER site (https://www.icf-docter.com/) to "get quick, data-driven insights into a set of search results" (Cawley 2022). DoCTER helped the information specialists to identify concepts to exclude from searches and to identify a priority set of studies within a larger set of results. There are many other tools available to identify additional relevant studies based on known relevant records (seeds) which make use of algorithms that seek studies that are similar to seed studies (Kreutz and Schenkel 2022). These include the "Similar articles" option within PubMed and resources that search for similar sentences within a large corpus (which can include full text documents), such as LitSense (https://www.ncbi.nlm.nih.gov/research/litsense/) or Elicit (https://elicit.com/) (Allot et al 2019).

As well as ready-to-use internet tools, researchers have created software tools to carry out term identification and other search tasks that can be run in R or programming languages such as Python. Examples include litsearchr (https://elizagrames.github.io/litsearchr/) (Crisan et al 2019, Grames et al 2019, McGowan 2021), Ananse (https://baasare.github.io/ananse/ build/html/index.html)(Kwabena et al 2022) and PMIDigest (https://github.com/JNovoaR/PMIDigest)(Novoa et al 2023). R code, shinyapps and tools in programming languages such as Python can be accessed from public resources such as GitHub (https://github.com/) (Mesgarpour et al 2016). Using these tools may require some knowledge of programming languages or how to download and run the tools. The advantage of many of these tools is that they offer searchers more control over text analyses and more sophisticated functions than the easy to use "off the shelf" tools described above. Developing new tools will typically require a knowledge of R or programming languages.

Instead of using text mining tools to identify search terms and build strategies, researchers are exploring machine learning as an alternative to the typical approach of searching databases with detailed search strategies (Roth and Wermer-Colan 2023). If machine learning software can be trained to identify relevant records, then a sensitive search approach could be used to

identify records from a series of databases or large internet resources such as Dimensions AI (https://www.dimensions.ai/) or OpenAlex (https://openalex.org/)(Hair et al 2021). The records could then be processed using machine learning software to identify relevant studies. This means less time might be spent in searching, with searches being very broad (highly sensitive) rather than detailed and relatively focused on the search topic. This approach could mean that searching time might be reallocated to training software to recognize relevant records. This approach is already being used by health technology assessment teams and seems to be of particular interest in the context of updating reviews (Del Fiol et al 2018, Muller et al 2021, Røst et al 2021, Stansfield et al 2022, Shemilt et al 2024). The COVID-19 pandemic has been a recent key testbed for the development of resources using combinations of text mining and machine learning (Wang and Lo 2021). Examples of resources built with automated techniques include the Cochrane COVID-19 Study Register (https://covid-19.cochrane.org/) (Shemilt et al 2022) and the Epistemonikos L·OVE (Living OVerview of Evidence) platform (https://iloveevidence.com/)(Metzendorf et al 2023). The ACCESSSS database (https://www.accessss.org/)) of pre-appraised evidence is being built using machine learning to identify high quality research studies (Lokker et al 2023).

Text mining and machine learning have the potential to involve the acquisition and interrogation of large volumes of search results, so it is important to be aware of both licensing and copyright issues involved in handling large sets of records (Sag 2019).

Although text mining tools and machine learning tools have great potential to support search strategy design, there are many variants and options to choose from and little guidance about what works best and when and for which questions. There is a need for more case studies and for more parallel research to show where benefits may lie. Text mining carries with it challenges in terms of documentation of the processes used and there is little guidance available on how best to report the use of text mining for strategy development.

There is growing interest in the use of general purpose chatbots such as ChatGPT or Claude, built using Large Language Models (LLMs) and more specialized artificial intelligence (AI) tools trained on academic literature (such as Consensus and SciSpace), to assist in systematic review processes. Systematic reviews and mapping reviews of the use of AI in the context of systematic review searches so far indicate the emergence of many tools (Felizardo and Carver 2020, Napoleão et al 2021, van Dinter et al 2021, Cierco Jimenez et al 2022, Sallam 2023, Hersh 2024, Jin et al 2024), but limited applications in search strategy design. Learning from published research is challenging because AI tools are developing quickly and the relevance of findings may change rapidly. LLMs are typically trained on data up to a certain date, which may be important in fields where terminology is changing quickly or if the tools are being used for record retrieval as well as search design. Researchers have noted that LLMs are trained on published studies which in the case of systematic reviews often have poor quality searches, leading to concerns about the quality of search-specific recommendations from AI tools (Levay and Craven 2023). Experiments using ChatGPT for developing Boolean queries in January 2023 suggested the need for well-developed prompts to specify and refine the search query, but also

warned that ChatGPT may generate incorrect MeSH and that different terms may be suggested if a prompt is run at different times (Wang et al 2023). A user evaluation of AI search tools iris.ai and Yewno by the Royal Library of Denmark found that users could see the potential for the tools, but only in the exploratory phase of the literature search (Wildgaard et al 2023). Unrelated to the search prompts being used, the results of searches undertaken with general AI services should be carefully assessed since ChatGPT, for example, has known problems in terms of so-called "hallucinations" (fabricated references) and inaccurate references (Bhattacharyya et al 2023). It should be noted that "hallucinations" have been reported to be less of an issue in ChatGPT-4 and with chatbots trained on academic literature (Gillin 2024).

There are AI tools that search PubMed and the research literature more widely and offer search strategies to support the answers they provide. For example, EvidenceHunt (https://evidencehunt.com/) is a search plugin for ChatGPT that identifies and summarizes evidence that answers research questions using PubMed and offers search strategy suggestions. A comparison of ChatGPT and Elicit to evaluate their performance in finding known studies for a completed review reported that ChatGPT generated so-called "hallucinations" (fabricated publications) but Elicit was more helpful in terms of accurate recall (Enomoto et al 2023).

There is a large and evolving literature comparing the results of queries of general purpose AI tools such as ChatGPT to known best evidence and best practice, for a range of conditions and questions. Before using general AI tools to search for research evidence it might be useful to search for studies that have evaluated the accuracy and reliability of the results for questions in the topic of interest. Alternatively, AI tools might be used that are trained on the biomedical or wider academic literature.

Database providers are offering AI-informed research discovery tools as add-ons or features within their standard interfaces. Examples include the Summon discovery tool offered by Clarivate, the EBSCO Discovery service, Elsevier's ClinicalKey AI and Scopus AI. At present there are few formal evaluations of these facilities, but if they are offered by a provider then they may be worth exploring to evaluate how they may support strategy development.

Large language models, and AI tools that make use of them, seem likely to help with the development of search strategies and the identification of relevant studies, but should be used cautiously for the reasons described above. They should be used as a further search approach, in addition to the more traditional search methods described in this Handbook.

3.3 Synonyms, related terms, variant spellings, truncation and wildcards

In order to be as comprehensive as possible, it is necessary to include a wide range of free-text terms for each of the concepts selected. This might include the use of truncation and wildcards.

It is mandatory, for Cochrane Reviews of interventions, to identify appropriate spelling variants, synonyms, acronyms and truncation (MECIR C33). For example:

- synonyms: ‘pressure sore’ OR ‘decubitus ulcer’
- related terms: ‘brain’ OR ‘head’ and
- variant spellings: ‘tumour’ OR ‘tumor’.

Database interfaces offer functionality to capture these variations through truncation and wildcards. For example:

- truncation: random* (for random or randomised or randomized or randomly, etc.) and
- wildcard: wom?n (for woman or women).

These features vary across different database interfaces, especially with respect to truncation length (e.g. number of characters) and position (e.g. mid-word or end-of-word), and should be checked carefully before adapting a search strategy to a different database and/or interface from that for which it was originally designed. For further details refer to the respective database help files. It should also be noted that many service providers incorporate fuzzy logic searching into their search interfaces and this automatically includes variant endings by default including singular and plural variants.

3.4 Boolean operators (AND, OR and NOT)

Boolean operators are used to join together the search terms within a search strategy. The most widely used Boolean operators are:

- AND: combines *different concepts* to make a set of results that is usually smaller than the smallest concept (i.e. terms from all concepts need to be present in records for them to be retrieved)
- OR: gathers terms *within a concept* and this usually makes the set of results larger (i.e. at least one term needs to be present in records for them to be retrieved) and
- NOT: excludes terms or concepts (one term or concept can be excluded from the set of results and the set will usually reduce in size – but see caveats below).

Generally speaking, a search strategy should build up the controlled vocabulary terms, text words, synonyms and related terms for each concept (such as the intervention), one concept at a time. Terms within a concept should normally be combined with the Boolean ‘OR’ operator: see demonstration search strategies in Box 3.h and Box 3.i . This means records will be retrieved that contain at least one of these search terms. Sets of terms should usually be developed for the different concepts being searched such as the healthcare condition and intervention(s) with or without the study design. These two/three concepts (sets of terms) can then be combined using the ‘AND’ operator. This combination step results in a set of records that are likely to address both the health condition of interest and the intervention(s) to be

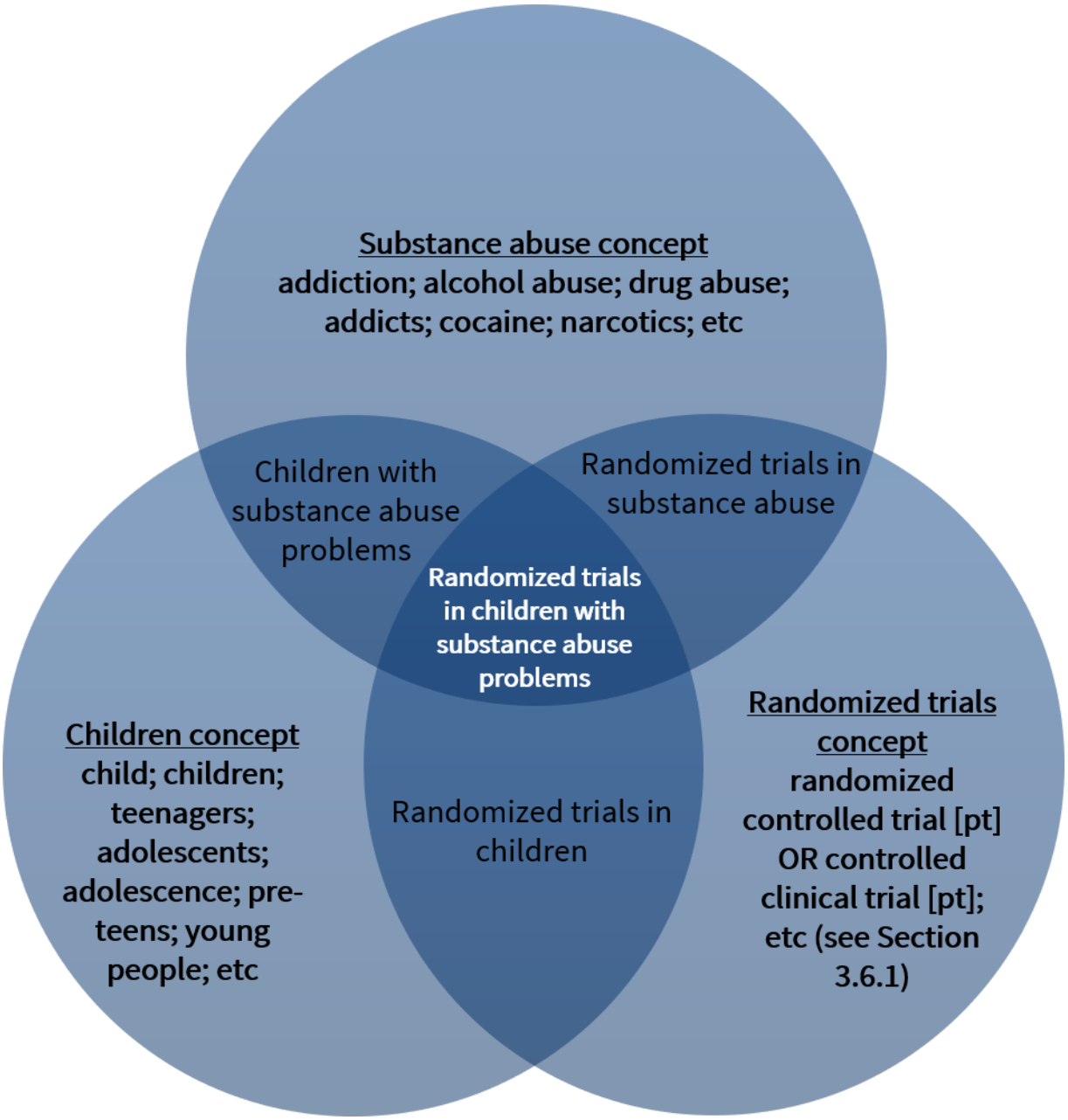
evaluated, together with the appropriate study design, if study design was also included as a concept (see Figure 3.a). It is mandatory, for Cochrane Reviews of interventions, to ensure correct use of the 'AND' and 'OR' operators (MECIR C32).

A note of caution about this approach is warranted. If a record does not contain at least one term from each of these two or three sets, it will not be identified. For example, if an index term has not been added to the record for the intervention and the intervention is not mentioned in the title or abstract, the record would be missed by the strategy. The best approach is to begin with as few concepts as possible and only add additional concepts if record numbers are unmanageable. So a search might begin with only one concept, and the study design concept might only be added if essential. Note that for searches in the Cochrane Central Register of Controlled Trials (CENTRAL), where all records should refer to controlled trials, the study design concept should not be included in the search strategy.

The 'NOT' operator should be avoided where possible to avoid inadvertently removing from the search set any records that might be relevant. For example, when searching for records indexed as female, the use of 'NOT male' would remove any record that was about both males and females. NOT can be used in some situations where care is taken to ensure that relevant records are not lost, for example in the animal exclusion algorithm used within the MEDLINE, Embase and CINAHL search filters to identify RCTs (see Section 3.6 and subsections).

Searches to identify studies for Cochrane Reviews can sometimes be extremely long, often including over 100 search lines. It can be tedious to type in the combinations of these search sets, for example as '#1 OR #2 OR #3 OR #4 …. OR #100'. Some service providers offer alternatives to this. For example, in CENTRAL and Ovid it is possible to combine sets using the syntax {OR #1-#100} or 'or/1-100' respectively. For those service providers where this is not possible, it has been recommended that the search string above could be created in full and saved, for example, as a Word document and the requisite number of combinations copied and pasted into the search as required. Alternatively, the search string could be copied from another search strategy, checking carefully for any typographical errors before use. Having created/identified the string with the # symbols as above, a second string can be generated by globally replacing the # symbol with nothing to create the string '1 OR 2 OR 3 OR 4 …. OR 100' to be used for those service providers where the search interface does not use the # symbol.

Figure 3.a Combining concepts as search sets



3.5 Proximity operators (NEAR, NEXT and ADJ)

Proximity operators identify search terms which are near to each other but not necessarily directly adjacent to each other, nor necessarily in that specific order. Where the operator dictates that the search terms must be directly adjacent to each other, they are often referred to as adjacency operators or, more specifically, phrase searching. It is mandatory, for Cochrane reviews of interventions, to ensure that proximity operators are used appropriately (MECIR C33). Use of proximity operators helps to ensure that searches are more sensitive than would

be the case with direct adjacency or phrase searching, and can also facilitate ease of searching where there are multiple possible variations of a phrase which would otherwise need to be typed in full.

