## Appendix A: Glossary

This glossary contains definitions of important terms used in this article.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database coverage</td>
<td>The proportion of eligible records available in a database. A database has high coverage of a topic if a large percentage of topic-relevant records found by any means are available in this database. For example, an SR may have searched ten different information sources, but a retrospective analysis could find all included studies are indexed in MEDLINE. This indicates that MEDLINE has high coverage of this topic.</td>
</tr>
<tr>
<td>Eligibility criteria (Also: inclusion &amp; exclusion criteria, PICOS)</td>
<td>Criteria defining the parameters for the inclusion and exclusion of studies in a review. They are based on the research question and defined during the protocol phase. Preliminary searches and known relevant records usually help define the finalised eligibility criteria. The full systematic search (database and supplementary searching) can only be conducted based on the finalised eligibility criteria.</td>
</tr>
<tr>
<td>Grey literature</td>
<td>Non-published evidence, or evidence released outside of traditional commercial or academic publishing, in contrast to evidence released through traditional publishing (i.e., in a journal article or a book). This can include various digital (and printed) documents, such as study register entries, preprints (but not in-press versions of journal articles), reports of government and non-government organizations, dissertations and theses, conference abstracts and proceedings, etc.</td>
</tr>
<tr>
<td>Information specialist</td>
<td>In the context of evidence synthesis development, information specialist are professional searchers who have expertise in the methodology of systematic searching. Information specialists often have a background in Library and Information Science. (see also Appendix C, section 1.1)</td>
</tr>
<tr>
<td>Known relevant records (Also: key studies, test set, benchmark records)</td>
<td>Literature relevant to the research question that is already known before the systematic search strategy is finalised and searches are run. These texts are usually found through preliminary searching. Ideally, they should meet the eligibility criteria at least at abstract screening level. Known relevant studies can be used to identify search terms, choose relevant information sources, and validate search strategies. (See Appendix C, section 1.3.)</td>
</tr>
<tr>
<td>Limiting function (Also: limits, filters)</td>
<td>Features built into a database or interface that allow search results to be narrowed by specific criteria (e.g., publication date ranges, language, gender, age group, publication type). Some filtering functions are based on validated search filters (see below), but not all. To avoid unintentionally removing eligible studies from the search result, the searcher must understand the underlying mechanism of the limiting functions they intend to use.</td>
</tr>
</tbody>
</table>
| Precision                                                  | A metric to assess the effectiveness of a literature search. The proportion of eligible/relevant documents in a search result: \[
\frac{\text{Relevant documents identified}}{\text{Total search result}} \times 100
\] Precision and sensitivity are inversely related. A highly precise search will not be very sensitive. In comprehensive systematic searches, sensitivity is valued over precision. A focused systematic search aims to increase the precision of a search without sacrificing too much sensitivity. (In the context of systematic searching "relevant documents" are determined by a systematic literature selection process. They are identical with the references to studies that meet the inclusion criteria of the review and are included in the final review. Precision might be calculated for a single search source or across all information sources.) |
| **Preliminary searches**  
(Also: scoping searches, exploratory searches) | Any searches conducted during the protocol phase to inform the further review process. Unlike the systematic search proper, these searches are not usually documented or reported in the final review report.  
(See Appendix C, section 1.2.) |
| **Search filters**  
(Also: hedges) | Predefined database- and interface-specific search queries that can be used in a larger search strategy to identify references meeting certain criteria (e.g., study design, age group). Filters usually combine free-text terms (e.g., searching in title abstract) and subject headings (e.g., MeSH terms). Validated search filters have been assessed for sensitivity, specificity, and precision based on a gold standard set of eligible references. |
| **Search strategy** | The interface-specific Boolean search query run in a particular database (e.g., MEDLINE) using a particular interface (e.g., Ovid). When several databases are searched, this query is usually developed in a primary database (e.g., Ovid MEDLINE) and translated to other databases/interfaces. Different databases and interfaces offer different search options and search queries need to be adapted accordingly. |
| **Search syntax** | Syntax refers to the correct expression of a search query in a specific interface. It includes the fields that can be searched (e.g., title, abstract), the available operators (e.g., Boolean, proximity), and other functionalities such as truncation, wild cards, and types of phrase searching. Syntax is interface AND database-specific. For example, the same MEDLINE search query will look different if expressed for Ovid MEDLINE or PubMed. On the other hand, two search strategies searching MEDLINE and Embase via the Ovid interface will look broadly similar, except for database specific subject headings and data fields. |
| **Sensitivity**  
(Also: recall) | A metric to assess the effectiveness of a literature search. The proportion of relevant documents found with a specific search strategy (e.g., Ovid MEDLINE) divided by the total number of relevant documents included in the review.  
\[
\text{Sensitivity} = \left( \frac{\text{Relevant documents identified}}{\text{Total number of relevant documents (in the source)}} \right) \times 100
\]  
Precision and Sensitivity are inversely related. A highly precise search will not be very sensitive. In comprehensive systematic searches, sensitivity is valued over precision. A focused, systematic search aims to increase the precision of a search without sacrificing too much sensitivity. To validate a search strategy means assessing sensitivity of a database search based on the retrieval of benchmark records. (In the context of systematic searching "relevant documents" are determined by a systematic literature selection process. They are identical with the references to studies that meet the inclusion criteria of the review and are included in the final review. Sensitivity might be calculated for a single search source or across all information sources.) |
| **Study selection**  
(Also: literature screening, study selection, abstract and full-text screening) | Search results are assessed for their suitability for inclusion in the review based on the predefined eligibility criteria. (A separate paper in this series will focus on considerations for literature screening in RRs.) |
| **Systematic search** | Information retrieval processes aimed to systematically identify eligible studies for a review. The systematic search combines several information sources, and usually includes database search strategies and supplementary search techniques. The systematic search process must be as transparent and reproducible as possible and |
be reported in the final review report.
Appendix B: Useful resources

Guidance for systematic searching


Templates for planning & documenting searches


Tools

IEBH Systematic Review Accelerator: https://sr-accelerator.com/#/

- Includes tools for text analysis, search strategy translation, deduplication of search results, and citation searching
PubMed PubReMiner: https://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi
  ➤ Simple word frequency analysis of MEDLINE records

Citation Chaser: https://estech.shinyapps.io/citationchaser/
  ➤ Pulls backwards and forwards citations for a set of seed papers