The following proximity and adjacency operators are illustrated with reference to the Cochrane Library.

NEXT

The Cochrane Library uses the proximity operator 'NEXT' to identify search terms which are directly adjacent to each other and in the specified order. For example, diabetes NEXT screening retrieves 'diabetes screening', but not 'screening diabetes'.

'NEXT' functions in the Cochrane Library in the same way as searching for phrases within quotation marks such as "diabetes screening". Use 'NEXT' for including truncation '*' or wildcards '?' in a phrase, such as 'diabet* NEXT screen*'.

NEAR

The Cochrane Library uses the operator 'NEAR/*n*' to search for search terms within a specified number of words, where *n* specifies the maximum number of words either search term is from the other search term in any order. For example:

- diabetes NEAR/1 screening retrieves 'diabetes screening' and 'screening diabetes'
- diabetes NEAR/2 screening retrieves 'diabetes *x* screening' and 'screening *x* diabetes' where *x* is the maximum number of intervening words
- diabetes NEAR/3 screening retrieves 'diabetes *x x* screening' and 'screening *x x* diabetes' where *x* is the maximum number of intervening words.

If using only NEAR, with no number specified, then this defaults to a maximum of 6 words in the Cochrane Library. Thus 'diabetes NEAR screening' retrieves 'diabetes *x x x x x* screening' and 'screening *x x x x x* diabetes'.

Syntax variation between databases

Other database interfaces use different operators, for example, 'N*n*' in the EBSCO*host* interface or 'ADJ*n*' in the Ovid interface. Links to help pages on proximity operators for each of the main database providers are detailed at the end of this section.

It is important to note that interfaces also vary in how the number *n* relates to the specified search terms. In the Cochrane Library, Embase.com and Ovid interfaces *n* specifies the maximum number of words that either search term is from the other search term, i.e. to find a maximum of *x* words between two search terms *n* should equal *x* + 1. In the EBSCO*host*, ProQuest, Scopus and Web of Science interfaces *n* specifies the maximum number of words between the specified search terms, i.e. to find a maximum of *x* words in between two search

terms *n* should equal *x.* For example, if *n* is set to 2 it functions as shown below in the Ovid and EBSCO*host* interfaces, respectively, where *x* is an intervening word:

- diabetes N2 screening retrieves 'diabetes *x x* screening' and 'screening *x x* diabetes' (EBSCO*host*) where x is the maximum number of intervening words
- diabetes ADJ2 screening retrieves 'diabetes *x* screening' and 'screening *x* diabetes' (Ovid) where x is the maximum number of intervening words.

If *n* is set to 1 in the Ovid interface it functions as shown below:

- diabetes ADJ1 screening retrieves 'diabetes screening' and 'screening diabetes'

Searching using ADJ in the Ovid interface without specifying *n* operates in the same way as NEXT in the Cochrane Library, i.e. the search terms are retrieved but only in the specified order.

When searching using two or more search terms without quotation marks in EBSCO databases, the search terms are automatically combined using the proximity setting N5. This can be overridden by placing the terms in quotation marks, using a different proximity operator value, or combining the search terms using a Boolean operator.

Proximity operators in PubMed

Proximity search capability was recently added to PubMed (US National Library of Medicine 2022). To carry out proximity searching in PubMed, the search should be structured as follows:

"search terms" [field:~N]

- Search terms = two or more terms enclosed within double quotation marks
- Field = the search field within which the terms should appear (this is limited to either Title or Title/Abstract fields)
- N = maximum number of words which can appear between the specified search terms.

For example, to search PubMed for records where the words "hip" and "pain" appear with no more than two words in between them in the Title/Abstract field, use the following search:

"hip pain"[Title/Abstract:~2]

It is not possible to use the truncation operator (*) with the proximity search function in PubMed. It is also not possible to use phrase searching with the proximity operator function, i.e. you cannot specify that a phrase should appear next to another search term or phrase with a maximum number of words in between. More detail on using the proximity operator function in PubMed is available here.

PubMed does not support phrase searching for phrases which are not in the PubMed phrase index. To search for a phrase which is not in the phrase index, use a proximity operator setting of 0 to search for the required phrase, e.g. "cognitive impairment in multiple sclerosis"[Title/Abstract:~0].

Retaining the order of search terms

As noted above, the NEAR operator in the Cochrane Library and the equivalent operators used in other interfaces identify the specified search terms in any order. There is no option in the Cochrane Library for specifying the maximum number of words between search terms and retaining the specified order of the search terms. Some database providers do offer this option. For example, the EBSCO*host* and ProQuest interfaces retain the specified order of search terms when using the 'Wn' and 'pre/n' operators, respectively, as shown below:

- diabetes W2 screening retrieves 'diabetes *x x* screening' where x is the maximum number of intervening words (EBSCO*host*)
- diabetes pre/2 screening retrieves 'diabetes *x x* screening' where x is the maximum number of intervening words (ProQuest).

Help pages for proximity operators

Listed below are help links on how to use proximity operators produced by the main database providers. Some of these links go directly to the proximity operators help section and others require searching for the proximity operators section within them.

The Cochrane Library databases

https://www.wiley.com/en-us/customer-success/cochrane/cochrane-library-user-guide

EBSCO databases

https://connect.ebsco.com/s/article/How-do-I-create-a-proximity-search?language=en_US

Ovid databases (scroll down to Adjacency/Defined Adjacency)

https://ospguides.ovid.com/OSPguides/medline.htm

ProQuest databases

https://parlipapers.proquest.com/help/parlipapers/Search_Tips.html

PubMed database (Automatic Term Mapping)

https://pubmed.ncbi.nlm.nih.gov/help/#automatic-term-mapping

PubMed database (Searching for a Phrase)

https://pubmed.ncbi.nlm.nih.gov/help/#searching-for-a-phrase

Scopus database (Elsevier)

https://blog.scopus.com/posts/6-simple-search-tips-lessons-learned-from-the-scopus-webinar

Web of Science databases (Clarivate Analytics)

https://images.webofknowledge.com/WOKRS58B4/help/WOS/hs_search_operators.html#dsy862-TRS_proximity

3.6 Search filters

This section should be read in conjunction with Chapter 4, Section 4.4.7.

3.6.1 The Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE

The first Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE was designed by one of the authors (CL) and published in 1994 (Dickersin et al 1994). This strategy was thereafter published in subsequent editions of this *Handbook* and has been adapted and updated as necessary over time. The Cochrane Highly Sensitive Search Strategies for MEDLINE, in subsequent sections, are adapted from strategies first published in 2006 as a result of a frequency analysis of MeSH terms and free-text terms occurring in the titles and abstracts of MEDLINE-indexed records of reports of randomized trials (Glanville et al 2006), using methods of search strategy design first developed by the authors to identify systematic reviews in MEDLINE (White et al 2001).

Two strategies are offered: a sensitivity-maximizing version and a sensitivity- and precision-maximizing version. It is recommended that searches for trials for inclusion in Cochrane Reviews begin with the sensitivity-maximizing version in combination with a highly sensitive subject search. If this retrieves an unmanageable number of references the sensitivity- and precision-maximizing version should be used instead. See Sections 2.1.1, 2.1.2, 2.2.1 and 2.2.2. for details as to how these search strategies and others have been run centrally in Cochrane over the years and relevant records included in CENTRAL, to avoid unnecessary duplication of effort.

The strategies have been updated pragmatically, to reflect changes in search syntax and changes in indexing policy introduced by the US National Library of Medicine since the original analysis. These changes include:

- the change of the MeSH term CLINICAL TRIALS to CLINICAL TRIALS AS TOPIC
- no longer assigning ‘Clinical Trial’ as a Publication Type to all records indexed with ‘Randomized Controlled Trial’ or ‘Controlled Clinical Trial’ as a Publication Type; and

- the introduction of two more specific (i.e. narrower) headings to the Publication Type RANDOMIZED CONTROLLED TRIAL, which means it should now be exploded.

A recent performance review has demonstrated that the Cochrane RCT filter continues to rank highly in terms of sensitivity (Glanville et al 2020). The strategies are given in Box 3.a and Box 3.b for PubMed and in Box 3.c and Box 3.d for Ovid.

The strategies below are based on data derived from MEDLINE-indexed records and were designed to be run in MEDLINE. These strategies were not specifically designed to retrieve non-MEDLINE records in PubMed or, for example, those records in the Ovid segments: 'Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations'. However, as noted above, the Cochrane RCT filter performed highly in terms of sensitivity in a recent performance test suggesting that the concerns around performance in unindexed studies may be allayed (Glanville et al 2020). However, for those who do have concerns that the filters may miss unindexed studies, the filters could be run in the 'Ovid MEDLINE(R) ALL/PubMed(R) 1946 to Present' Ovid segment and the status field (ST) limits below could be used to isolate the MEDLINE-indexed and the non-indexed records as follows:

- all records in the database: *docz.dz.*
- MEDLINE status: *medline.st.* (i.e. MEDLINE-indexed).
- Publisher - ahead of print status: *publisher.st.*
- In-process & non-indexed citations: *("in data review" or in process or "pubmed not medline").st.*
- Pmcbooks: *nb$.bk.*

The use of the various status limits and how they add up to all records in the entire MEDLINE on Ovid database (generated by the search term docz.dz.) is shown below:

Ovid MEDLINE(R) ALL <1946 to February 20, 2024>

#	Searches	Results
1	docz.dz.	36,913,763
2	limit 1 to medline	31,202,830
3	limit 1 to publisher	334,800
4	limit 1 to ("in data review" or in process or "pubmed not medline")	5,342,882

5	nb$.bk.	33,251
6	2 or 3 or 4 or 5	36,913,763

For identifying non-indexed records a range of truncated free-text terms would be required, such as (in Ovid syntax) random$, placebo$, trial$, etc., and the search must not be limited to humans (as the records may not yet be indexed as humans).

As discussed in Section 2.1.1, MEDLINE has been searched from 1966 to 2004 inclusive, using previous versions of the Cochrane Highly Sensitive Search Strategy for identifying randomized trials, and more recent MEDLINE records (from 2011) have been searched as part of the current Embase project. All reports of trials identified in these ways (predominantly on the basis of the titles and abstracts only) are now included in CENTRAL (see Sections 2.1.1 and 2.1.2). For further guidance as to the appropriate use of these Highly Sensitive Search Strategies see Section 2.2.2. Alternatively, other search filters for identifying reports of RCTs in MEDLINE can be identified from the InterTASC Information Specialists' Sub-Group's Search Filter Resource (https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/rcts. Filter selection should be informed by an assessment of the filter's performance.

Box 3.a Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format

#1	randomized controlled trial [pt]
#2	controlled clinical trial [pt]
#3	randomized [tiab]
#4	placebo [tiab]
#5	drug therapy [sh]
#6	randomly [tiab]
#7	trial [tiab]
#8	groups [tiab]
#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10	animals [mh] NOT humans [mh]
#11	#9 NOT #10

PubMed search syntax (for Box 3.a above and Box 3.b below):

[pt] denotes a Publication Type term;

[tiab] denotes a word in the title or abstract;

[sh] denotes a subheading;

[mh] denotes a Medical Subject Heading (MeSH) term 'exploded';

[mesh:noexp] denotes a Medical Subject Heading (MeSH) term not 'exploded';

[ti] denotes a word in the title.

Box 3.b Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); PubMed format

#1	randomized controlled trial [pt]
#2	controlled clinical trial [pt]
#3	randomized [tiab]
#4	placebo [tiab]
#5	clinical trials as topic [mesh:noexp]
#6	randomly [tiab]
#7	trial [ti]
#8	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7
#9	animals [mh] NOT humans [mh]
#10	#8 NOT #9

The *search syntax* is explained above under Box 3.a

Box 3.c Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2023 revision); Ovid format

1	exp randomized controlled trial/
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.

5	drug therapy.fs.
6	randomly.ab.
7	trial.ab.
8	groups.ab.
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10	exp animals/ not humans.sh.
11	9 not 10

Ovid search syntax (for Box 3.c above and Box 3.d below):

exp denotes a Medical Subject Heading (MeSH) term 'exploded';

.pt. denotes a Publication Type term;

.ab. denotes a word in the abstract;

.fs. denotes a 'floating' subheading, that is a subheading irrespective of the MeSH term to which it is attached;

.sh. denotes a Medical Subject Heading (MeSH) term not 'exploded';

.ti. denotes a word in the title.

Box 3.d Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2023 revision); Ovid format

1	exp randomized controlled trial/
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	clinical trials as topic.sh.
6	randomly.ab.
7	trial.ti.
8	1 or 2 or 3 or 4 or 5 or 6 or 7

9	exp animals/ not humans.sh.
10	8 not 9

The *search syntax* is explained above under Box 3.c.

3.6.2 Search filters for identifying randomized trials in Embase

As discussed in Section 2.1.2, Embase has been searched with various filters from 1980 to date (and from 1974 to 1979 for some search terms), and records of reports of trials (predominantly based on screening the titles and abstracts only) have been included in CENTRAL. Cochrane funded the development of a highly sensitive search strategy for identifying reports of controlled trials in Embase (Glanville et al 2019a). This search filter was designed for the Embase database via the Ovid interface and was developed, tested and validated in 2016.

After the development of the filter, the Cochrane Centralized Search Service (CSS) decided to move to conducting regular searches for reports of RCTs and CCTs using the Embase.com interface, maintained by Elsevier. This move required a translation of the Ovid Embase RCT filter (Glanville et al 2019a). Variations of this filter have been used over time to identify reports of controlled trials in Embase for inclusion in CENTRAL. For the current version of the Embase.com filter used by Cochrane for identifying trials for CENTRAL, see the 'How CENTRAL is created' file at https://www.cochranelibrary.com/central/central-creation. The filters described above were optimized for identifying reports of trials for CENTRAL and were not optimized for use by individuals searching Embase to identify RCTs. A proposed filter for searchers for identifying trials in Embase.com is shown in Box3e and the Ovid version is shown in Box 3.f below. This filter was updated in April 2023 to reflect the addition of more specific (i.e. narower) Emtree terms to the 'Randomized controlled trial' Emtree heading. The revision has involved exploding that heading in the two lines where it is used in the filter. Alternatively, other search filters for identifying reports of RCTs in Embase can be identified from the InterTASC Information Specialists' Sub-Group's Search Filter Resource (https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/rcts). As noted above, filter selection should be informed by an assessment of the filter's performance.