Citation Finder: https://citation-finder.vercel.app/
  ➤ Can be used to create RIS or Bibtex files from a reference list

Risklick Deduklick: https://www.risklick.ch/products/deduklick/
  ➤ A web-application for automated deduplication of search results

Websites

Summarized Research in Information Retrieval for HTA (SuRe Info):
https://sites.google.com/york.ac.uk/sureinfo/home
  ➤ Web resource that provides research-based information relating to the information retrieval aspects of producing systematic reviews and health technology assessments

MIAR: https://miar.ub.edu/
  ➤ Identifies which databases index a specific journal

The ISSG Search Filter Resource: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home
  ➤ Collection of -mainly published and validated- search filters/hedges for various study designs and topics

University Library Vrije Universiteit Amsterdam - Search blocks: https://blocks.bmi-online.nl/
  ➤ Collection of non-validated search strings for wide variety of topics

searchRxiv: https://searchrxiv.org/
  • Repository for search strategies

  • Various search filters for MEDLINE and Embase via the Ovid interface

Resources for grey literature searching


Glanville, J. and C. Lefebvre. (18 May 2022). "Finding clinical trials, research registers and research results." from https://sites.google.com/a/york.ac.uk/yhectrialsregisters/home


Text analysis and text mining for search strategy design


Appendix C: Elaboration and practical considerations

Content
1 Preparation and planning .............................................................................................................................................. 8
   1.1 Involvement of information specialists ................................................................................................................. 8
   1.2 Preliminary searches .................................................................................................................................................... 8
   1.3 Known relevant records .................................................................................................................................................. 9
2 Information sources and search methods ........................................................................................................................ 10
   2.1 General considerations for selection of RR information sources .............................................................................. 10
   2.2 Example: Reviews based on randomised controlled trials (RCTs) ........................................................................... 11
      2.2.1 What is known .......................................................................................................................................................... 11
      2.2.2 Possible streamlined search approaches .................................................................................................................. 12
   2.3 Example: COVID-19 reviews ......................................................................................................................................... 12
      2.3.1 What is known .......................................................................................................................................................... 12
      2.3.2 Considerations for streamlined search approaches ............................................................................................... 13
   2.4 Testing the coverage of the intended primary database ............................................................................................. 13
3 Search strategies ............................................................................................................................................................. 13
   3.1 Text analysis ................................................................................................................................................................. 14
   3.2 Improving search precision and sensitivity .................................................................................................................. 14
      3.2.1 Database/Interface-specific search functionalities ................................................................................................. 16
   3.3 Translating search strategies ........................................................................................................................................... 16
   3.4 Removing specific document types form database search results .............................................................................. 17
      3.4.1 Ovid MEDLINE ....................................................................................................................................................... 17
      3.4.2 Ovid Embase ......................................................................................................................................................... 17
      3.4.3 Embase.com ......................................................................................................................................................... 17
      3.4.4 Scopus.com ......................................................................................................................................................... 17
      3.4.5 Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Library ............................................ 17
   3.5 Search filters ................................................................................................................................................................. 18
4 Quality assurance and search strategy peer review ........................................................................................................ 18
   4.1 Validation of search strategies ...................................................................................................................................... 18
   4.2 Search peer review ....................................................................................................................................................... 19
5 Reporting and record management ................................................................................................................................... 20
   5.1 Search reporting ........................................................................................................................................................... 20
   5.2 Record management ................................................................................................................................................... 20
   5.3 Deduplication of search results ................................................................................................................................. 21
1 Preparation and planning

1.1 Involvement of information specialists
The involvement of information specialists (or librarians) is widely recommended in systematic search guidance.1-4

In the restricted timeframe of a rapid review (RR), author teams might fear that involving such an expert causes unnecessary delays and try to save time by designing and running the searches themselves. However, information specialist involvement can save time and improve review quality in the long run.5-9 Because RR searches should be conducted as efficiently as possible, it is advisable to involve an information specialist that already has experience in conducting systematic searches:

- Having more routine, experienced searchers can develop and conduct searches more efficiently and are less likely to make grave errors;
- they can assist with formulating searchable review questions;
- they are more likely to be familiar with a wide range of information sources and methods and able to assess their suitability for the RR topic;
- they are more likely to be familiar with search reporting and quality assurance standards.

1.2 Preliminary searches
Taking time to plan and prepare the search process adequately will save time in later review steps. A crucial step is conducting preliminary/scoping/exploratory searches1,10 to:

1) improve the topic understanding and refine the key questions,
2) identify existing systematic reviews (SR) and a first set of potentially relevant primary studies ("known relevant records"), and
3) estimate the resources necessary to perform the RR.

Unlike the systematic search proper, these searches are not usually documented or reported in the final review report. They do not have to meet the systematic search’s standards for transparency and reproducibility. However, a limited documentation of the preliminary search process can be useful, in particular for broad or complex topics where the refinement of the research question may take several iterations.

An important step for preliminary searching is identifying existing SRs. If a recent SR on the topic already exists, conducting an RR may be unnecessary. However, updating an older SR can also be a viable approach for an RR.

SRs can also be mined for included studies that are likely to be eligible for the new RR (i.e., known relevant records). In this case, the existing SR does not need to investigate the same question as the planned RR. It only has to include some studies that could be eligible for the planned RR.

Useful sources to find SRs are:

- Cochrane Database of Systematic Reviews (https://www.cochranelibrary.com/)
- Epistemonikos (https://www.epistemonikos.org/)
- Health Evidence (https://www.healthevidence.org/)
- International HTA database (https://database.inahta.org/)
- KSR Evidence (https://ksrevidence.com/)
- Trip Database (https://www.tripdatabase.com/)
Preliminary searches are usually iterative. If no SRs are found, highly precise search queries can be used to find a few promising primary studies. For this step, an information source that allows sorting by relevance or “best match” is useful, for example, PubMed (https://pubmed.ncbi.nlm.nih.gov/) or Google Scholar (https://scholar.google.com/).

Based on these known, and potentially relevant studies, a citation-based search method can be used. This means using "seed articles" to identify further citations of interest. Citation-based search methods include: backward citation searches (i.e., reference list checking) and forward citation searches (i.e. "cited by" searching);¹¹ co-cited references (i.e., articles that are cited together with the seed article) and co-citing references (i.e., articles that share similar reference lists with the seed article);¹¹ and similar article functions in databases and search engines.¹² The advantage of collecting studies included in SRs and other citation-based methods is that the search results are not produced by search terms the searcher chooses. Thus, the known relevant records are less likely to be biased by the terminology the searcher is already aware of. (This is important if search terms for the systematic search strategy are based on text analysis of known relevant records.)

Some useful sources for citation-based searches are:

- CoCites (https://www.cocites.com/, co-cited references)
- Connected Papers (https://www.connectedpapers.com, similarity based co-cited references & co-citing references)
- Google Scholar (https://scholar.google.com/, similar articles, forward citation searching)
- Lens.org Scholarly Search and Analysis (https://www.lens.org/lens/search/scholar/structured, backward & forward citation searching, similar articles)
- Scopus (backward & forward citation searching, co-citing references; subscription required)
- Web of Science (backward & forward citation searching, co-citing references; subscription required)

Academic search engines that use artificial intelligence/machine learning support to generate search results may also be useful in preliminary searches. Their functionalities tend to be suited to simple, precision focused searches that do not require high sensitivity, reproducibility or transparency. Like citation-based searches, these search engines can help to find literature that does not contain the specific search terms the searcher chooses.

Some examples are:

- Elicit: https://elicit.org/
- Scite : https://scite.ai/
- Semantic Scholar: https://www.semanticscholar.org/

### 1.3 Known relevant records

At the end of the preliminary search process, a set of studies that meet the RR eligibility criteria at the abstract level should be available. The aim for these “known relevant records” is to be as representative of the evidence base as possible. This is particularly important if text analysis methods are to be used for search term identification.¹³ In this regard, studies identified through previous SRs are particularly useful to get a broad picture of the evidence available.

These known relevant records can be further used to 1) test the coverage of chosen databases (i.e., Are relevant records available in this database?), 2) identify relevant search terms, 3) test the
performance of search strategies (i.e., Does the search find the relevant studies?), and 4) as a starting
point for some additional search methods (e.g., citation and similar articles searching).

The optimal number of known relevant records depends largely on the scope of the topic and the
terminological variety: A topic like the "effectiveness of Mogamulizumab in Cutaneous T-cell
lymphoma" is both focussed and has a low terminological variety. This means a small set of records
may be representative of the entire evidence base. On the other hand, the topic of "methods for
reducing aggression in psychiatric wards" is likely to use a more varied terminology, because eligible
studies will likely contain a wide variety of interventions (e.g., pharmacological, behavioural,
administrative). In this case, a larger set of records may be necessary to represent the evidence base.

Depending on these considerations, we suggest aiming for 5-20 references for the set of known
relevant records.

2 Information sources and search methods

In SRs, a comprehensive search process maximises the likelihood of finding all topic-relevant studies,
both published and unpublished (i.e., grey literature). This is achieved using multiple information
sources, including bibliographic databases and supplementary methods. One advantage of this
approach is, that because of the content-overlap between the various sources, relevant studies
missed by one search (e.g., a MEDLINE search strategy) are likely to be picked up by another search
(e.g., a CENTRAL search strategy, reviewing reference lists of existing reviews, etc.). For this reason,
MECIR requires 'stacking' of sources, and defines a topic-independent minimum set of sources
comprised of bibliographic databases (i.e., MEDLINE, EMBASE, CENTRAL), study registers (i.e.,
ClinicalTrials.gov, WHO ICTRP), and citation-based searching (i.e., reference lists of included studies).
To these standard sources, additional sources need be added depending on the topic (e.g., a topic-
specific bibliographic database). That way, even if some of these sources do not have a high topic
coverage, or if a single search strategy is not very sensitive, the combination of all sources increases
the likelihood of finding all topic-relevant studies.

In RRs, reducing the number of information sources used is a practical way to reduce the time spent
on searching and the size of the search result. However, this can impact the sensitivity of searches.
While a lower degree of certainty in RR conclusions is an acceptable trade-off for quicker results, RR
searches still aim for high enough sensitivity to draw conclusions.25

This means it is essential that RR searches select information sources and search methods that are
most likely to identify relevant studies.

Retrospective studies have shown that, on average, MEDLINE in particular has very high coverage of
studies included in healthcare SRs.16-18 This also indicates that grey literature often plays a minor role
in SRs. However, when selecting a limited number of information sources for RRs, several important
cautives must be considered.

2.1 General considerations for selection of RR information sources

- The coverage of MEDLINE can vary widely across topics. For example, Frandsen et al.17 found
  that Cochrane reviews of the Cystic Fibrosis and Genetic Disorders Group had an average
  PubMed coverage of less than 50%. Levay et al.19 found an even lower coverage for some
  NICE public health reviews due to the large percentage of grey literature included in the
  reviews.

- Coverage is not the same as sensitivity. While a high percentage of relevant studies are often
  available in MEDLINE, that does not mean a search strategy will identify them with a
  manageable number of results. Consequently, studies show that combining two MECIR
databases is more likely to find a high percentage of relevant literature than a single
database.\textsuperscript{16}  
- Depending on the topic, specialised or transdisciplinary databases can play an important
role. For example, the combination of CINAHL and MEDLINE has been shown to be efficient
in finding qualitative studies in diabetes.\textsuperscript{20} In contrast, PsycInfo had the highest coverage for
faith-based mental health interventions,\textsuperscript{21} and Scopus and Web of Science had the highest
coverage of health care management research.\textsuperscript{22}  
An analysis of NICE public health guidelines\textsuperscript{23} showed that MEDLINE and Embase were part of
the minimum required databases to find all included studies. However, the full set of
minimum required information sources varied by topic, and always included supplementary
search techniques, in particular website searching. (See COVID-19 below as an example of a topic where less common information sources
should be considered.)
- Research that evaluates the impact of information sources on the results or conclusions of
SRs (as opposed to search sensitivity), is based mainly on recalculating meta-analyses of RCTs
simulating different search approaches.\textsuperscript{15,24,25} It is not clear how applicable those findings are
to SRs without meta-analyses or SRs relying on other study designs. (See example below for
relevant information sources for reviews of RCTs.)

2.2 Example: Reviews based on randomised controlled trials (RCTs)
Information sources and search methods for reviews of RCTs are relatively well-researched. This
makes it possible to make some evidence-based decisions on streamlined search approaches.