Box 3.e Cochrane Highly Sensitive Search Strategy used for identifying controlled trials in Embase for CENTRAL: (2023 revision); Embase.com format (Glanville et al 2019a)

1	'randomized controlled trial'/exp
2	'controlled clinical trial'/de
3	random*:ti,ab,tt
4	'randomization'/de
5	'intermethod comparison'/de

6 placebo:ti,ab,tt

7 (compare:ti,tt OR compared:ti,tt OR comparison:ti,tt)

8 ((evaluated:ab OR evaluate:ab OR evaluating:ab OR assessed:ab OR assess:ab) AND (compare:ab OR compared:ab OR comparing:ab OR comparison:ab))

9 (open NEXT/1 label):ti,ab,tt

10 ((double OR single OR doubly OR singly) NEXT/1 (blind OR blinded OR blindly)):ti,ab,tt

11 'double blind procedure'/de

12 (parallel NEXT/1 group*):ti,ab,tt

13 (crossover:ti,ab,tt OR 'cross over':ti,ab,tt)

14 ((assign* OR match OR matched OR allocation) NEAR/6 (alternate OR group OR groups OR intervention OR interventions OR patient OR patients OR subject OR subjects OR participant OR participants)):ti,ab,tt

15 (assigned:ti,ab,tt OR allocated:ti,ab,tt)

16 (controlled NEAR/8 (study OR design OR trial)):ti,ab,tt

17 (volunteer:ti,ab,tt OR volunteers:ti,ab,tt)

18 'human experiment'/de

19 trial:ti,tt

20 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19

21 (((random* NEXT/1 sampl* NEAR/8 ('cross section*' OR questionnaire* OR survey OR surveys OR database or databases)):ti,ab,tt) NOT ('comparative study'/de OR 'controlled study'/de OR 'randomised controlled':ti,ab,tt OR 'randomized controlled':ti,ab,tt OR 'randomly assigned':ti,ab,tt))

22 ('cross-sectional study'/de NOT ('randomized controlled trial'/exp OR 'controlled clinical study'/de OR 'controlled study'/de OR 'randomised controlled':ti,ab,tt OR 'randomized controlled':ti,ab,tt OR 'control group':ti,ab,tt OR 'control groups':ti,ab,tt))

23 ('case control*':ti,ab,tt AND random*:ti,ab,tt NOT ('randomised controlled':ti,ab,tt OR 'randomized controlled':ti,ab,tt))

24 ('systematic review':ti,tt NOT (trial:ti,tt OR study:ti,tt))

25 (nonrandom*:ti,ab,tt NOT random*:ti,ab,tt)

26 'random field*':ti,ab,tt

27 ('random cluster' NEAR/4 sampl*):ti,ab,tt

28 (review:ab AND review:it) NOT trial:ti,tt

29 ('we searched':ab AND (review:ti,tt OR review:it))

30 'update review':ab

31 (databases NEAR/5 searched):ab

32 ((rat:ti,tt OR rats:ti,tt OR mouse:ti,tt OR mice:ti,tt OR swine:ti,tt OR porcine:ti,tt OR murine:ti,tt OR sheep:ti,tt OR lambs:ti,tt OR pigs:ti,tt OR piglets:ti,tt OR rabbit:ti,tt OR rabbits:ti,tt OR cat:ti,tt OR cats:ti,tt OR dog:ti,tt OR dogs:ti,tt OR cattle:ti,tt OR bovine:ti,tt OR monkey:ti,tt OR monkeys:ti,tt OR trout:ti,tt OR marmoset*:ti,tt) AND 'animal experiment'/de)

33 ('animal experiment'/de NOT ('human experiment'/de OR 'human'/de))

34 #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33

35 #20 NOT #34

Embase.com search syntax
/exp denotes an exploded index term (Emtree indexing term)

/de denotes an index term (Emtree indexing term)

:ti denotes a word in the article title

:ab denotes a word in the abstract

:tt denotes a word in the original non-English title

:it denotes a publication type (item type).

The Cochrane Embase RCT filters for Embase.com and Ovid – **version for searchers** – (2023 revision) are also available from https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/rcts/embase-rct-filter.

Box 3.f Cochrane Highly Sensitive Search Strategy used for identifying controlled trials in Embase for CENTRAL: 2023 revision; Ovid format (adapted from (Glanville et al 2019a))

1 exp randomized controlled trial/

2 controlled clinical trial/

3 random$.ti,ab.

4 randomization/

5 intermethod comparison/

6 placebo.ti,ab.

7 (compare OR compared OR comparison).ti,ab.

8 ((evaluated OR evaluate OR evaluating OR assessed OR assess) AND (compare OR compared OR comparing OR comparison)).ab.

9 (open adj label).ti,ab.

10 ((double OR single OR doubly OR singly) adj (blind OR blinded OR blindly)).ti,ab.

11 double blind procedure/

12 parallel group$1.ti,ab.

13 (crossover OR cross over).ti,ab.

14 ((assign$ OR match OR matched OR allocation) adj5 (alternate OR group$1 OR intervention$1 OR patient$1 OR subject$1 OR participant$1)).ti,ab.

15 (assigned OR allocated).ti,ab.

16 (controlled adj7 (study OR design OR trial)).ti,ab.

17 (volunteer OR volunteers).ti,ab.

18 human experiment/

19 trial.ti.

20 or/1-19

21 (random$ adj sampl$ adj7 ("cross section$" OR questionnaire$1 OR survey$ OR database$1)).ti,ab. NOT (comparative study/ OR controlled study/ OR randomi?ed controlled.ti,ab. OR randomly assigned.ti,ab.)

22 cross-sectional study/ NOT (exp randomized controlled trial/ OR controlled clinical trial/ OR controlled study/ OR randomi?ed controlled.ti,ab. OR control group$1.ti,ab.)

23 (((case adj control$) AND random$.ti,ab.) NOT randomi?ed controlled).ti,ab.

24 systematic review.ti,ab. NOT (trial OR study).ti.

25 (nonrandom$ NOT random$).ti,ab.

26 "random field$".ti,ab.

27 (random cluster adj3 sampl$).ti,ab.

28 (review.ab. AND review.pt.) NOT trial.ti.

29 "we searched".ab. AND (review.ti. OR review.pt.)

30 "update review".ab.

31 (databases adj4 searched).ab.

32 (rat OR rats OR mouse OR mice OR swine OR porcine OR murine OR sheep OR lambs OR pigs OR piglets OR rabbit OR rabbits OR cat OR cats OR dog OR dogs OR cattle OR bovine OR monkey OR monkeys OR trout OR marmoset$1).ti. AND animal experiment/

33 animal experiment/ NOT (human experiment/ OR human/)

34 or/21-33

35 20 NOT 34

Embase on Ovid search syntax
exp denotes an Emtree term 'exploded'

/ denotes an Emtree term not 'exploded'

.ti,ab. denotes a word in the title or abstract

.ab. denotes a word in the abstract

.ti. denotes a word in the title

.pt. denotes a Publication Type term.

The Cochrane Embase RCT filters for Embase.com and Ovid – versions for searchers – (2023 revision) are also available from https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/rcts/embase-rct-filter.

3.6.3 Search filters for identifying randomized trials in CINAHL versions without full text

A search filter for identifying randomized trials in the then CINAHL Plus was prepared by the Cochrane Centralized Search Service (CSS) and was published in February 2019 (Glanville et al 2019b). Note that this search filter is optimized for CINAHL without the full text publications and searchers with access to versions with full text publications should consider additional search approaches to make use of the option to search the full text (see Section 3.6.1 above). The strategy was modified in May 2023 to explode MH randomized controlled trials to reflect NLM's changes to that heading which were also adopted in CINAHL.

Box 3.g Cochrane Highly Sensitive Search Strategy for identifying controlled trials in CINAHL EBSCO versions without full text (May 2023 version)

S1 MH randomized controlled trials+

S2 MH double-blind studies

S3 MH single-blind studies

S4 MH random assignment

S5 MH pretest-posttest design

S6 MH cluster sample

S7 TI (randomised OR randomized)

S8 AB (random*)

S9 TI (trial)

S10 MH (sample size) AND AB (assigned OR allocated OR control)

S11 MH (placebos)

S12 PT (randomized controlled trial)

S13 AB (control W5 group)

S14 MH (crossover design) OR MH (comparative studies)

S15 AB (cluster W3 RCT)

S16 MH animals+

S17 MH (animal studies)

S18 TI (animal model*)

S19 S16 OR S17 OR S18

S20 MH (human)

S21 S19 NOT S20

S22 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15

S23	S22 NOT S21

CINAHL on EBSCO search syntax

MH	CINAHL subject heading
+	explode subject heading
AB	Word in abstract
TI	Word in title
*	Truncated word
WN	Within N words

3.7 Demonstration search strategies

Box3h provides a demonstration search strategy for CENTRAL for the topic 'treating breast cancer with tamoxifen'. Note that it includes topic terms only and there is no limit to humans or trials, because CENTRAL only contains reports of randomized and controlled trials in humans. The strategy is provided for illustrative purposes only: searches of CENTRAL for studies to include in a systematic review would have many more search terms for each of the concepts.

Box3i provides a demonstration search strategy for MEDLINE (Ovid format) for the topic 'treating breast cancer with tamoxifen'. Note that both topic terms and a randomized trial filter are used for MEDLINE. The search is limited to studies that are likely to be about humans, by excluding studies that are only about animals. The strategy is provided for illustrative purposes only: searches of MEDLINE for studies to include in a systematic review would have many more search terms for each of the concepts.

Box 3.h Demonstration search strategy for CENTRAL, for the topic 'treating breast cancer with tamoxifen'

#1	[mh "breast neoplasms"]
#2	(breast NEAR cancer*):ti,ab,kw
#3	(breast NEAR neoplasm*):ti,ab,kw
#4	(breast NEAR carcinoma*):ti,ab,kw
#5	(breast NEAR tumour*):ti,ab,kw
#6	(breast NEAR tumor*):ti,ab,kw
#7	#1 OR #2 OR #3 OR #4 OR #5 OR #6

#8	[mh tamoxifen]
#9	tamoxifen:ti,ab,kw
#10	#8 OR #9
#11	#7 AND #10

Cochrane CENTRAL search syntax for the Cochrane Library via the Wiley interface

mh denotes a MeSH term

The 'NEAR' operator defaults to within six words

'*' indicates truncation

:ti,ab,kw denotes a word or phrase in the title, abstract or keyword field.

Box 3.i Demonstration search strategy for MEDLINE (Ovid format), for the topic 'treating breast cancer with tamoxifen'

1	exp randomized controlled trial/
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	drug therapy.fs.
6	randomly.ab.
7	trial.ab.
8	groups.ab.
9	or/1-8
10	exp animals/ not humans/
11	9 not 10
12	exp Breast Neoplasms/
13	(breast adj6 cancer$).ti,ab,kf.

14	(breast adj6 neoplasm$).ti,ab,kf.
15	(breast adj6 carcinoma$).ti,ab,kf.
16	(breast adj6 tumour$).ti,ab,kf.
17	(breast adj6 tumor$).ti,ab,kf.
18	or/12-17
19	exp Tamoxifen/
20	tamoxifen.ti,ab,kf.
21	19 or 20
22	11 and 18 and 21

Ovid search syntax

exp denotes a Medical Subject Heading (MeSH) term 'exploded'

.pt. denotes a Publication Type term

.ab. denotes a word in the abstract

.fs. denotes a 'floating' subheading, that is a subheading irrespective of the MeSH term to which it is attached

'adj6' operator indicates within six words

'$' indicates truncation

.ti,ab,kf. denotes a word in the title, abstract or Keyword Heading Word [Word Indexed].

3.8 Adapting search strategies across databases/sources and interfaces

Search strategies need to be customized for each database and search interface. Special caution is warranted when adapting a search strategy developed for a specific database in a specific interface to other databases and/or interfaces. This process requires a thorough knowledge of the specifications of both the new database and the new interface, including the controlled vocabulary being used to index the database's content and the availability of Boolean and proximity operators, as well as the specific syntax for wildcards and truncation and definitions of date fields. These vary across databases and interfaces and need to be taken into account before running a strategy. Searchers should be particularly vigilant with respect to wildcard and truncation symbols, which in some cases have the opposite meaning in different database interfaces. Additionally, a search for health economics in a general

healthcare database such as MEDLINE will require different natural language (free-text) terminology/search terms from the terminology required in a specialized economics database. Review authors are, therefore, encouraged to work together with a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, who can provide advice on the accuracy of adaptations carried out by the review authors themselves or may be able to provide adaptations of the principal, generally MEDLINE, search strategy into the databases and trials registers, which will be searched for the review. Some attempts have been made to simplify through automation the adaptation of search syntax across service providers:

- Bond University Centre for Research in Evidence-Based Practice "Systematic Review Accelerator Polyglot Search" http://sr-accelerator.com/#/polyglot

- Erasmus University Medical Centre (Bramer et al 2017b) http://www.stationsweb.nl/emcmb_cursus/bestanden/macros.html

- MEDLINE Transpose from the College of Physicians and Surgeons of British Columbia (CPSBC) and the Collaboration for Leadership in Applied Health Research and Care South West Peninsula (PenCLAHRC) https://medlinetranspose.github.io/about.html (Wanner and Baumann 2018).

None of the above, however, addresses the complexities outlined above regarding differences in natural language (free-text) terminology or controlled vocabulary.

With respect to date fields, the table below indicates the equivalent date fields between Ovid and PubMed. For example, it is important to note that the Publication Date (DP) field in PubMed (for the date that the article was published) is not equivalent to the Year of Publication (YR) field in Ovid MEDLINE – see Table 3.8.a.

For further information on MEDLINE/PubMed date field descriptions see:

https://pubmed.ncbi.nlm.nih.gov/help/#search-tags

https://www.nlm.nih.gov/bsd/mms/medlineelements.html

For further information on how date fields in Ovid MEDLINE correlate with those in PubMed see:

https://tools.ovid.com/ovidtools/pdf/Ovid_MEDLINE_and_PubMed_compared.pdf

Table 3.8.a Equivalent date fields between Ovid MEDLINE and PubMed

PubMed Search	Ovid MEDLINE Search
("1950"[Electronic Publication Date] : "2023"[Electronic Publication Date])	EP - Electronic Date of Pub.: 19500101:20231231.(ep).
("1950"[Date - Publication] : "2023"[Date - Publication])	Note: Use both YR and EP: 1950:2023.(yr). or 19500101:20221231.(ep).
("1950"[Date - MeSH] : "2023"[Date - MeSH])	DA - MeSH date: 19500101:20231231.(da).
("1950"[Date - Entrez] : "2023"[Date - Entrez])	EZ - Entrez date: 19500101:20231231.(ez).
("1950"[Date - Create] : "2023"[Date - Create])	DT - Create date: 19500101:20231231.(dt).
("1950"[Date - Completion] : "2023"[Date - Completion])	ED - Entry date: 19500101:20231231.(ed).

3.9 Identifying fraudulent studies, other retracted publications, errata and comments: further considerations

This section should be read in conjunction with Chapter 4, Section 4.4.6. It is mandatory, for authors of Cochrane Reviews of interventions, to examine any relevant retraction statements and errata for information (MECIR C48). Identifying retraction statements and published errata or comments (and their associated original retracted articles or corrected articles) can help to avoid errors that impact on the overall estimates in systematic reviews. It is essential at the original search stage to ascertain whether any retractions or errata have been published for studies to be included in the original review and also at the update stage to ascertain whether any retractions or errata have been published subsequently for studies previously included in the original review.

Reports of studies indexed in MEDLINE that have been retracted (as fraudulent or for other reasons) will have the Publication Type term 'Retracted Publication' added to the record (since 1989). The article giving notice of the retraction (the retraction notice) will have the Publication Type term 'Retraction of Publication' assigned (since 1991).

How to search for retraction notices and retracted publications in Ovid MEDLINE:

- retracted publication.pt. or retraction of publication.pt.