2.2.1 What is known
- Most Cochrane intervention reviews include RCTs only. Limiting meta-analyses of these
reviews to studies available in MEDLINE only changes the results in rare cases.\textsuperscript{24,25}  
- MEDLINE, Embase, and CENTRAL each have very high coverage of RCTs, and combining any
two of these databases is usually enough to find a majority of available RCTs. This
combination often suffices to arrive at the same review conclusions as a comprehensive
search.\textsuperscript{15}  
- A large number of validated RCT search filters for MEDLINE exist.\textsuperscript{26} There are also various
search filters for most other commonly used databases.\textsuperscript{27}  
- CENTRAL is a database for RCTs and quasi-RCTs. This means no search filter is necessary.
Unlike MEDLINE and Embase, CENTRAL also contains grey literature, in particular study
register entries and conference abstracts.\textsuperscript{28} It has a very high coverage of RCTs but is updated
less frequently than databases like MEDLINE or Embase.\textsuperscript{29}  
- In general, RCTs are more likely to be published than observational studies.\textsuperscript{30} In most cases,
excluding grey literature has little to no effect on meta-analyses of RCTs.\textsuperscript{31,32} Of the different
types of grey literature, excluding conference abstracts seems to have the least impact.\textsuperscript{32}  
- The vast majority of RCTs of newly approved drugs are registered in ClinicalTrials.gov and can
be found in this study register.\textsuperscript{34} Non-pharmacological RCTs seem to be less consistently
registered in ClinicalTrials.gov.\textsuperscript{35}  
- ClinicalTrials.gov does not only include prospective study registration data. In some cases,
results are also available. Additionally, PubMed-indexed articles containing an NCT Number
will automatically be listed on the ClinicalTrials.gov entry.\textsuperscript{36}
Follow up of eligible study register entries is a practical way to identify completed pharmacological trials missed by database searches, but rarely affects the results of meta-analyses.37

Combining a single database search with reference list checking of included articles will improve search sensitivity. However, combining MEDLINE and reference list checking might be less effective than combining the method with Embase or CENTRAL search strategies.35

The combination of a simple MEDLINE search plus PubMed Similar Articles based on selected studies has been shown to be an effective search approach for updating SRs based on RCTs on clinically-focused topics.12,38 There is some indication that this approach also works for newly created reviews on clinically-focused topics.39

2.2.2 Possible streamlined search approaches

For RRs based on RCTs, combining two of the MECIR-required databases (i.e., MEDLINE, CENTRAL, Embase) is likely to be the most robust approach.

However, depending on the topic, combining MEDLINE with a supplementary search technique may also be acceptable. This has the advantage that access to both MEDLINE via PubMed and supplementary sources like ClinicalTrials.gov and PubMed similar articles is free.

2.3 Example: COVID-19 reviews

The COVID-19 pandemic is an example of a rapidly evolving topic. This means on the one hand, that search approaches developed for more stable topics are not necessarily effective. On the other hand, search approaches for COVID-19 topics also changed over time because information sources, publishing characteristics, etc. changed.

2.3.1 What is known

At the beginning of the pandemic, there was no standardized terminology for Sars-Cov-2/COVID-19, leading to a large variety of potential search terms. Standardised subject headings for MEDLINE and Embase only became available after several months.40

Analyses up to the 2nd quarter of 2020 show that the majority of Sars-Cov-2/COVID-19 research findings was available through pre-prints, not published (peer-reviewed) journal articles.41,42 Of the available journal articles, about a quarter were Chinese.41 Most study results used in the Cochrane living systematic review and network meta-analysis on COVID-19 (COVID-NMA) came from pre-prints.43

In early SRs, RCTs played a minor role. The most commonly included study design were cohort studies.44 Modelling studies also played an important role.45

The vast majority of registered Covid-19 RCTs seem to concern pharmacological topics.46,47

Available specialised information sources and their search functions changed over time, meaning that search approaches for review updates needed to be adapted constantly.45

At the time of writing (March 2022), there are several databases for COVID-19/Sars-Cov-2. Typically, these are meta-databases that collect references to published articles, preprints, and study register entries from various sources. Some examples are:

- Cochrane COVID-19 Study Register (https://COVID-19.cochrane.org/)
- COVID-19 L·OVE (https://app.iloveevidence.com/loves/5e6fdb9669c00e4ac072701d)
- COVID-NMA (https://covid-nma.com/)
Some evaluations of coverage of COVID-19 L·OVE and the Cochrane COVID-19 Study Register are available.\textsuperscript{48-50} Based on studies included in SRs, both had very high coverage, with L·OVE containing nearly all cited studies.

There are various Sars-Cov-2/COVID-19 search strategies available for MEDLINE, Embase, and other healthcare research databases\textsuperscript{51,52}, but there are no current validated search filters at the time of writing. Database providers have also incorporated search filters into their interfaces.\textsuperscript{53-55}

As of 2021, the National Institute for Health and Care Excellence (NICE) conducted a single comprehensive weekly surveillance search for all of their COVID-19 rapid guidelines instead of separate searches for each topic.\textsuperscript{40}

2.3.2 Considerations for streamlined search approaches

- Many analyses of document types, study designs, and information sources are based on analyses of data from 2020 and might not reflect the current situation.
- Available information sources and their functionalities have changed throughout the pandemic. Searchers preparing or updating a search on a COVID-19 topic should first understand the currently available sources and not solely rely on what worked in the past.
- Currently, there are several high-coverage specialised databases. Searches should use at least one of them. A combination of two specialised databases is likely to retrieve almost all available evidence.
- The evidence base is still evolving rapidly, so RR searches should include pre-prints to find the most recent study results. For pharmacological topics, study registers can also be a useful information source.

2.4 Testing the coverage of the intended primary database

Databases generally release information about their scope and coverage of journals or topics.\textsuperscript{56-58} Additionally, methods studies assessing coverage of databases for specific use-cases are often available.\textsuperscript{16,17,22,49} Searchers should be aware of the general coverage of the databases they intend to use.

To assess the usefulness of databases for a specific RR, known relevant records identified during the preliminary search can be used. In particular, it is helpful to search these records in the intended primary database (i.e., the database where the search strategy is developed, usually MEDLINE). The Accession Number (e.g., PMID in MEDLINE) should be recorded if they are available in the database. (The Accession Numbers will be later used to validate the search strategy, see next chapter.)