How to search for retraction notices and retracted publications in PubMed:

- retracted publication [pt] OR retraction of publication [pt]

The above searches should be supplemented with a free-text search of both of the terms 'retracted' and 'retraction' limited to the title fields, to pick up records not (yet) indexed as such but this will inevitably result in false positives, i.e. irrelevant records.

Retraction notices indexed in Embase until April 2017 were identified by the Publication Type 'erratum' and were additionally indexed with the Preferred Term 'retracted article'. There was no link, prior to April 2017, back from the retraction notice to the original retracted article, as there is in MEDLINE.

How to search for retraction notices and retracted publications in Ovid Embase:

- Erratum.pt. or Retracted article/ or Tombstone.pt. or yes.ne.

As above for MEDLINE, the above search in Embase should be supplemented with a free-text search of both of the terms 'retracted' and 'retraction' limited to the title fields, to pick up records not (yet) indexed as such but this will inevitably result in false positives, i.e. irrelevant records.

Prior to any decision being taken to retract an article, articles may be published that refer to an original article and raise concerns of this sort. A new MeSH Publication Type was introduced in 2018 to cover this: Expression of Concern. This is defined in the Scope Note as: "A notification about the integrity of a published article that is typically written by an editor and should be labelled prominently in the item title. It is the responsibility of the editor to initiate appropriate investigative procedures, discover the outcome of the investigation, and notify readers of that outcome in a subsequent published item. The outcome may require the publication of a retraction notice."

To search for "expressions of concern" prior to 2018, search for the phrase "expression of concern".

Search in Ovid, across all dates, as:

expression of concern.pt. or "expression of concern".af.

Search in PubMed, across all dates, as:

"expression of concern"[Publication Type] OR "expression of concern"[All Fields]

As noted above, MEDLINE/PubMed reports of randomized trials that have been retracted and indexed as such in the MEDLINE, will include the 'Retracted Publication' term in the Publication Type field (since 1989). This is also the case for those retracted articles in CENTRAL which are sourced from MEDLINE. This is not, however, the case for the majority of records from Embase (prior to 2017) or from other sources.

In addition, articles may have been partially retracted (previously indexed in MEDLINE as Partial Retraction but since 2016 indexed as Erratum), corrected through a published erratum or may have been corrected and re-published in full. It is therefore important to search MEDLINE for the latest version of the citations to the records for the (previously) included studies when updating a review. In some display formats of some versions of MEDLINE the retracted publication, erratum and comment statements are included in the citation data together with the title and are, therefore, highly visible. This is not, however, always the case so care should be taken to ensure that this information is always retrieved in all searches by downloading the appropriate fields together with the citation data.

Retraction Watch is a resource listing retracted publications (since late 2010). Review authors and others interested in keeping abreast of this area can subscribe to their blog by email. They can also search the Retraction Watch blog and archives by category (http://retractionwatch.com/) or search the Retraction Watch Database (http://retractiondatabase.org/), which has over 50,000 entries (as of June 2024). A user guide is available at https://retractionwatch.com/retraction-watch-database-user-guide/. Some reference management and related programmes, such as Edifix, EndNote, LibKey, Papers, and Zotero, link with the Retraction Watch Database to notify the user automatically when a reference to a study matches a retraction in the Retraction Watch Database. Some other services such as Crossref, the Lens and OpenAlex include Retraction Watch data within their data. Further information on this functionality is available from the various reference management software providers.

3.10 Summary points

- Cochrane Review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, on designing search strategies.
- Authors of non-Cochrane reviews should seek advice from a medical/healthcare librarian or information specialist, with experience of conducting searches for studies for systematic reviews.
- Avoid too many *different* search concepts but use a wide variety of synonyms and related terms *within* each concept.
- Appropriate controlled vocabulary (e.g. MeSH, Emtree, including 'exploded' terms) and free-text terms should be identified (considering, for example, spelling variants, synonyms, acronyms, truncation and proximity operators).
- Ensure correct use of the 'AND' and 'OR' operators.
- Avoid use of the 'NOT' operator in combining search sets.

- Specially designed and tested search filters should be used where appropriate including the Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE, Embase and CINAHL.
- Do not use filters in pre-filtered databases e.g. do not use a randomized trial or human studies filter in CENTRAL or a systematic review filter in a database consisting solely of systematic reviews.
- For identifying randomized trials in MEDLINE, begin with a highly sensitive search filter such as the sensitivity-maximizing version of the Cochrane Highly Sensitive Search Strategy. If this retrieves an unmanageable number of references, use the sensitivity- and precision-maximizing version instead. (See Sections 2.1.1 and 2.2.2 for details as to how these search strategies have already been run centrally in Cochrane over the years and relevant records included in CENTRAL, to avoid unnecessary duplication of effort.)
- Searches designed for a specific database and service provider will need to be adapted for use in another database or service provider.
- Ensure awareness of any retracted publications (e.g. fraudulent publications), errata and comments.
- Consideration should be given to searching indexed records and non-indexed/in-process records separately in databases such as MEDLINE and Embase which include both indexed and non-indexed content.

4 Managing references

4.1 Reference management software

Reference management software is used to import, de-duplicate, and store references from database searches and searches of other sources. Specially designed bibliographic or reference management software such as EndNote (https://endnote.com/), Mendeley (https://www.mendeley.com/), RefWorks (https://about.proquest.com/en/products-services/refworks) and Zotero (https://www.zotero.org/) is useful and relatively easy to use to keep track of references to and other records of studies (Lorenzetti and Ghali 2013). Reference management software varies in terms of cost, operating system, and ease of database and record sharing, among other characteristics. The choice of which software to use is likely to be influenced by what is available and thus supported at the review author's institution. In May 2024, there were approximately 40 different software tools listed in the reference management section of the Systematic Review Toolbox at: http://systematicreviewtools.com/. This service is currently unavailable (September 2024) whilst the developers seek and transition to a new hosting solution. For a comparison of several of the main products see https://en.wikipedia.org/wiki/Comparison_of_reference_management_software.

Reference management software usually provides import file formats (import filters) that allow text files exported from sources such as APA PsycInfo, CENTRAL, CINAHL, Embase, MEDLINE, PubMed and others to be imported into the reference management database. Some reference management software can also be used to search sources such as PubMed from within the database of citations and to import retrieved records directly from those sources. Using reference management software to carry out complex searches, such as those for identifying studies for systematic reviews, is, however, discouraged (Gomis et al 2008).

Reference management software facilitates storage of information about the methods and process of a search. For example, unused record fields can be used to store information such as 1) the name of the database or other source details from which a trial record was identified, 2) when and from where a document was ordered and the date of document receipt, 3) when and with whom the search results were shared, and 4) whether the study associated with a record/document was included in or excluded from a review and, if excluded, the reasons for exclusion.

Software is increasingly being developed to manage a range of functions within the systematic review process and many of these products also have some level of reference management capacity. See Chapter 4, Section 4.6.6.1 of the *Handbook* for a discussion of software that can be used to assist in record screening and selection. An evaluation of several software tools has been published recently (Harrison et al 2020). Up-to-date information about these and other software tools will be available from the Systematic Review Toolbox at http://systematicreviewtools.com/, when a new hosting solution for the website is identified as described above.

4.2 Which fields to download

In addition to the fields that are essential for identifying a reference (e.g. author, title, source, year) several additional key fields should be considered for downloading from databases where they are available. Some of these key fields are listed below. The list below is intended, where possible, to be generic across databases. For the full range of fields in PubMed, see:

https://www.nlm.nih.gov/bsd/mms/medlineelements.html.

Abstract: abstracts can be used to eliminate clearly irrelevant reports, obviating the need to obtain the full text of those reports or to return to the bibliographic database at a later time.

Accession number/unique identifier: it is advisable to allocate an unused field or fields to store the unique identifier(s)/accession number(s) of records downloaded, such as the PubMed ID number (PMID). This allows subsequent linkage to the full database record and also facilitates information management such as duplicate detection and removal (i.e. de-duplication).

Affiliation/address: may include the institutional affiliation and/or email address of the author/investigator.

Article identifier/digital object identifier (DOI): can be used to cite and link to the full record. In PubMed, the Article Identifier (AID) field includes article identifiers submitted by journal publishers such as DOI (digital object identifier). In Ovid MEDLINE, the Digital Object Identifier (DO) field contains the DOI.

Author identifier: can be used to disambiguate authors with similar names. The identifier may be an ORCID (https://info.orcid.org/what-is-orcid), an International Standard Name Identifier (ISNI) https://isni.org/, or from the Virtual International Authority File (VIAF) https://viaf.org/.

Clinical trial number: if the record contains a clinical trial number, such as the registration number assigned by ClinicalTrials.gov or the ISRCTN schemes, or a number allocated by the sponsor of the trial, these should be downloaded to aid linking of trial reports to the original studies. In PubMed, the Secondary Source ID field [SI] contains information from secondary sources such as ClinicalTrials.gov and ISRCTN. Similarly, in Ovid MEDLINE, the Secondary Source Linking (SL) field contains the URL to ClinicalTrials.gov and ISRCTN resources where these are mentioned in MEDLINE records. In Embase, the Clinical Trial Number (CN) field contains clinical trial numbers associated with the record.

Index terms/thesaurus terms/keywords: These help indicate why records were retrieved if the title and abstract lack detail.

Investigator name: this field contains personal names of individuals (e.g. collaborators and investigators) who are not authors of the article but rather are listed in the article as members of a collective/corporate group that is an author of the article.

Language: this is the language (or languages) of publication of the original document.

Location identifier: in PubMed, this field includes the DOI or publisher ID that serves the role of pagination to locate an online article.

Original title/Title: if the original title of the document is not in English and both the original title and the translated title are available, then both titles should be downloaded into separate database fields, to aid correct identification of the reference and de-duplication. See also Transliterated title below with respect to PubMed.

Other term: in PubMed this field contains largely non-MeSH subject terms (also referred to as Keywords) that describe the content of the article. Author-supplied keywords are included here in PubMed (since 2013). (Note: this field is searched automatically in PubMed when the field TIAB, or the field KW is searched).

Registry Number/EC Number and Substance Name: these fields provide supplementary subject information regarding substances (chemicals, drugs and enzymes).

Transliterated title: in PubMed, this field contains the title of each item originally published in a non-English language, in that language. Transliterations of article titles in some Cyrillic languages (Greek, Bulgarian, Russian, Serbian and Ukrainian) were added to this field until 2004. This field can be useful for de-duplication.

Comments, corrections, errata, retractions and updates:

It is mandatory, for Cochrane Reviews of interventions, to examine any relevant retraction statements and errata for information (MECIR C48). All fields that relate to subsequently published comments, corrections, errata, retractions and updates should be selected for inclusion in the download, so that any impact of these subsequent publications can be taken into account. The MECIR standard specifies: "Care should be taken to ensure that this information is retrieved in all database searches by downloading the appropriate fields, together with the citation data". For example, the most important fields to consider, in relation to comments, errata, retractions, etc., together with their field labels in PubMed, are provided in
Box 4.a.

Box 4.a Important field labels in PubMed in relation to comments, errata, retractions, etc.

CIN: 'Comment in'
CON: 'Comment on'
CRI: 'Corrected and Republished in'
CRF: 'Corrected and Republished from'
EIN: 'Erratum in'
EFR: 'Erratum for'
ECI: Expression of Concern in
ECF: Expression of Concern for
RIN: 'Retraction in'
ROF: 'Retraction of'
RRI: 'Retracted and Republished in'
RRF: 'Retracted and Republished from'

RPI: ‘Republished in’

RPF: ‘Republished from’

UIN: ‘Update in’

UOF: ‘Update of’

See:

https://www.nlm.nih.gov/bsd/mms/medlineelements.html#cc

https://www.nlm.nih.gov/bsd/policy/errata.html

https://pubmed.ncbi.nlm.nih.gov/help/#comment-correction

The above list is provided as an example of the relevant fields in PubMed and as an indicator of the equivalent fields in other databases and service providers.

4.3 De-duplicating references

Because searching to inform systematic reviews is intended to be extensive, thousands of records may be retrieved from multiple sources. References to the same article (that is, exactly the same bibliographic reference) may be downloaded multiple times from different sources and duplicates can even be found within individual databases. The failure to remove duplicate records leads to extra time and effort during the screening phase and may lead to difficulty in specifying the total number of non-duplicate records in the PRISMA flow diagram (Page et al 2021b, Page et al 2021a). Failure to identify duplicate records may even lead to mistakenly including duplicate data in systematic reviews (Tramèr et al 1997). On the other hand, deleting non-duplicate records from search retrievals in error may lead to omitting relevant studies from systematic reviews. This is a particular concern in using simple filters for reducing retrieval of duplicate records during searching. For example, recent testing of the Exclude MEDLINE Journals filter in Embase suggests that this approach may result in loss of unique records (Premji and Ganshorn 2020). Many Cochrane Information Specialists de-duplicate records so that review authors see only search results that have already been de-duplicated.

Formatting of citation information often varies across sources, and automated identification of duplicate references from within reference management software may lead to false positives (removing non-duplicate records) and false negatives (retaining duplicate records). Meanwhile, de-duplication through visual examination of each record is time-consuming and often impractical. Several strategies have been developed to address these issues. Methods for modifying duplicate detection algorithms within reference management software have been developed and tested (Kwon et al 2015, Bramer et al 2016b). An online method to identify search results that are duplicates of PubMed citations has been reported (Sampson et al 2006).

Comparisons of false positives, false negatives, and time involved in use of different searching methods and reference software for de-duplication have been tested, with no clear advantage to any one method (Kwon et al 2015). Both commercial deduplication programmes (Borissov et al 2022) and open-source software programs for online duplicate detection have also been developed (Wallace et al 2012, Jiang et al 2014, Rathbone et al 2015b, Rathbone et al 2015a, Hair et al 2021, Escaldelai et al 2022). Records may also be exported from reference management software into systematic review production software that provides automated duplicate detection (e.g. Covidence, Rayyan) and deduplicated there (Guimarães et al 2022). A recent comparison of several de-duplication approaches including the Ovid multifile search, reference management software, and systematic review production software has revealed that different methods have different combinations of false positives and false negatives, with systematic review production software being among the most accurate methods (McKeown and Mir 2021). When practical, manual methods may still be best, depending on the available alternatives. For example, a recent case study comparing manual and automated tools for completion of systematic review tasks found that manual duplicate detection performed better than automated detection (Clark et al 2021). There is no consensus on the optimal method for duplicate detection, and the most appropriate method will most likely depend upon the size of the combined dataset, the identity, number and output format of the resources searched, access to software tools, and the skill and comfort level of the operator. A combination of automated methods and visual inspection is often used.

After de-duplication of search results, records may be screened for inclusion from within the reference management database, exported into dedicated screening software, or imported into systematic review production software (where further deduplication may take place). See Chapter 4, Section 4.6.6.1 for discussion of software to support the screening process. Records for included and excluded studies can be exported and uploaded into systematic review software such as RevMan. Instructions for importing references from reference management software into RevMan 5 can be found at https://community.cochrane.org/help/tools-and-software/revman-5/support-revman-5/revman-5-faq and directions for importing references into RevMan Web can be found at https://documentation.cochrane.org/revman-kb/import-references-for-included-and-excluded-studies-95422738.html.

4.4 Summary points

- Cochrane Review authors should seek advice from a Cochrane Information Specialist, or other medical/healthcare librarian or information specialist, on managing references.
- Authors of non-Cochrane reviews should seek advice from a medical/healthcare librarian or information specialist, with experience of managing references for systematic reviews.
- Use of reference management software is recommended.
- Ensure that all the necessary fields are downloaded.

- Remove duplicate references before screening.
- Either screen references within the reference management software and export references for the included and excluded studies into systematic review software, or export references to specialized screening software.

5 Supplement information

Authors: Carol Lefebvre, Julie Glanville, Simon Briscoe, Robin Featherstone, Anne Littlewood, Maria-Inti Metzendorf, Anna Noel-Storr, Robin Paynter, Tamara Rader, James Thomas and L. Susan Wieland on behalf of the Cochrane Information Retrieval Methods Group

6 Acknowledgements

This chapter has been developed from sections of previous editions of this Handbook since 1995. Authors of previous editions included Kay Dickersin, Julie Glanville, Kristen Larson, Carol Lefebvre, Eric Manheimer, Chris Marshall and Farhad Shokraneh. Many of the sources listed in this Technical Supplement, the accompanying Appendix and Chapter 4 of the Cochrane Handbook for Systematic Reviews of Interventions have been brought to our attention by a variety of people over the years and we should like to acknowledge this. We should like to acknowledge: Ruth Foxlee, (formerly) Information Specialist, Cochrane Editorial Unit; Miranda Cumpston, (formerly) Head of Learning & Support, Cochrane Central Executive; Colleen Finley, Product Manager, John Wiley and Sons, for checking sections relating to searching the Cochrane Library; the (UK) National Institute for Health and Care Excellence, the German Institute for Quality and Efficiency in Health Care (IQWiG) and the authors of a living systematic review on the problems with systematic reviews (Uttley et al 2023) for support in identifying some of the references; the (US) Agency for Healthcare Research and Quality (AHRQ) Effective Healthcare Program Scientific Resource Center for their previous Article Alert service; Tianjing Li, Co-Convenor, Comparing Multiple Interventions Methods Group, for text and references that formed the basis of the re-drafting of parts of Section 4.6 Selecting studies; Lesley Gillespie, Cochrane author and former Editor and Trials Search Co-ordinator of the Cochrane Bone, Joint and Muscle Trauma Group, for copy-editing an early draft; the Cochrane Information Specialists' Executive, the previous Cochrane Information Specialists' Support Team, Cochrane Information Specialists, members of the Cochrane Information Retrieval Methods Group and others for comments on drafts; Su Golder, Co-Convenor, Adverse Effects Methods Group and Steve McDonald, Co-Director, Cochrane Australia for peer review of earlier versions.

7 References

Abhijnhan A, Surcheva Z, Wright J, Adams CE. Searching a biomedical bibliographic database from Bulgaria: the ABS database. *Health Information and Libraries Journal* 2007; **24**: 200-203.

Adam GP, Springs S, Trikalinos T, Williams JW, Jr., Eaton JL, Von Isenburg M, Gierisch JM, Wilson LM, Robinson KA, Viswanathan M, Middleton JC, Forman-Hoffman VL, Berliner E, Kaplan RM. Does information from ClinicalTrials.gov increase transparency and reduce bias? Results from a five-report case series. *Syst Rev* 2018; **7**: 59.

Adam GP, Wallace BC, Trikalinos TA. Semi-automated Tools for Systematic Searches. *Methods Mol Biol* 2022; **2345**: 17-40.

Adam GP, Paynter R. Development of literature search strategies for evidence syntheses: pros and cons of incorporating text mining tools and objective approaches. *BMJ Evid Based Med* 2023; **28**: 137-139.

Al-Hajeri AA, Fedorowicz Z, Amin FA, Eisinga A. The handsearching of 2 medical journals of Bahrain for reports of randomized controlled trials. *Saudi Medical Journal* 2006; **27**: 526-530.

Allot A, Chen Q, Kim S, Vera Alvarez R, Comeau DC, Wilbur WJ, Lu Z. LitSense: making sense of biomedical literature at sentence level. *Nucleic Acids Research* 2019; **47**: W594-W599.

Almerie MQ, Matar H, Jones V. Searching the Polish Medical Bibliography (Polska Bibliografia Lekarska) for trials. *Health Information and Libraries Journal* 2007; **24**: 283-286.

Alqaidoom Z, Nguyen PY, Awadh M, Page MJ. Impact of searching clinical trials registers in systematic reviews of pharmaceutical and non-pharmaceutical interventions: Reanalysis of meta-analyses. *Research Synthesis Methods* 2023; **14**: 52-67.

Anderson ML, Chiswell K, Peterson ED, Tasneem A, Topping J, Califf RM. Compliance with results reporting at ClinicalTrials.gov. *N Engl J Med* 2015; **372**: 1031-1039.

Arruda H, Silva ER, Lessa M, Proença D, Jr., Bartholo R. VOSviewer and Bibliometrix. *J Med Libr Assoc* 2022; **110**: 392-395.

Atsawawaranunt K, Adams CE, Roberts S. Searching for randomised controlled trials and clinical controlled trials in Thai online bibliographical biomedical databases. *Health Information and Libraries Journal* 2011; **28**: 68-76.

Bagg MK, O'Hagan E, Zahara P, Wand BM, Hübscher M, Moseley GL, McAuley JH. Systematic reviews that include only published data may overestimate the effectiveness of analgesic medicines for low back pain: a systematic review and meta-analysis. *J Clin Epidemiol* 2020; **124**: 149-159.

Balan PF, Gerits A, Vanduffel W. A practical application of text mining to literature on cognitive rehabilitation and enhancement through neurostimulation. *Front Syst Neurosci* 2014; **8**: 182.

Banno M, Tsujimoto Y, Kataoka Y. Using the Cochrane Central Register of Controlled Trials to identify clinical trial registration is insufficient: a cross-sectional study. *BMC Med Res Methodol* 2020; **20**: 200.

Barnabas J, Yamuna B, Parthasarathy V, Venkatesh S, Tharyan P. Access to evidence from countries in South Asia: the South Asian Database of Controlled Clinical Trials and the South Asian Cochrane Network and Centre's Digital Library. 17th Cochrane Colloquium; 2009; Singapore. https://abstracts.cochrane.org/2009-singapore/access-evidence-countries-south-asia-south-asian-database-controlled-clinical-trials.

Baudard M, Yavchitz A, Ravaud P, Perrodeau E, Boutron I. Impact of searching clinical trial registries in systematic reviews of pharmaceutical treatments: methodological systematic review and reanalysis of meta-analyses. *BMJ* 2017; **356**: j448.

Becker JE, Krumholz HM, Ben-Josef G, Ross JS. Reporting of results in ClinicalTrials.gov and high-impact journals. *JAMA* 2014; **311**: 1063-1065.

Beckles Z, Glover S, Ashe J, Stockton S, Boynton J, Lai R, Alderson P. Searching CINAHL did not add value to clinical questions posed in NICE guidelines. *J Clin Epidemiol* 2013; **66**: 1051-1057.

Bethel A, Rogers M. A checklist to assess database-hosting platforms for designing and running searches for systematic reviews. *Health Information and Libraries Journal* 2014; **31**: 43-53.

Bhattacharyya M, Miller VM, Bhattacharyya D, Miller LE. High Rates of Fabricated and Inaccurate References in ChatGPT-Generated Medical Content. *Cureus* 2023; **15**: e39238.

Blumle A, Antes G. Ten years of handsearching in Germany: results and future prospects [poster]. 13th Cochrane Colloquium; 2005; Melbourne, Australia. http://abstracts.cochrane.org/2005-melbourne/ten-years-handsearching-germany-results-and-future-prospects.

Boeker M, Vach W, Motschall E. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC Med Res Methodol* 2013a; **13**: 131.

Boeker M, Vach W, Motschall E. Time-dependent migration of citations through PubMed and OvidSP subsets: a study on a series of simultaneous PubMed and OvidSP searches. *Studies in health technology and informatics* 2013b; **192**: 1196.

Boesen K, Jørgensen KJ, Gøtzsche PC. Clinical trials were missing from regulatory documents of extended-release methylphenidate for ADHD in adults: a case study of public documents. *J Clin Epidemiol* 2022; **143**: 242-253.

Boluyt N, Tjosvold L, Lefebvre C, Klassen TP, Offringa M. The usefulness of systematic review search strategies in finding child health systematic reviews in MEDLINE. *Archives of Pediatrics & Adolescent Medicine* 2008; **162**: 111-116.

Bonfill X, Osorio D, Posso M, Sola I, Rada G, Torres A, Garcia Dieguez M, Pina-Pozas M, Diaz-Garcia L, Tristan M, Gandarilla O, Rincon-Valenzuela DA, Marti A, Hidalgo R, Simancas-Racines D, Lopez L, Correa R, Rojas-De-Arias A, Loza C, Gianneo O, Pardo H, Iberoamerican Cochrane Network. Identification of biomedical journals in Spain and Latin America. *Health Information and Libraries Journal* 2015; **32**: 276-286.

Booth A. Cochrane or cock-eyed? How should we conduct systematic reviews of qualitative research? Qualitative Evidence-based Practice Conference, Taking a Critical Stance; 2001; Coventry. http://www.leeds.ac.uk/educol/documents/00001724.htm.

Booth A, Briscoe S, Wright JM. The "realist search": A systematic scoping review of current practice and reporting. *Res Synth Methods* 2020; **11**: 14-35.

Borissov N, Haas Q, Minder B, Kopp-Heim D, von Gernler M, Janka H, Teodoro D, Amini P. Reducing systematic review burden using Deduklick: a novel, automated, reliable, and explainable deduplication algorithm to foster medical research. *Syst Rev* 2022; **11**: 172.

Boynton J, Glanville J, McDaid D, Lefebvre C. Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. *Journal of information science* 1998; **24**: 137-157.

Bramer WM, Giustini D, Kramer BMR, Anderson PF. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. *Systematic reviews* 2013; **2**: 115.

Bramer WM. Variation in number of hits for complex searches in Google Scholar. *Journal of the Medical Library Association* 2016; **104**: 143-145.

Bramer WM, Giustini D, Kramer BM. Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. *Syst Rev* 2016a; **5**: 39.

Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc* 2016b; **104**: 240-243.

Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev* 2017a; **6**: 245.

Bramer WM, Rethlefsen ML, Mast F, Kleijnen J. Evaluation of a new method for librarian-mediated literature searches for systematic reviews. *Res Synth Methods* 2017b; **9**: 510-520.

Brand-de Heer DL. A comparison of the coverage of clinical medicine provided by PASCAL BIOMED and MEDLINE. *Health Information and Libraries Journal* 2001; **18**: 110-116.

Brassey JR. Turning Research Into Practice (TRIP). *Journal of the Medical Library Association* 2007; **95**: 215-216.

Briscoe S, Cooper C. The British Nursing Index and CINAHL: a comparison of journal title coverage and the implications for information professionals. *Health Info Libr J* 2014; **31**: 195-203.

Briscoe S. A review of the reporting of web searching to identify studies for Cochrane systematic reviews. *Research Synthesis Methods* 2018; **9**: 89-99.

Briscoe S, Bethel A, Rogers M. Conduct and reporting of citation searching in Cochrane systematic reviews: A cross-sectional study. *Research Synthesis Methods* 2020a; **11**: 169-180.

Briscoe S, Nunns M, Shaw L. How do Cochrane authors conduct web searching to identify studies? Findings from a cross-sectional sample of Cochrane Reviews. *Health Information and Libraries Journal* 2020b; **37**: 293-318.

Briscoe S, Rogers M. An alternative screening approach for Google Search identifies an accurate and manageable number of results for a systematic review (case study). *Health Info Libr J* 2021; **41**: 149-155.

Briscoe S, Abbott R, Lawal H, Shaw L, Coon JT. Feasibility and desirability of screening search results from Google Search exhaustively for systematic reviews: A cross-case analysis. *Res Synth Methods* 2023; **14**: 427-437.

Burns CS, Shapiro RMn, Nix T, Huber JT. Search results outliers among MEDLINE platforms. *J Med Libr Assoc* 2019; **107**: 364-373.

Burns CS, Nix T, Shapiro RMn, Huber JT. MEDLINE search retrieval issues: A longitudinal query analysis of five vendor platforms. *PLoS One* 2021; **16**: e0234221.

Cawley MA. Using Machine Learning to Locate Evidence More Efficiently: New Roles for Academic Research Librarians. In: Mani NS, Cawley MA, editors. *Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems*. Hershey, PA, USA: IGI Global; 2022. p. 144-168. https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-9702-6.ch008.

Chen KY, Borglund EM, Postema EC, Dunn AG, Bourgeois FT. Reporting of clinical trial safety results in ClinicalTrials.gov for FDA-approved drugs: A cross-sectional analysis. *Clin Trials* 2022; **19**: 442-451.

Chibuzor M, Meremikwu M. Preliminary assessment of handsearching programmes for randomized controlled trials in Nigeria [Poster]. 17th Cochrane Colloquium; 2009; Singapore. https://abstracts.cochrane.org/2009-singapore/preliminary-assessment-handsearching-program-mes-randomized-controlled-trials.

Chokkalingam A, Scherer R, Dickersin K. Agreement of data in abstracts compared to full publications. *Contemporary clinical trials* 1998; **19**: S61-S62.

Cierco Jimenez R, Lee T, Rosillo N, Cordova R, Cree IA, Gonzalez A, Indave Ruiz BI. Machine learning computational tools to assist the performance of systematic reviews: A mapping review. *BMC Med Res Methodol* 2022; **22**: 322.

Clark J, McFarlane C, Cleo G, Ishikawa Ramos C, Marshall S. The Impact of Systematic Review Automation Tools on Methodological Quality and Time Taken to Complete Systematic Review Tasks: Case Study. *JMIR Med Educ* 2021; **7**: e24418.

Clark OAC, Castro AA, Atallah AN. Searching LILACS database improves systematic reviews. 6th Cochrane Colloquium; 1998; Baltimore (MD), USA. https://abstracts.cochrane.org/1998-baltimore/searching-lilacs-database-improves-systematic-reviews.

Clark OAC, Castro AA. Cochrane reviews must use LILACS database-like source of articles. 9th Annual Cochrane Colloquium Abstracts; 2001; Lyon. https://cmr.cochrane.org/?CRGReportID=3675.

Clark OAC, Castro AA. Searching the Literatura Latino Americana e do Caribe em Ciencias da Saude (LILACS) database improves systematic reviews. *International journal of epidemiology* 2002; **31**: 112-114.

Cochrane Information Specialist Support Team. Section 5: Handsearching. 2021a. In: Cochrane Information Specialists' Handbook [Internet]. Available from: https://training.cochrane.org/resource/cochrane-information-specialists-handbook.

Cochrane Information Specialist Support Team. Section 3: The Cochrane Register of Studies (CRS). 2021b. In: Cochrane Information Specialists' Handbook [Internet]. Available from: https://training.cochrane.org/resource/cochrane-information-specialists-handbook.

Cochrane Information Specialist Support Team. Section 4: Maintaining and developing a Specialised Register. 2021c. In: Cochrane Information Specialists' Handbook [Internet]. Available from: https://training.cochrane.org/resource/cochrane-information-specialists-handbook.