Ideally, the primary database should contain most if not all published known relevant records. This is to make sure that the primary database is highly relevant for the chosen topic: Suppose a large percentage of these known records are not available in the selected database. In that case, the searcher needs to assess why this is the case and consider the implication for selecting information sources. For example, if the missing records are published journal articles but not available in the assessed database, it is important to check if there is another bibliographic database with higher topic coverage. Conversely, if the majority of missing records are grey literature, supplementary searches might be important to answer the research question.

3 Search strategies

The construction of a search strategy to optimise the balance between sensitivity and specificity according to the intended purpose of the search is critical to all evidence synthesis products. Achieving this balance is even more important in the context of an RR. A strategy retrieving too many search results is prohibitive in terms of the time to screen results. Conversely, a strategy with too
many limitations and a correspondingly small result set undermines confidence in findings, limiting the capacity of a decision-maker to make recommendations for action.15

3.1 Text analysis
Tools such as PubMed PubReminer (https://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi), Systematic Review Accelerator Word Frequency Analyzer (https://sr-accelerator.com/#/), or Voyant Tools (https://voyant-tools.org/) can be used to identify terms or subject headings that frequently appear in the known relevant records found by preliminary searching. This analysis can inform the selection of the most relevant search terms for a topic.

In general, there are two ways to use this type of analysis for search strategy development. First, as support for selecting search terms that will also be supplemented by other means (e.g., synonyms listed for MeSH or Emtree terms, the inclusion of word variants that did not appear in the known relevant records) and secondly, as a unique source of search terms in an “objectively-derived” search strategy.59,60 For the objective approach, the known relevant records (called “test set”) have to be a “quasi-gold standard” (i.e., they must be representative of the evidence available for the research question). A truly objectively-derived search strategy also needs prospectively defined rules for the inclusion of terms into to the search strategy.

Note that even methods that automatically generate Boolean queries require substantial human input to be effective (e.g., to identify relevant concepts and terms61, to refine automatically generated queries62).

Overall, text analysis methods can improve the sensitivity of search strategies and accelerate search strategy development.59,63 However, the gains likely depend on the techniques used, the experience of the searcher with the technology, and the test set quality.64

Appendix B lists references to guidance for using text analysis in search strategy development.

3.2 Improving search precision and sensitivity
An important way to reduce the number of records retrieved by a database search is to increase search precision without sacrificing too much sensitivity.

For SRs it is not uncommon to present all search terms for a particular element (e.g., population) within a single search string. This assumes that each term is equally useful in retrieving relevant references. A more appropriate approach for RRs is to assume diminishing returns as one moves from core terms to less useful ones. Dividing search terms for each element into “definites”, “probables” and “possibles” allows you to calibrate your strategy for relevance and numbers of results. Text analysis can be useful to identify which terms are most commonly used to describe a particular element in the literature. For example, “heart failure”, “cardiac failure”, and “myocardial failure” are synonyms. But “heart failure” is the most commonly used, making it a “definite” search term. “Cardiac failure” is not as common and would be a “probable” search term. “Myocardial failure” is very rarely used and could be a “possible” search term. Omitting such a possible search term from the search strategy is likely to reduce the size of the search result without losing relevant studies.

The impact of adding and removing search terms should be evaluated during search strategy development. So, for example, you might decide only to use “definites” for the population while using “definites” and “probables” for the intervention. Two important considerations in that regard are the sensitivity – are known relevant records retrieved? – and the number of search results generated by each set. This process is closely related to validating the search strategy using known relevant records.
The Table 1 shows an example of how dividing search terms can be used to evaluate the sensitivity and precision of different search term configurations. Result 1, using definite terms, has a sensitivity of 81% and a precision of 9% based on the known relevant records. Result 2, using definite and probable terms, has a sensitivity of 92% and precision of 9%. Result 3 adds possible terms and has 100% sensitivity and 4% precision. This means including the “probable” search terms (Result 2) increased sensitivity without sacrificing precision, making it the most efficient option. Note that while Result 3 found two additional known relevant records, the number of overall results increased by 2.6 times from 267 to 698. This is a huge increase compared to Result 2, which only adds 31 records to Result 1.

**Table 1: Ovid MEDLINE Search strategy: contraceptive use and exercise performance**

<table>
<thead>
<tr>
<th>#</th>
<th>Searches</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>exp Contraceptives, Oral/</td>
<td>51105</td>
</tr>
<tr>
<td>2</td>
<td>(oral adj (contraceptiv* or estradiol or progestin or progestogen or progestrone)).ti,ab,kf.</td>
<td>27279</td>
</tr>
<tr>
<td>3</td>
<td>1 or 2</td>
<td>60719</td>
</tr>
<tr>
<td>4</td>
<td>exp Athletic Performance/</td>
<td>59804</td>
</tr>
<tr>
<td>5</td>
<td>exp Physical Endurance/</td>
<td>36115</td>
</tr>
<tr>
<td>6</td>
<td>Exercise Test/</td>
<td>66882</td>
</tr>
<tr>
<td>7</td>
<td>Muscle Strength/</td>
<td>24054</td>
</tr>
<tr>
<td>8</td>
<td>(exercise adj2 (performance or response? or recovery or exhaustion or tolerance or test*).ti,ab,kf.</td>
<td>57349</td>
</tr>
<tr>
<td>9</td>
<td>(endurance adj1 (physical or performance or test*).ti,ab,kf.</td>
<td>3974</td>
</tr>
<tr>
<td>10</td>
<td>(fitness adj1 (physical or cardio*).ti,ab,kf.</td>
<td>18970</td>
</tr>
<tr>
<td>11</td>
<td>or/4-10</td>
<td>176527</td>
</tr>
<tr>
<td>12</td>
<td>3 and 11</td>
<td>236</td>
</tr>
<tr>
<td>13</td>
<td>(peak adj4 (performance or output or capacity)).ti,ab,kf.</td>
<td>8172</td>
</tr>
<tr>
<td>14</td>
<td>((anaerobic or aerobic) adj (capacity or power)).ti,ab,kf.</td>
<td>9049</td>
</tr>
<tr>
<td>15</td>
<td>(vo2peak or vo2max or &quot;v'o2peak&quot; or &quot;v'o2max&quot;).ti,ab,kf.</td>
<td>12988</td>
</tr>
<tr>
<td>16</td>
<td>((max* or peak) adj2 (muscle action? or force production)).ti,ab,kf.</td>
<td>540</td>
</tr>
<tr>
<td>17</td>
<td>(muscle adj (recovery or strength)).ti,ab,kf.</td>
<td>26673</td>
</tr>
<tr>
<td>18</td>
<td>cardiorespiratory response.ti,ab,kf.</td>
<td>360</td>
</tr>
<tr>
<td>19</td>
<td>or/13-18</td>
<td>53544</td>
</tr>
<tr>
<td>20</td>
<td>3 and (11 or 19)</td>
<td>267</td>
</tr>
<tr>
<td>21</td>
<td>exp Exercise/</td>
<td>226979</td>
</tr>
<tr>
<td>22</td>
<td>exp Sports/</td>
<td>203054</td>
</tr>
<tr>
<td>23</td>
<td>(exercise or sport? or physical* activ*).ti,kf.</td>
<td>221780</td>
</tr>
<tr>
<td>24</td>
<td>or/21-23</td>
<td>426679</td>
</tr>
<tr>
<td>25</td>
<td>3 and (11 or 19 or 24)</td>
<td>698</td>
</tr>
<tr>
<td>26</td>
<td>Known relevant records</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>12 and 26</td>
<td>21</td>
</tr>
<tr>
<td>28</td>
<td>20 and 26</td>
<td>24</td>
</tr>
<tr>
<td>29</td>
<td>25 and 26</td>
<td>26</td>
</tr>
<tr>
<td>30</td>
<td>20 not 12</td>
<td>29</td>
</tr>
<tr>
<td>31</td>
<td>(3 and (4 or 5 or 6)) not (8 or 9 or 10)</td>
<td>88</td>
</tr>
<tr>
<td>32</td>
<td>(3 and (8 or 9 or 10)) not (4 or 5 or 6)</td>
<td>76</td>
</tr>
</tbody>
</table>
For search strategies developed in PubMed, Systematic Review Accelerator SearchRefinery (https://sr-accelerator.com/#/) can also be used to assess the sensitivity and precision of each search term in a search strategy. The tool visualizes how many records each term retrieves and how many of the benchmark records it finds, making it easier to decide which terms to remove from the search string.