Cohen JF, Korevaar DA, Wang J, Spijker R, Bossuyt PM. Should we search Chinese biomedical databases when performing systematic reviews? *Systematic reviews* 2015; **4**: 23.

Coleman S, Wright JM, Nixon J, Schoonhoven L, Twiddy M, Greenhalgh J. Searching for Programme theories for a realist evaluation: a case study comparing an academic database search and a simple Google search. *BMC Medical Research Methodology* 2020; **20**: 217.

Cooper C, Rogers M, Bethel A, Briscoe S, Lowe J. A mapping review of the literature on UK-focused health and social care databases. *Health Information and Libraries Journal* 2015; **32**: 5-22.

Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. *Systematic reviews* 2017a; **6**: 234.

Cooper C, Lovell R, Husk K, Booth A, Garside R. Supplementary search methods were more effective and offered better value than bibliographic database searching: A case study from public health and environmental enhancement. *Res Synth Methods* 2017b; **9**: 195-223.

Cooper C, Snowsill T, Worsley C, Prowse A, O'Mara-Eves A, Greenwood H, Boulton E, Strickson A. Handsearching had best recall but poor efficiency when exporting to a bibliographic tool: case study. *J Clin Epidemiol* 2020; **123**: 39-48.

Cooper C, Lorenc T, Schauberger U. What you see depends on where you sit: The effect of geographical location on web-searching for systematic reviews: A case study. *Research Synthesis Methods* 2021a; **12**: 557-570.

Cooper C, Court R, Kotas E, Schauberger U. A technical review of three clinical trials register resources indicates where improvements to the search interfaces are needed. *Research Synthesis Methods* 2021b; **12**: 384-393.

Cooper C, Dawson S, Lefebvre C. Searching for medical devices - Practical guidance. *Res Synth Methods* 2022; **13**: 144-154.

Craane B, Dijkstra PU, Stappaerts K, De Laat A. Methodological quality of a systematic review on physical therapy for temporomandibular disorders: influence of hand search and quality scales. *Clin Oral Investig* 2012; **16**: 295-303.

Craven J, Jefferies J, Kendrick J, Nicholls D, Boynton J, Frankish R. A comparison of searching the Cochrane Library databases via CRD, Ovid and Wiley: implications for systematic searching and information services. *Health Info Libr J* 2014; **31**: 54-63.

Crisan A, Munzner T, Gardy JL. Adjutant: an R-based tool to support topic discovery for systematic and literature reviews. *Bioinformatics* 2019; **35**: 1070-1072.

Curkovic M, Kosec A. Bubble effect: including internet search engines in systematic reviews introduces selection bias and impedes scientific reproducibility. *BMC Medical Research Methodology* 2018; **18**: 130.

Dal-Re R, Ross JS, Marusic A. Compliance with prospective trial registration guidance remained low in high-impact journals and has implications for primary end point reporting. *J Clin Epidemiol* 2016; **75**: 100-107.

De Oliveira GS, Jr., Jung MJ, McCarthy RJ. Discrepancies between randomized controlled trial registry entries and content of corresponding manuscripts reported in anesthesiology journals. *Anesth Analg* 2015; **121**: 1030-1033.

Del Fiol G, Michelson M, Iorio A, Cotoi C, Haynes RB. A Deep Learning Method to Automatically Identify Reports of Scientifically Rigorous Clinical Research from the Biomedical Literature: Comparative Analytic Study. *J Med Internet Res* 2018; **20**: e10281.

Devine J, Egger-Sider F. *Going Beyond Google Again: Strategies for Using and Teaching the Invisible Web*. London: Facet Publishing; 2013.

DeVito NJ, Bacon S, Goldacre B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* 2020; **395**: 361-369.

Dhippayom T, Rattanachaisit N, Wateemongkollert A, Napim R, Chaiyakunapruk N. Should CINAHL be used as one of the main databases for evidence synthesis of health services intervention? *Cochrane Evidence Synthesis and Methods* 2023; **1**: e12019.

Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994; **309**: 1286-1291.

Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 2002; **25**: 38-64.

Diezel K, Pharoah FM, Adams CE. Abstracts of trials presented at the Vth World Congress of Psychiatry (Mexico, 1971): a cohort study. *Psychological medicine* 1999; **29**: 491-494.

Dogpile.com. Different engines, different results: Web searches not always finding what they're looking for 2007. http://www.assessmentpsychology.com/metasearch-analysis.pdf.

Doshi P. FDA to begin releasing clinical study reports in pilot programme. *BMJ* 2018; **360**: k294.

Duddy C, Roberts N. Identifying evidence for five realist reviews in primary health care: A comparison of search methods. *Res Synth Methods* 2022; **13**: 190-203.

Earley A, Lau J, Uhlig K. Haphazard reporting of deaths in clinical trials: a review of cases of ClinicalTrials.gov records and matched publications; a cross-sectional study. *BMJ Open* 2013; **3**: e001963.

Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *HEALTH TECHNOLOGY ASSESSMENT* 2003; **7**: 1-76.

Egilman AC, Kapczynski A, McCarthy ME, Luxkaranayagam AT, Morten CJ, Herder M, Wallach JD, Ross JS. Transparency of Regulatory Data across the European Medicines Agency, Health Canada, and US Food and Drug Administration. *J Law Med Ethics* 2021; **49**: 456-485.

Eisinga A, Siegfried N, Clarke M. The sensitivity and precision of search terms in Phases I, II and III of the Cochrane Highly Sensitive Search Strategy for identifying reports of randomized trials in MEDLINE in a specific area of health care - HIV/AIDS prevention and treatment interventions. *Health Info Libr J* 2007; **24**: 103-109.

Enomoto M, Tseng CH, Hsu YC, Thuy LTT, Nguyen MH. Collaborating with AI in literature search-An important frontier. *Hepatol Commun* 2023; **7**.

Escaldelai FMD, Escaldelai L, Bergamaschi DP. Systematic Review Support software system: web-based solution for managing duplicates and screening eligible studies. *Rev Bras Epidemiol* 2022; **25**: e220030.

EUnetHTA JA3WP6B2-2 Authoring Team. Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness. Diemen (The Netherlands): EUnetHTA; 2019. https://www.eunethta.eu/.

Eysenbach G, Tuische J, Diepgen TL. Evaluation of the usefulness of Internet searches to identify unpublished clinical trials for systematic reviews. *Med Inform Internet Med* 2001; **26**: 203-218.

Falzon L, Trudeau KJ. Developing a database of behavioural medicine interventions. *Health Information and Libraries Journal* 2007; **24**: 257-266.

Farace DJ, Frantzen J. Third International Conference on Grey Literature: Perspectives on the design and transfer of scientific and technical information; 1997; Luxembourg. Amsterdam: TransAtlantic; 1997.

Farace DJ, Frantzen J. Sixth International Conference on Grey Literature: Work on Grey in Progress; 2005; New York. Amsterdam: GreyNet, Grey Literature Network Service; 2005.

Farrah K, Mierzwinski-Urban M. Almost half of references in reports on new and emerging nondrug health technologies are grey literature. *J Med Libr Assoc* 2019; **107**: 43-48.

Fedorowicz Z, Amin F, Eisinga A, Al-Sayyad J. Handsearching for 'buried' randomized trials in Bahrain medical journals [Poster]. 13th Cochrane Colloquium; 2005; Melbourne, Australia. https://abstracts.cochrane.org/2005-melbourne/handsearching-buried-randomized-trials-bahrain-medical-journals.

Felizardo KR, Carver JC. Automating Systematic Literature Review. In: Felderer M, Travassos G, editors. *Contemporary empirical methods in software engineering*. Switzerland: Springer Nature; 2020. p. 327-356. https://doi.org/10.1007/978-3-030-32489-6_12.

Fontaine JF, Barbosa-Silva A, Schaefer M, Huska MR, Muro EM, Andrade-Navarro MA. MedlineRanker: flexible ranking of biomedical literature. *Nucleic Acids Research* 2009; **37**: W141-146.

Gandhi R, Jan M, Smith HN, Mahomed NN, Bhandari M. Comparison of published orthopaedic trauma trials following registration in Clinicaltrials.gov. *BMC musculoskeletal disorders* 2011; **12**: 278.

Garfield E. The evolution of the Science Citation Index. *Int Microbiol* 2007; **10**: 65-69.

Gill CJ. How often do US-based human subjects research studies register on time, and how often do they post their results? A statistical analysis of the Clinicaltrials.gov database. *BMJ Open* 2012; **2**: e001186.

Gillin P. AI hallucinations: The 3% problem no one can fix slows the AI juggernaut:. siliconAngle [Internet]. 2024. Available from: https://siliconangle.com/2024/02/07/ai-hallucinations-3-problem-no-one-can-fix-slows-ai-juggernaut/.

Glanville J, Lefebvre C. Identifying systematic reviews: key resources. *ACP Journal Club* 2000; **132**: A11-12.

Glanville J, Lefebvre C, Manson P, Robinson S, Brbre I, Woods L, (editors). ISSG Search Filter Resource [Internet]. York (UK): The InterTASC Information Specialists' Sub-Group 2006 updated [20 May 2024, Cited 1 July 2024]. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home.

Glanville J, Foxlee R, Wisniewski S, Noel-Storr A, Edwards M, Dooley G. Translating the Cochrane EMBASE RCT filter from the Ovid interface to Embase.com: a case study. *Health Information and Libraries Journal* 2019a; **36**: 264-277.

Glanville J, Dooley G, Wisniewski S, Foxlee R, Noel-Storr A. Development of a search filter to identify reports of controlled clinical trials within CINAHL Plus. *Health Information and Libraries Journal* 2019b; **36**: 73-90.

Glanville J, Kotas E, Featherstone R, Dooley G. Which are the most sensitive search filters to identify randomized controlled trials in MEDLINE? *J Med Libr Assoc* 2020; **108**: 556-563.

Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *J Med Libr Assoc* 2006; **94**: 130-136.

Glanville JM, Duffy S, McCool R, Varley D. Searching ClinicalTrials.gov and the International Clinical Trials Registry Platform to inform systematic reviews: what are the optimal search approaches? *J Med Libr Assoc* 2014; **102**: 177-183.

Godin K, Stapleton J, Kirkpatrick SI, Hanning RM, Leatherdale ST. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. *Systematic reviews* 2015; **4**: 1-10.

Goldacre B, Turner E, on behalf of the OpenTrials team. You can now search FDA approval documents easily at fda.opentrials.net. *BMJ* 2017; **356**: j677.

Goldacre B, DeVito NJ, Heneghan C, Irving F, Bacon S, Fleminger J, Curtis H. Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource. *BMJ* 2018; **362**: k3218.

Gomis M, Gall C, Brahmi FA. Web-based citation management compared to EndNote: options for medical sciences. *Medical Reference Services Quarterly* 2008; **27**: 260-271.

Grames EM, Stillman AN, Tingley MW, Elphick CS. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution* 2019; **10**: 1645-1654.

Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005; **331**: 1064-1065.

Greiner B, Corcoran A, Wheeler D. Clinical trial registry searches are under-utilized in systematic reviews from critical care journals: A bibliometric analysis. *J Crit Care* 2021; **63**: 175-178.

Gresham G, Meinert JL, Gresham AG, Piantadosi S, Meinert CL. Update on the clinical trial landscape: analysis of ClinicalTrials.gov registration data, 2000-2020. *Trials* 2022; **23**: 858.

Greyson D, Rafferty E, Slater L, MacDonald N, Bettinger JA, Dubé È, MacDonald SE. Systematic review searches must be systematic, comprehensive, and transparent: a critique of Perman et al. *BMC Public Health* 2019; **19**: 153.

Guimarães NS, Ferreira AJF, Ribeiro Silva RC, de Paula AA, Lisboa CS, Magno L, Ichiara MY, Barreto ML. Deduplicating records in systematic reviews: there are free, accurate automated ways to do so. *J Clin Epidemiol* 2022; **152**: 110-115.

Hackenbroich S, Kranke P, Meybohm P, Weibel S. Include or not to include conference abstracts in systematic reviews? Lessons learned from a large Cochrane network meta-analysis including 585 trials. *Syst Rev* 2022; **11**: 178.

Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLoS One* 2015; **10**: e0138237.

Haddaway NR, Bayliss HR. Shades of grey: Two forms of grey literature important for reviews in conservation. *Biol Conserv* 2015; **191**: 827-829.

Haddaway NR, Callaghan MW, Collins AM, Lamb WF, Minx JC, Thomas J, John D. On the use of computer-assistance to facilitate systematic mapping. *Campbell Systematic Reviews* 2020; **16**: e1129.

Haddaway NR, Grainger MJ, Gray CT. citationchaser: An R package and Shiny app for forward and backward citations chasing in academic searching. 2021.

Haddaway NR, Grainger MJ, Gray CT. Citationchaser: A tool for transparent and efficient forward and backward citation chasing in systematic searching. *Res Synth Methods* 2022; **13**: 533-545.

Hair K, Bahor Z, Macleod M, Liao J, Sena ES. The Automated Systematic Search Deduplicator (ASySD): a rapid, open-source, interoperable tool to remove duplicate citations in biomedical systematic reviews. *bioRxiv* 2021: 2021.2005.2004.442412.

Hannink G, Gooszen HG, Rovers MM. Comparison of registered and published primary outcomes in randomized clinical trials of surgical interventions. *Ann Surg* 2013; **257**: 818-823.

HarmoniSR Working Group, for the Cochrane Information Specialists' Executive. HarmoniSR: final project report 2015. http://community.cochrane.org/sites/default/files/uploads/inline-files/HarmoniSR%20-%20final%20report%20-Sept%202015_22-09-2015.pdf.

Harrison H, Griffin SJ, Kuhn I, Usher-Smith JA. Software tools to support title and abstract screening for systematic reviews in healthcare: an evaluation. *BMC Med Res Methodol* 2020; **20**: 7.

Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Medical Research Methodology* 2017; **17**: 64.

Hartung DM, Zarin DA, Guise JM, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials.gov Results Database and peer-reviewed publications. *Ann Intern Med* 2014; **160**: 477-483.

Harzing A-W. Publish or Perish 2006updated August 2023. https://harzing.com/resources/publish-or-perish.

Hausner E, Waffenschmidt S, Kaiser T, Simon M. Routine development of objectively derived search strategies. *Syst Rev* 2012; **1**: 19.

Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. *J Clin Epidemiol* 2015; **68**: 191-199.

Hausner E, Knelangen M, Waffenschmidt S. Use of text mining tools in the development of search strategies - Comparison of different approaches. *J Clin Epidemiol* 2022; **149**: 254-256.

Hersh W. Search still matters: information retrieval in the era of generative AI. *J Am Med Inform Assoc* 2024; **31**: 2159-2161.

Hinde S, Spackman E. Bidirectional citation searching to completion: an exploration of literature searching methods. *Pharmacoeconomics* 2015; **33**: 5-11.

Hirt J, Nordhausen T, Fuerst T, Ewald H, Appenzeller-Herzog C. Guidance on terminology, application, and reporting of citation searching: the TARCiS statement. *BMJ* 2024; **385**: e078384.

Hocking R. Yale MeSH Analyzer [Product review]. *Journal of the Canadian Health Libraries Association* 2017; **38**: 125-126.

Hodkinson A, Dietz KC, Lefebvre C, Golder S, Jones M, Doshi P, Heneghan C, Jefferson T, Boutron I, Stewart L. The use of clinical study reports to enhance the quality of systematic reviews: a survey of systematic review authors. *Syst Rev* 2018; **7**: 117.

Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev* 2007a; **2**: MR000010.

Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. *Cochrane Database Syst Rev* 2007b; **2**: MR000001.

Hopewell S, Clarke MJ, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database of Systematic Reviews* 2007c; **2**: MR000011.

Hopkins AM, Modi ND, Rockhold FW, Hoffmann T, Menz BD, Veroniki AA, McKinnon RA, Rowland A, Swain SM, Ross JS, Sorich MJ. Accessibility of clinical study reports supporting medicine approvals: a cross-sectional evaluation. *J Clin Epidemiol* 2024; **167**: 111263.

Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database of Systematic Reviews* 2011; **8**: MR000026.

Hunt DL, McKibbon KA. Locating and Appraising Systematic Reviews. *Ann Intern Med* 1997; **126**: 532-538.

Hunter KE, Webster AC, Page MJ, Willson M, McDonald S, Berber S, Skeers P, Tan-Koay AG, Parkhill A, Seidler AL. Searching clinical trials registers: guide for systematic reviewers. *BMJ* 2022; **377**: e068791.

Huser V, Cimino JJ. Evaluating adherence to the International Committee of Medical Journal Editors' policy of mandatory, timely clinical trial registration. *J Am Med Inform Assoc* 2013a; **20**: e169-e174.

Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to track results of interventional human clinical trials. *PLoS One* 2013b; **8**: e68409.

Isojarvi J, Glanville J. Evidence-based searching for health technology assessment: keeping up to date with SuRe Info. *Int J Technol Assess Health Care* 2021; **37**: e51.

Jefferson T, Jones MA, Doshi P, Del Mar CB, Hama R, Thompson MJ, Spencer EA, Onakpoya I, Mahtani KR, Nunan D, Howick J, Heneghan CJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database of Systematic Reviews* 2014; **4**: CD008965.

Jefferson T, Doshi P, Boutron I, Golder S, Heneghan C, Hodkinson A, Jones M, Lefebvre C, Stewart LA. When to include clinical study reports and regulatory documents in systematic reviews. *BMJ evidence-based medicine* 2018; **23**: 210-217.

Jiang Y, Lin C, Meng W, Yu C, Cohen AM, Smalheiser NR. Rule-based deduplication of article records from bibliographic databases. *Database (Oxford)* 2014; **2014**: bat086.

Jin Q, Leaman R, Lu Z. PubMed and beyond: biomedical literature search in the age of artificial intelligence. *EBioMedicine* 2024; **100**: 104988.

Jones CW, Platts-Mills TF. Quality of registration for clinical trials published in emergency medicine journals. *Annals of emergency medicine* 2012; **60**: 458-464.e451.

Jones CW, Handler L, Crowell KE, Keil LG, Weaver MA, Platts-Mills TF. Non-publication of large randomized clinical trials: cross sectional analysis. *BMJ* 2013; **347**: f6104.

Jorgensen L, Gotzsche PC, Jefferson T. Index of the human papillomavirus (HPV) vaccine industry clinical study programmes and non-industry funded studies: a necessary basis to address reporting bias in a systematic review. *Syst Rev* 2018; **7**: 8.

Kohl C, McIntosh EJ, Unger S, Haddaway NR, Kecke S, Schiemann J, Wilhelm R. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. *Environmental Evidence* 2018; **7**: 8.

Kreutz CK, Schenkel R. Scientific paper recommendation systems: a literature review of recent publications. *Int J Digit Libr* 2022; **23**: 335-369.

Kulkarni AV, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *JAMA* 2009; **302**: 1092-1096.

Kwabena AE, Wiafe OB, John BD, Bernard A, Boateng FAF. An automated method for developing search strategies for systematic review using Natural Language Processing (NLP). *MethodsX* 2022; **10**: 101935.

Kwon Y, Lemieux M, McTavish J, Wathen N. Identifying and removing duplicate records from systematic review searches. *J Med Libr Assoc* 2015; **103**: 184-188.

Ladanie A, Ewald H, Kasenda B, Hemkens LG. How to use FDA drug approval documents for evidence syntheses. *BMJ* 2018; **362**: k2815.

Lanera C, Minto C, Sharma A, Gregori D, Berchialla P, Baldi I. Extending PubMed searches to ClinicalTrials.gov through a machine learning approach for systematic reviews. *J Clin Epidemiol* 2018; **103**: 22-30.

Lauer M. NIH Clinical Trials Reporting Compliance: A Shared Commitment. 2023. https://nexus.od.nih.gov/all/2023/03/24/nih-clinical-trials-reporting-compliance-a-shared-commitment/.

Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. *BMC Med Res Methodol* 2012; **12**: 51.

Lefebvre C, Clarke M. Identifying randomised trials. In: Egger M, Davey Smith G, Altman DG, editors. *Systematic Reviews in Health Care: Meta-analysis in Context*. 2nd ed. London (UK): BMJ Publication Group; 2001.

Lefebvre C, Eisinga A, McDonald S, Paul N. Enhancing access to reports of randomized trials published world-wide - the contribution of EMBASE records to the Cochrane Central Register of Controlled Trials (CENTRAL) in *The Cochrane Library*. *Emerg Themes Epidemiol* 2008; **5**: 13.

Levay P, Raynor M, Tuvey D. The Contributions of MEDLINE, Other Bibliographic Databases and Various Search Techniques to NICE Public Health Guidance. *Evidence Based Library and Information Practice* 2015; **10**: 50-68.

Levay P, Ainsworth N, Kettle R, Morgan A. Identifying evidence for public health guidance: a comparison of citation searching with Web of Science and Google Scholar. *Res Synth Methods* 2016; **7**: 34-45.

Levay P, Craven J. Systematic Searching in a Post-Pandemic World: New Directions for Methods, Technology, and People. *Evidence Based Library and Information Practice* 2023; **18**: 93-104.

Lexchin J, Herder M, Doshi P. Canada finally opens up data on new drugs and devices. *BMJ* 2019; **365**: l1825.

Linder SK, Kamath GR, Pratt GF, Saraykar SS, Volk RJ. Citation searches are more sensitive than keyword searches to identify studies using specific measurement instruments. *J Clin Epidemiol* 2015; **68**: 412-417.

Loan FA, Parray UY, Khan AM. Performance of search engines in harvesting grey literature. *Library Hi Tech News* 2022; **39**: 21-23.

Loder E, Loder S, Cook S. Characteristics and publication fate of unregistered and retrospectively registered clinical trials submitted to The BMJ over 4 years. *BMJ Open* 2018; **8**: e020037.

Lokker C, Bagheri E, Abdelkader W, Parrish R, Afzal M, Navarro T, Cotoi C, Germini F, Linkins L, Haynes RB, Chu L, Iorio A. Deep learning to refine the identification of high-quality clinical research articles from the biomedical literature: Performance evaluation. *J Biomed Inform* 2023; **142**: 104384.

Lorenzetti DL, Ghali WA. Reference management software for systematic reviews and meta-analyses: an exploration of usage and usability. *BMC Med Res Methodol* 2013; **13**: 141.

Lunny C, McKenzie JE, McDonald S. Retrieval of overviews of systematic reviews in MEDLINE was improved by the development of an objectively derived and validated search strategy. *J Clin Epidemiol* 2015; **74**: 107-118.

Mahood Q, Van Eerd D, Irvin E. Searching for grey literature for systematic reviews: challenges and benefits. *Res Synth Methods* 2014; **5**: 221-234.

Manriquez JJ. Searching the LILACS database could improve systematic reviews in dermatology. *Archives of dermatology* 2009; **145**: 947-948.

Marshall IJ, Noel-Storr A, Kuiper J, Thomas J, Wallace BC. Machine learning for identifying Randomized Controlled Trials: an evaluation and practitioner's guide. *Research Synthesis Methods* 2018; **9**: 602-614.

McDonald S. Improving access to the international coverage of reports of controlled trials in electronic databases: a search of the Australasian Medical Index. *Health Info Libr J* 2002; **19**: 14-20.

McGowan BS. Using Text Mining Tools to Inform Search Term Generation: An Introduction for Librarians. *Libraries and the Academy* 2021; **21**: 603-618.

McKeown S, Mir ZM. Considerations for conducting systematic reviews: evaluating the performance of different methods for de-duplicating references. *Systematic reviews* 2021; **10**: 38.

Mesgarpour B, Mesgarpour M, Herkner H. SensPrecOptimizer: a software tool that combined search queries to design efficient search strategies. *J Clin Epidemiol* 2016; **71**: 122-123.

Metzendorf MI, Weibel S, Reis S, McDonald S. Pragmatic and open science-based solution to a current problem in the reporting of living systematic reviews. *BMJ Evid Based Med* 2023; **28**: 267-272.

Miron L, Gonçalves RS, Musen MA. Obstacles to the reuse of study metadata in ClinicalTrials.gov. *Sci Data* 2020; **7**: 443.

Montori VM, Wilczynski NL, Morgan D, Haynes RB, for the Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005; **330**: 68.

Moseley AM, Sherrington C, Elkins MR, Herbert RD, Maher CG. Indexing of randomised controlled trials of physiotherapy interventions: a comparison of AMED, CENTRAL, CINAHL, EMBASE, Hooked on Evidence, PEDro, PsycINFO and PubMed. *Physiotherapy* 2009; **95**: 151-156.

Muller A, Ames H, Himmels J, Jardim P, Nguyen L, Rose C, Van de Velde S. Implementation of machine learning in evidence syntheses in the Cluster for Reviews and Health Technology Assessments: Final report 2020-20212021:[1-83 pp.]. Available from: 10.13140/RG.2.2.32822.42562.

Napoleão B, Petrillo F, Hallé S. Automated support for searching and selecting evidence in software engineering : a cross-domain systematic mapping. 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA); 2021: IEEE. https://ieeexplore.ieee.org/document/9582563.

Nasser M, Al Hajeri A. A comparison of handsearching versus EMBASE searching of the Archives of Iranian Medicine to identify reports of randomized controlled trials. *Archives of Iranian Medicine* 2006; **9**: 192-195.

Noel-Storr A, Wisniewski S. Evaluation of Cochrane's Evidence Pipeline for populating CENTRAL. Global Evidence Summit; 2024 September 10-13, 2024; Czech Republic, Prague. https://programme.ges24.myeventapp.com/.

Noel-Storr AH, Dooley G, Wisniewski S, Glanville J, Thomas J, Cox S, Featherstone R, Foxlee R. Cochrane Centralised Search Service showed high sensitivity identifying randomized controlled trials: A retrospective analysis. *J Clin Epidemiol* 2020; **127**: 142-150.

Novoa J, Chagoyen M, Benito C, Moreno FJ, Pazos F. PMIDigest: Interactive Review of Large Collections of PubMed Entries to Distill Relevant Information. *Genes (Basel)* 2023; **14**.

O'Keefe H, Rankin J, Wallace SA, Beyer F. Investigation of text-mining methodologies to aid the construction of search strategies in systematic reviews of diagnostic test accuracy—a case study. *Research Synthesis Methods* 2023; **14**: 79-98.

O'Mara-Eves A, Brunton G, McDaid D, Kavanagh J, Oliver S, Thomas J. Techniques for identifying cross-disciplinary and 'hard-to-detect' evidence for systematic review. *Research Synthesis Methods* 2014; **5**: 50-59.

O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic reviews* 2015; **4**: 5.

Ogilvie D, Hamilton V, Egan M, Petticrew M. Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *J Epidemiol Community Health* 2005; **59**: 804-808.

Oxman A, Chalmers I, Clarke M, Enkin M, Schulz K, Starr M, Dickersin K, Herxheimer A, Silagy C (editors). Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Collaboration; 1994.

Paez A. Gray literature: An important resource in systematic reviews. *J Evid Based Med* 2017; **10**: 233-240.

Page MJ, Shamseer L, Tricco AC. Registration of systematic reviews in PROSPERO: 30,000 records and counting. *Syst Rev* 2018; **7**: 32.

Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hrobjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, McKenzie JE. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021a; **372**: n160.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hrobjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021b; **372**: n71.

Pallath A, Zhang Q. Paperfetcher: A tool to automate handsearching and citation searching for systematic reviews. *Res Synth Methods* 2023; **14**: 323-335.

Paludan-Müller AS, Maclean-Nyegaard IR, Munkholm K. Substantial delays in clinical data published by the European Medicines Agency - a cross sectional study. *J Clin Epidemiol* 2022; **146**: 68-76.

Papaioannou D, Sutton A, Carroll C, Booth A, Wong R. Literature searching for social science systematic reviews: consideration of a range of search techniques. *Health Information and Libraries Journal* 2010; **27**: 114-122.

Paynter R, Banez L, Berliner E, Erinoff E, Lege-Matsuura J, Potter S, Uhl S. EPC Methods: An exploration of the use of text-mining software in systematic reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016. https://www.ncbi.nlm.nih.gov/books/NBK362044/.

Paynter RA, Featherstone R, Stoeger E, Fiordalisi C, Voisin C, Adam GP. A prospective comparison of evidence synthesis search strategies developed with and without text-mining tools. *J Clin Epidemiol* 2021; **139**: 350-360.

Paynter RA, Featherstone R, Stoeger E, Fiordalisi C, Voisin C, Adam GP. Reply to Hausner et al. Re: use of text mining tools in the development of search strategies-comparison of different approaches. *J Clin Epidemiol* 2022; **149**: 256-257.

Petticrew M, Song F, Wilson P, Wright K. Quality-assessed reviews of health care interventions and the database of abstracts of reviews of effectiveness (DARE). *International Journal of Technology Assessment in Healthcare* 1999; **15**: 671-678.

Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. *JAMA* 1999; **281**: 1110-1111.

Premji Z, Ganshorn H. Can database-level MEDLINE exclusion filters in Embase and CINAHL be used to remove duplicate records without loss of relevant studies in systematic reviews? An exploratory study. *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada* 2020; **41**: 3-15.

Ramachandran R, Morten CJ, Ross JS. Strengthening the FDA's Enforcement of ClinicalTrials.gov Reporting Requirements. *JAMA* 2021; **326**: 2131-2132.

Rathbone J, Carter M, Hoffmann T, Glasziou P. Better duplicate detection for systematic reviewers: evaluation of Systematic Review Assistant-Deduplication Module. *Syst Rev* 2015a; **4**: 6.

Rathbone J, Hoffmann T, Glasziou P. Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. *Syst Rev* 2015b; **4**: 80.

Rathbone J, Carter M, Hoffmann T, Glasziou P. A comparison of the performance of seven key bibliographic databases in identifying all relevant systematic reviews of interventions for hypertension. *Syst Rev* 2016; **5**: 27.

Richards D. Handsearching still a valuable element of the systematic review. *Evidence-Based Dentistry* 2008; **9**: 85.

Roberts S, Atsawawaranunt K, Adams CE. Thai online bibliographical biomedical databases [Poster] 17th Cochrane Colloquium; 2009; Singapore. https://abstracts.cochrane.org/2009-singapore/thai-databases.