If the representativeness of the known relevant records is uncertain, another option is to review the difference in search results. Query 30 in Table 1 retrieves only records found by Result 2 but not Result 1. The searcher can assess the relevance of these records by reviewing the top results for the query. If it contains articles that are likely to be eligible, these search terms should be included in the final search strategy. Query 31 shows articles in Result 1 that were retrieved by the MeSH search but do not contain the “definitive” free-text search terms. Conversely, query 32 shows Result 1 articles retrieved by the free-text search, not the MeSH search. These results can be reviewed for additional free-text search terms (query 31) or MeSH terms (query 32) that could be used in search strategy.

3.2.1 Database/Interface-specific search functionalities
The functionalities of the database (e.g., MEDLINE, Embase) and interface (e.g., Ovid, ProQuest) also play an important part in improving the balance of search sensitivity and precision:

- Most database interfaces offer proximity/adjacency operators. These operators allow to set a maximum distance between two search terms. They are more precise than AND-combinations but less restrictive than phrases. They can be used to cover a wider variety of phrase combinations. For example, the Ovid MEDLINE query (exercise adj2 response?).ti,ab. finds "exercise response", "exercise responses", "response to exercise", "responses to exercise", etc.
- Most databases with a controlled vocabulary offer the option to restrict subject headings to records where they are the focus of the article (e.g., major descriptors in MeSH).
- Limiting free-text terms to occurrence in the title also can improve search precision: The most important concepts of a study are often mentioned in the title. For example, searching for exercise.ti. will mainly retrieve articles where exercise is the focus of the study.
- Ovid interfaces also offer a frequency operator, which allows a searcher to define a minimum number of times a term has to appear in a certain field. For example, methods.ab./freq=2 will only retrieve articles where the word “methods” appears at least twice in the abstract. This helps to retrieve articles of methods research: Structured abstracts will mention “methods” at least once. However, if the word appears more than once, it is more likely that topic of the article is methods research.

It is also important to note that search functionalities, syntax, subject headings, and even coverage may change over time. Searchers need to keep up with these changes to ensure the correct use of the database/interface.

3.3 Translating search strategies
When translating a primary search strategy to different databases/interfaces, using Systematic Review Accelerator Polyglot Search (https://sr-accelerator.com/#/) for automated syntax translation can reduce time and translation errors. However, some manual adjustments are generally still necessary. For example, subject headings will need to be reviewed manually because they are database specific (e.g., MeSH and Emtree are similar, but the terms are not identical). Additionally, some elements of the search strategy may not be relevant for all subsequent databases (e.g., using an RCT filter in CENTRAL is unnecessary).
3.4 Removing specific document types from database search results

To efficiently design search strategies, knowledge of the type of database contents is essential. Some databases include conference abstracts, study register entries, theses and dissertations, or other non-research publications (e.g., book reviews). It therefore may be sensible to exclude some document types from the search result.

Below are some examples for database- and interface-specific strategies to exclude results by document type.

3.4.1 Ovid MEDLINE

- Remove animal studies: `not (animals/ not humans/)`
- Limit to database entry dates (this is not the same as publication date!):
The Create Date (DT) field contains the date when it was added to PubMed. [https://ospguides.ovid.com/OSPguides/medline.htm#DT](https://ospguides.ovid.com/OSPguides/medline.htm#DT)
- Remove animal studies, comments, editorials, reviews, guidelines, and case reports:
  `not ((animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/ or hi.fs. or case report.mp.)`
When using this string please cite: Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. [https://dx.doi.org/10.1002/jrsm.1425](https://dx.doi.org/10.1002/jrsm.1425)

3.4.2 Ovid Embase

- Remove MEDLINE records: `not medline.cr.`
- Remove animal studies: `not (exp animal/ not exp human/)`
- Remove conference abstracts or conference reviews: `not (Conference Abstract or Conference Review).pt.`
- Limit to database entry dates (this is not the same as publication date!):
The Date Created (DC) field contains the date of the last activity on the citation before creation of an XML file for delivery to Ovid. The date is loaded from the Elsevier XML element `<date-created>`. [https://ospguides.ovid.com/OSPguides/embase.htm#DC](https://ospguides.ovid.com/OSPguides/embase.htm#DC)

3.4.3 Embase.com

- Remove conference abstracts: `NOT 'conference abstract'/it`
- Remove records that are directly imported form MEDLINE: `NOT [medline]/lim NOT ([embase classic]/lim AND [medline]/lim)`
- Remove preprints: `NOT [preprint]/lim`

3.4.4 Scopus.com

- Remove MEDLINE records: `AND NOT INDEX(medline)`

3.4.5 Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Library

- Remove study register records:
  `not (clinicaltrials or trialsearch or ANZCTR or ensaiosclinicos or chictr or cris or ctri or registroclinicor or clinicaltrialregister or DRKS or IRCT or rctporta or JapicCTI or IMACCT or jRCT or JPRN or UMIN or trialregister or PACTR or REPEC or SLCTR):so`
- Remove conference abstracts: `not (conference:pt or abstract:so)`
- Remove non-english records:
3.5 Search filters

Validated search filters are database- and interface-specific search queries that can be used in a larger search strategy (e.g., to identify study designs). Their sensitivity, specificity, and precision has been assessed, and often two or more versions of a filter exist: one focusing on sensitivity, one on specificity, and one “best balance option”. In SRs, the most sensitive filter is often chosen, but depending on the topic and the number of search results, a specificity-maximizing or “best balance” filter could be the better option for an RR.

The impact of a search filter on a specific search strategy can be assessed in the same way as the search term selection shown in Table 1. Depending on the size of the unfiltered search result, a more sensitive or a more precise filter could be chosen to achieve a manageable search result. For example, the search Result in Table 2 (query 12) retrieves less than 300 records without a study design filter. If we want to limit the search to RCTs, using the Cochrane sensitivity-maximizing RCT filter\(^{67}\) (96% sensitivity, 14% precision)\(^{26}\) will ensure that we miss very few eligible RCTs while still having a reasonably small search result. If the unfiltered search result was 3000 records, using the Cochrane Sensitivity and precision maximizing RCT filter\(^{67}\) (sensitivity 93%, 46% precision)\(^{26}\) might be preferable. It is slightly less sensitive but more precise and will thus reduce the search considerably.

The InterTASC Information Specialists’ Sub-Group (ISSG) Search Filter Resource (https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home) collects validated and unvalidated search filters. It also provides a list of publications that have evaluated the performance of some search filters and links to literature related to search filters (e.g., impact on search results, utility of built-in filters).

A newer approach to identifying particular study designs are machine learning classifiers.\(^{68}\) These are sometimes integrated into systematic review production platforms (e.g., EPPI-Reviewer\(^{69}\)) and can be applied to uploaded search results.\(^{70}\)

4 Quality assurance and search strategy peer review

4.1 Validation of search strategies

We recommend that the primary database search strategy is validated using known relevant records. This is usually part of the search strategy development and means testing if the primary search strategy retrieves known relevant records found through preliminary searching. If some known relevant records are not identified, the searcher has to assess the reasons and decide if revisions are necessary and feasible. Note that improving the sensitivity of a search will increase the size of the search result. Using known relevant records to assess the search result makes it easier to find a balance of sensitivity and precision appropriate for the specific RR project.

We recommend that a search strategy that optimizes precision and sensitivity aims to identify at least 80-90% of known relevant records. This recommendation is based on benchmarks commonly used in developing validated search filters. Search filters generally aim for 95% sensitivity or more for
highly sensitive filters, but sensitivity-precision-maximising filters tend to have a sensitivity of 80% or more.\textsuperscript{71,72} However, using such a cut-off for general search strategy development depends on the known relevant records found by the preliminary search. If the searcher is not confident that the set is fairly representative of the available evidence, it is advisable to aim for retrieval of all known relevant records.

Table 2 shows assessment steps of a search strategy. The search finds 81% of the 26 known relevant records (query 14), which is slightly above the proposed cut-off. Further investigation shows that Element 1 finds all known records (query 15), but Element 2 misses 5 (query 16). These 5 records should be assessed to determine what terminologies are used to describe the concept of “exercise performance”. The next step would be adding these additional search terms to the search strategy and evaluating if the sensitivity can be improved without a large increase in the number of records retrieved.

<table>
<thead>
<tr>
<th>#</th>
<th>Searches</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Element 1: oral contraceptive use</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>exp Contraceptives, Oral/</td>
<td>51105</td>
</tr>
<tr>
<td>2</td>
<td>(oral adj (contraceptiv* or estradiol or progestin or progestogen or progesterone)).ti,ab,kf.</td>
<td>27279</td>
</tr>
<tr>
<td>3</td>
<td>1 or 2</td>
<td>60719</td>
</tr>
<tr>
<td><strong>Element 2: exercise performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>exp Athletic Performance/</td>
<td>59804</td>
</tr>
<tr>
<td>5</td>
<td>exp Physical Endurance/</td>
<td>36115</td>
</tr>
<tr>
<td>6</td>
<td>Exercise Test/</td>
<td>66882</td>
</tr>
<tr>
<td>7</td>
<td>Muscle Strength/</td>
<td>24054</td>
</tr>
<tr>
<td>8</td>
<td>(exercise adj2 (performance or response? or recovery or exhaustion or tolerance or test*)).ti,ab,kf.</td>
<td>57349</td>
</tr>
<tr>
<td>9</td>
<td>(endurance adj1 (physical or performance or test*)).ti,ab,kf.</td>
<td>3974</td>
</tr>
<tr>
<td>10</td>
<td>(fitness adj1 (physical or cardi*)).ti,ab,kf.</td>
<td>18970</td>
</tr>
<tr>
<td>11</td>
<td>or/4-10</td>
<td>176527</td>
</tr>
<tr>
<td><strong>Result</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3 and 11</td>
<td>236</td>
</tr>
<tr>
<td><strong>Known relevant records</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>(&quot;32666247&quot; or &quot;31663173&quot; or &quot;31686212&quot; or &quot;30167957&quot; or &quot;28906053&quot; or &quot;28497386&quot; or &quot;27898641&quot; or &quot;25694209&quot; or &quot;25519952&quot; or &quot;24504652&quot; or &quot;2296028&quot; or &quot;22948447&quot; or &quot;22446669&quot; or &quot;2203922&quot; or &quot;21848445&quot; or &quot;21399539&quot; or &quot;20227547&quot; or &quot;18054842&quot; or &quot;17990209&quot; or &quot;17157107&quot; or &quot;16596112&quot; or &quot;15598669&quot; or &quot;15618333&quot; or &quot;14707778&quot; or &quot;12706609&quot; or &quot;12381756&quot;).ui.</td>
<td>26</td>
</tr>
<tr>
<td><strong>Found by Result</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>12 and 13</td>
<td>21</td>
</tr>
<tr>
<td><strong>Found by E1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>3 and 13</td>
<td>26</td>
</tr>
<tr>
<td><strong>Found by E2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>11 and 13</td>
<td>21</td>
</tr>
<tr>
<td><strong>NOT found by E2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>13 not 16</td>
<td>5</td>
</tr>
</tbody>
</table>