Rogers M, Bethel A, Talens-Bou J, Briscoe S. Chasing references: a comparison of Scopus, Web of Science and Google Scholar for forward citation searching. CILIP Health Libraries Group Conference; 2016; Scarborough. https://www.cilip.org.uk/members/group_content_view.asp?group=200697&id=856306.

Rogers M, Bethel A, Abbott R. Locating qualitative studies in dementia on MEDLINE, EMBASE, CINAHL, and PsycINFO: A comparison of search strategies. *Res Synth Methods* 2017; **9**: 579-586.

Rogers M, Bethel A, Briscoe S. Resources for forwards citation searching for implementation studies in dementia care: A case study comparing Web of Science and Scopus. *Research Synthesis Methods* 2020; **11**: 379-386.

Røst TB, Slaughter L, Nytrø Ø, Muller AE, Vist GE. Using neural networks to support high-quality evidence mapping. *BMC Bioinformatics* 2021; **22**: 496.

Roth S, Wermer-Colan A. Machine Learning Methods for Systematic Reviews: A Rapid Scoping Review. *Dela J Public Health* 2023; **9**: 40-47.

Royle PL, Bain L, Waugh NR. Sources of evidence for systematic reviews of interventions in diabetes. *DIABETIC MEDICINE* 2005; **22**: 1386-1393.

Sadek J, Inskip A, Woltmann J, Wilkins G, Marshall C, Pokora M, Vedpathak A, Jadrevska A, Craig D, Trenell M. ScanMedicine: An online search system for medical innovation. *Contemp Clin Trials* 2023; **125**: 107042.

Sag M. The New Legal Landscape for Text Mining and Machine Learning (February 27, 2020). *Journal of the Copyright Society of the USA* 2019; **66**: 291.

Salami K, Alkayed K. Publication bias in pediatric hematology and oncology: analysis of abstracts presented at the annual meeting of the American Society of Pediatric Hematology and Oncology. *Pediatric hematology and oncology* 2013; **30**: 165-169.

Saldanha IJ, Smith BT, Ntzani E, Jap J, Balk EM, Lau J. The Systematic Review Data Repository (SRDR): descriptive characteristics of publicly available data and opportunities for research. *Systematic reviews* 2019; **8**: 334.

Saleh AA, Ratajeski MA, Bertolet M. Grey Literature Searching for Health Sciences Systematic Reviews: A Prospective Study of Time Spent and Resources Utilized. *Evid Based Libr Inf Pract* 2014; **9**: 28-50.

Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel)* 2023; **11**.

Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Development of an efficient search filter to retrieve systematic reviews from PubMed. *J Med Libr Assoc* 2021; **109**: 561-574.

Sampson M, McGowan J, Cogo E, Horsley T. Managing database overlap in systematic reviews using Batch Citation Matcher: case studies using Scopus. *Journal of the Medical Library Association* 2006; **94**: 461-463.

Sampson M, de Bruijn B, Urquhart C, Shojania K. Complementary approaches to searching MEDLINE may be sufficient for updating existing systematic reviews. *J Clin Epidemiol* 2016; **78**: 108-115.

Saric L, Vucic K, Dragicevic K, Vrdoljak M, Jakus D, Vuka I, Jelicic Kadic A, Saldanha IJ, Puljak L. Comparison of conference abstracts and full-text publications of randomized controlled trials presented at four consecutive World Congresses of Pain: Reporting quality and agreement of results. *Eur J Pain* 2019; **23**: 107-116.

Scells H, Zuccon G, Koopman B. A comparison of automatic Boolean query formulation for systematic reviews. *Information Retrieval Journal* 2021; **24**: 3-28.

Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2018; **11**: MR000005.

Scherer RW, Saldanha IJ. How should systematic reviewers handle conference abstracts? A view from the trenches. *Syst Rev* 2019; **8**: 264.

Schmidt L, Sinyor M, Webb RT, Marshall C, Knipe D, Eyles EC, John A, Gunnell D, Higgins JPT. A narrative review of recent tools and innovations toward automating living systematic reviews and evidence syntheses. *Z Evid Fortbild Qual Gesundhwes* 2023; **181**: 65-75.

Schmidt L, Mohamed S, Meader N, Bacardit J, Craig D. Automated data analysis of unstructured grey literature in health research: A mapping review. *Res Synth Methods* 2024; **15**: 178-197.

Schmucker CM, Blumle A, Schell LK, Schwarzer G, Oeller P, Cabrera L, von Elm E, Briel M, Meerpohl JJ, on behalf of the OPEN consortium. Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. *PLoS One* 2017; **12**: e0176210.

Schoonbaert D. SPIRS, WinSPIRS, and OVID: a comparison of three MEDLINE-on-CD-ROM interfaces [see comment]. *Bulletin of the Medical Library Association* 1996; **84**: 63-70.

Scopus. Scopus Factsheet 2024. 2024. https://www.elsevier.com/products/scopus.

Sheikh A, Siddique N, Qutab S, Khan MA, Mahmood K. An investigation of emerging COVID-19 research trends and future implications for LIS field: A bibliometric mapping and visualization. *J Librariansh Inf Sci* 2023; **55**: 3-17.

Shemilt I, Noel-Storr A, Thomas J, Featherstone R, Mavergames C. Machine learning reduced workload for the Cochrane COVID-19 Study Register: development and evaluation of the Cochrane COVID-19 Study Classifier. *Syst Rev* 2022; **11**: 15.

Shemilt I, Arno A, Thomas J, Lorenc T, Khouja C, Raine G, Sutcliffe K, Preethy D, Kwan I, Wright K, Sowden A. Cost-effectiveness of Microsoft Academic Graph with machine learning for automated study identification in a living map of coronavirus disease 2019 (COVID-19) research. *Wellcome Open Research* 2024; **6**: 210.

Shojania KG, Bero LA. Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. *Eff Clin Pract* 2001; **4**: 157-162.

Shokraneh F, Adams CE. Study-based registers of randomized controlled trials: Starting a systematic review with data extraction or meta-analysis. *BioImpacts* 2017; **7**: 209-217.

Simon C, Davidsen K, Hansen C, Seymour E, Barnkob MB, Olsen LR. BioReader: a text mining tool for performing classification of biomedical literature. *BMC Bioinformatics* 2019; **19**: 57.

Sinha A, Shen Z, Song Y, Ma H, Elde D, Hsu B-J, Wang K. An overview of Microsoft Academic Service (MAS) and applications. Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion); 2015; Florence. http://dx.doi.org/10.1145/2740908.2742839.

Slobogean GP, Verma A, Giustini D, Slobogean BL, Mulpuri K. MEDLINE, EMBASE, and Cochrane index most primary studies but not abstracts included in orthopedic meta-analyses. *J Clin Epidemiol* 2009; **62**: 1261-1267.

Stansfield C, Brunton G, Rees R. Search wide, dig deep: literature searching for qualitative research. An analysis of the publication formats and information sources used for four systematic reviews in public health. *Res Synth Methods* 2014; **5**: 142-151.

Stansfield C, Dickson K, Bangpan M. Exploring issues in the conduct of website searching and other online sources for systematic reviews: how can we be systematic? *Syst Rev* 2016; **5**: 191.

Stansfield C, O'Mara-Eves A, Thomas J. Text mining for search term development in systematic reviewing: A discussion of some methods and challenges. *Res Synth Methods* 2017; **8**: 355-365.

Stansfield C, Stokes G, Thomas J. Applying machine classifiers to update searches: Analysis from two case studies. *Res Synth Methods* 2022; **13**: 121-133.

Stevinson C, Lawlor DA. Searching multiple databases for systematic reviews: added value or diminishing returns? *Complementary therapies in medicine* 2004; **12**: 228-232.

Stovold E, Hansen S. Handsearching respiratory conference abstracts: a comparison with abstracts identified by an EMBASE search [Poster]. 19th Cochrane Colloquium; 2011; Madrid. https://abstracts.cochrane.org/2011-madrid/handsearching-respiratory-conference-abstracts-comparison-abstracts-identified-embase.

Subirana M, Sola I, Garcia JM, Gich I, Urrutia G. A nursing qualitative systematic review required MEDLINE and CINAHL for study identification. *J Clin Epidemiol* 2005; **58**: 20-25.

Talebi R, Redberg RF, Ross JS. Consistency of trial reporting between ClinicalTrials.gov and corresponding publications: one decade after FDAAA. *Trials* 2020; **21**: 675.

Thomas J, McDonald S, Noel-Storr A, Shemilt I, Elliott J, Mavergames C, Marshall IJ. Machine learning reduced workload with minimal risk of missing studies: development and evaluation of a randomized controlled trial classifier for Cochrane Reviews. *J Clin Epidemiol* 2021; **133**: 140-151.

Tramèr MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997; **315**: 635-640.

Treanor L, Frank RA, Cherpak LA, Dehmoobad Sharifabadi A, Salameh JP, Hallgrimson Z, Fabiano N, McGrath TA, Kraaijpoel N, Yao J, Korevaar DA, Bossuyt PM, McInnes MDF. Publication bias in diagnostic imaging: conference abstracts with positive conclusions are more likely to be published. *Eur Radiol* 2020; **30**: 2964-2972.

Tumin D, Akpan US, Kohler JA, Sr., Uffman JC. Publication Bias Among Conference Abstracts Reporting on Pediatric Quality Improvement Projects. *Am J Med Qual* 2020; **35**: 274-280.

Urhan TK, Rempel HG, Meunier-Goddik L, Penner MH. Information Retrieval in Food Science Research II: Accounting for Relevance When Evaluating Database Performance. *Journal of Food Science* 2019; **84**: 2729-2735.

Urlings MJE, Duyx B, Swaen GMH, Bouter LM, Zeegers MP. Citation bias and other determinants of citation in biomedical research: findings from six citation networks. *J Clin Epidemiol* 2021; **132**: 71-78.

US National Library of Medicine. PubMed Update: Proximity Search Now Available in PubMed. NLM Tech Bull [Internet]. 2022 May 31st 2023; 449(e4). Available from: https://www.nlm.nih.gov/pubs/techbull/nd22/nd22_pubmed_proximity_search_available.html.

Uttley L, Quintana DS, Montgomery P, Carroll C, Page MJ, Falzon L, Sutton A, Moher D. The problems with systematic reviews: a living systematic review. *J Clin Epidemiol* 2023; **156**: 30-41.

van Dinter R, Tekinerdogan B, Catal C. Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology* 2021; **136**: 106589.

van Driel ML, De Sutter A, De Maeseneer J, Christiaens T. Searching for unpublished trials in Cochrane reviews may not be worth the effort. *J Clin Epidemiol* 2009; **62**: 838-844.

Vickers AJ, Smith C. Incorporating data from dissertations in systematic reviews. *International Journal of Technology Assessment in Healthcare* 2000; **16**: 711-713.

Waffenschmidt S, Hausner E, Kaiser T. An evaluation of searching the German CCMed database for the production of systematic reviews. *Health Information and Libraries Journal* 2010; **27**: 262-267.

Wallace BC, Small K, Brodley CE, Lau J, Trikalinos TA. Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium; Miami, Florida, USA: Association for Computing Machinery; 2012. p. 819–824.

Wallace BC, Noel-Storr A, Marshall IJ, Cohen AM, Smalheiser NR, Thomas J. Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. *Journal of the American Medical Informatics Association* 2017; **24**: 1165-1168.

Wang LL, Lo K. Text mining approaches for dealing with the rapidly expanding literature on COVID-19. *Brief Bioinform* 2021; **22**: 781-799.

Wang S, Scells H, Koopman B, Zuccon G. Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search? *arXiv* 2023: 1426-1436.

Wanner A, Baumann N. Design and implementation of a tool for conversion of search strategies between PubMed and Ovid MEDLINE. *Res Synth Methods* 2018; **10**: 154-160.

Watson RJ, Richardson PH. Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of Embase, Medline and PsycINFO bibliographic databases. *BRITISH JOURNAL OF MEDICAL PSYCHOLOGY* 1999a; **72**: 535-542.

Watson RJ, Richardson PH. Accessing the literature on outcome studies in group psychotherapy: the sensitivity and precision of Medline and PsycINFO bibliographic database searching. *BRITISH JOURNAL OF MEDICAL PSYCHOLOGY* 1999b; **72**: 127-134.

White VJ, Glanville JM, Lefebvre C, Sheldon TA. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *Journal of information science* 2001; **27**: 357-370.

Wieland LS, Manheimer E, Sampson M, Barnabas JP, Bouter LM, Cho K, Lee MS, Li X, Liu J, Moher D, Okabe T, Pienaar ED, Shin BC, Tharyan P, Tsutani K, van der Windt DA, Berman BM. Bibliometric and content analysis of the Cochrane Complementary Medicine Field specialized register of controlled trials. *Syst Rev* 2013; **2**: 51.

Wilczynski NL, Haynes RB, for the Hedges Team. EMBASE search strategies achieved high sensitivity and specificity for retrieving methodologically sound systematic reviews. *J Clin Epidemiol* 2007; **60**: 29-33.

Wilczynski NL, Haynes RB. Consistency and accuracy of indexing systematic review articles and meta-analyses in MEDLINE. *Health Information and Libraries Journal* 2009; **26**: 203-210.

Wilczynski NL, McKibbon KA, Haynes RB. Sensitive Clinical Queries retrieved relevant systematic reviews as well as primary studies: an analytic survey. *J Clin Epidemiol* 2011; **64**: 1341-1349.

Wildgaard L, Vils A, Sandal Johnsen S. Reflections on tests of AI-search tools in the academic search process. *LIBER Quarterly: The Journal of the Association of European Research Libraries* 2023; **33**: 1-34.

Wright K, Golder S, Rodriguez-Lopez R. Citation searching: a systematic review case study of multiple risk behaviour interventions. *BMC Med Res Methodol* 2014; **14**: 73.

Wright K, Golder S, Lewis-Light K. What value is the CINAHL database when searching for systematic reviews of qualitative studies? *Syst Rev* 2015; **4**: 104.

Wu X-Y, Tang J-L, Mao C, Yuan J-Q, Qin Y, Chung VCH. Systematic reviews and meta-analyses of traditional Chinese medicine must search Chinese databases to reduce language bias. *Evid Based Complement Alternat Med* 2013: 812179.

Xia J, Wright J, Adams CE. Five large Chinese biomedical bibliographic databases: accessibility and coverage. *Health Information and Libraries Journal* 2008; **25**: 55-61.

Xue J, Chen W, Chen L, Gaudet L, Moher D, Walker M, Wen SW. Significant discrepancies were found in pooled estimates of searching with Chinese indexes versus searching with English indexes. *J Clin Epidemiol* 2016; **70**: 246-253.

Younger P, Boddy K. When is a search not a search? A comparison of searching the AMED complementary health database via EBSCOhost, OVID and DIALOG. *Health Information and Libraries Journal* 2009; **26**: 126-135.

Zhang L, Lu W, Chen H, Huang Y, Cheng Q. A comparative evaluation of biomedical similar article recommendation. *J Biomed Inform* 2022; **131**: 104106.

Ziai H, Zhang R, Chan A-W, Persaud N. Search for unpublished data by systematic reviewers: an audit. *BMJ Open* 2017; **7**: e017737.