4.2 Search peer review

The Peer Review Electronic Search Strategies (PRESS) checklist\textsuperscript{73} is a tool for information specialists/librarians. It assesses six domains of the database search strategy design: translation of the research question; use of Boolean and proximity operators; use of subject headings; implementation of text word searching; errors in spelling, syntax, and line numbers; and use of limits and filters.
Because it is a specialist tool, both the searcher and the person reviewing the search should be experts who are familiar with the methods and vocabulary of systematic searching. In cases were search experts review non-expert searches, using a more individualized approach adjusted to the searcher’s knowledge base might be necessary.

PRESS also only focusses on the database search strategy itself. Another important consideration is the appropriateness of the information sources overall. RRs that limit the number of information sources have to select the most likely sources to find relevant studies. For this reason, search peer reviewers should also assess whether the chosen sources or methods are likely to have high coverage of the relevant evidence and allow for the creation of sensitivity-and-precision-optimising searches.

A major reason for not conducting search peer review in RRs are concerns about turn-around time for the full process of finding a peer reviewer, giving the peer reviewer time to assess the search, and implementing changes suggested in peer review. However, if in-house information specialists are available, this turn-around time will usually be less than a day. In other cases, it is advisable to plan ahead by contacting potential peer reviewers as soon as work on the RR begins and agree on deadlines for delivery of the search strategy and feedback.

If expert peer review is not at all possible within the time frame, quality assurance is still necessary, in particular if the search was not designed by an information specialist. The primary search strategy must be checked for errors in spelling and use of Boolean operators (AND, OR, NOT). This means examining the entirety of the database search, and understanding how the search terms are connected to form a search strategy. The person reviewing the search strategy must have enough search experience to understand the syntax used and be familiar with systematic database searching (i.e., use of subject headings, free text search terms, and Boolean operators).

5 Reporting and record management

5.1 Search reporting
Any systematic search should be reported in as transparent and reproducible a manner as possible. This means that the search steps must be documented as the search is being developed and conducted. Appendix B provides some links to templates for search documentation.

At a minimum, search reporting includes a complete list of information sources and search techniques used, the search dates, and the complete search strategies of the database searches. Reporting should also state if the search strategies were peer-reviewed and how search results were deduplicated. PRISMA-5 provides a useful checklist and examples for search reporting. While not all items necessarily apply to a streamlined RR search, the overall guidance on reporting is still applicable to RRs.

5.2 Record management
Record management concerns the flow of citations throughout the whole review process, not just the search. This means keeping track of the number of records identified by each source (e.g., MEDLINE, citation-based searching), but also the deduplication process, the number screened at each stage, and the final number included or excluded.

As such the planning phase of the RR should also consider data management For example, what type of data and documents will be collected, how they will be stored (e.g., file formats, folder structures, naming conventions, software and platforms used), and what is required by the reporting standards (e.g., PRISMA-5, PRISMA-S).
Reference management software (e.g., EndNote, Zotero) should be used to manage large search results. This typically includes importing result sets from various databases and keeping track of the flow of citations (e.g., deduplication, screening). It is important to note that record management for database searches is less time consuming than for many types of supplementary or grey literature searches.\(^{77}\) Bibliographic databases usually allow bulk export of search results to the reference management software and provide detailed search histories for documentation of search strategies. In contrast, supplementary search methods often require manual input of each individual source into the reference management software and manual documentation of search process. This means the time required to manage the search results from different sources is also a consideration for the overall RR time plan.

Additionally, software platforms for systematic review production (e.g., Covidence\(^{78}\), EppiReviewer\(^{69}\)) can provide a unified way to keep track of records throughout the whole review process, which can improve management and save time. However, the time and cost investment necessary to use these tools has to be considered.\(^{79}\) These platforms are most likely useful for teams that regularly produce reviews.

5.3 Deduplication of search results

Because search results from various information sources are likely to overlap, special attention needs to be paid to the deduplication process.

Reference management software such as Endnote usually requires a stepwise deduplication process.\(^{80}\) Some duplicates can be automatically identified and discarded by the software (e.g., records where the elements author, title, publication date, journal are identical). Other duplicates can be flagged by the software but need to be manually reviewed (e.g., a record where most elements are identical, but perhaps the formatting of the author lists varies giving the records an appearance of being different). Further fully manual deduplication may also be required. This stepwise deduplication process requires time and expertise but gives more control over the deduplication results.

On the other hand, some platforms (e.g., Covidence, SR-Accelerator\(^{81}\), Deduklick\(^{82}\)) offer automated deduplication, but information about the deduplication methods (i.e., What data elements are compared?) and evaluations of their accuracy are only available in some cases.

When assessing what deduplication approaches are most efficient in a RR process, it is important to consider both time savings by deduplicating the results more quickly (e.g., through automation) and having a more accurate result of the deduplication process (i.e., missing as few duplicates as possible).

A high rate of false negatives (i.e., duplicate citations that should have been deleted but were not) will increase screening time. In contrast, false positives (i.e., citations that were incorrectly deleted as duplicates) may lower the sensitivity of the search result because relevant references could be lost. When considering the use of fully automated deduplication, time should be invested to assess the performance of the tool against the deduplication method commonly used by the team.

The review team also needs to decide what is considered a duplicate in the review context. The following are not considered true duplicates in the bibliographic sense but could be regarded as irrelevant for a RR:

- the preprint of an article where a later journal publication is available;
- different publications of the same conference abstract;
- re-prints of a research article;
older versions of an SR where a recent update is available.

References


6. Metzendorf M-I. Why medical information specialists should routinely form part of teams producing high quality systematic reviews--a Cochrane perspective. 2016


---


68. Thomas J, McDonald S, Noel-Storr A, et al. Machine learning reduced workload with minimal risk of missing studies: development and evaluation of a randomized controlled trial classifier for


