Research Synthesis Methods (2026), **17**: 1026–1044
doi:10.1017/rsm.2026.10084

Research Synthesis Methods

SOFTWARE FOCUS

CiteSource: An R package for data-driven search strategy development and enhanced evidence synthesis reporting

Trevor Riley[1], Sarah Young[2], Avery Paxton[3], Lukas Wallrich[4], Kaitlyn Hair[5] and Matthew Grainger[6]

[1]National Oceanic and Atmospheric Administration Central Library, USA
[2]University Libraries, Carnegie Mellon University, USA
[3]Southeast Fisheries Science Center, National Marine Fisheries Service, National, USA
[4]Birkbeck, University of London, Birkbeck Business School, UK
[5]EPPI Centre, UCL Social Research Institute, University College London, UK
[6]Knowledge Synthesis, Norwegian Institute for Nature Research, Norway

Corresponding author: Sarah Young; Email: sarahy@andrew.cmu.edu

Received: 30 June 2025; **Revised:** 12 December 2025; **Accepted:** 31 January 2026; published online 6 April 2026

Keywords: evidence synthesis; information retrieval; reproducibility; search strategy; systematic searching

Abstract

Evidence synthesis findings hinge upon well-designed, effective search strategies. When developing these strategies, evidence synthesis teams make multiple decisions (e.g., selecting information sources, developing search string architecture, and picking supplementary search methods) that directly affect the breadth of discovered evidence and thus evidence synthesis outcomes. Despite the number of decisions required when developing search strategies, limited guidance exists to inform these decisions using a data-driven approach. To help address this gap, we developed CiteSource, an R package and accompanying Shiny application, that supports data-driven search strategy development and reporting. CiteSource allows users to assign and retain metadata across three custom fields: *source, label*, and *string* to indicate where the records were found, what method or string was used to find them, and whether they were included after screening. CiteSource allows users to visually map the overlap between sets of records, create data summaries of citation records, and export citation records with the newly assigned metadata. CiteSource's analysis and visualization outputs can be harnessed for a variety of use cases, such as optimizing literature source selection, honing and understanding the effectiveness of search strings, and evaluating the impacts of literature sources and supplementary search methods. Overall, CiteSource provides a tool for evidence synthesizers to make informed data-driven decisions that boost the efficiency, rigor, and transparency of search strategies and associated reporting.

Highlights

What is already known?

- Search strategies in evidence synthesis significantly influence what evidence is considered during screening and what evidence is included in the final synthesis.
- Search strategies largely rely on the researcher's knowledge/experience and past analysis of information sources and methodologies.

This article was awarded Open Data and Open Materials badges for transparent practices. See the Data availability statement for details.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Research Synthesis Methodology. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

- Optimizing search strings and strategies is time-intensive due to constant changes in the content and functionality of information sources.

What is new?

- The CiteSource R package and Shiny app assist users in making data-driven decisions to improve search strategies.
- By tracking the provenance of citation data, users can analyze the overlap in citations across information sources and search methods.
- Underlying data can be viewed through various plots and tables, which can speed up the iterative process of validating search strategies.

Potential impact for Research Synthesis Methods readers

- Researchers can quickly assess the potential impact of information sources and the effectiveness of search strings to optimize strategies based on data.
- Search strategies and string development can be systematically developed and reported to increase the quality and transparency of syntheses.
- Reporting the impact of information sources and search methods can assist in developing further methodological guidance for search strategy development across synthesis methods.

1. Introduction

Evidence syntheses, including systematic maps and reviews, require well-designed and effective search strategies to optimize the discovery of relevant evidence. Although a thoughtful and well-designed search strategy is only one of many components of evidence synthesis (e.g., question formulation, literature screening, data extraction, and data synthesis), it can have an outsized impact on evidence synthesis quality.[1,2] This is because search strategies are instrumental in governing the scope and extent of the evidence base that will be discovered.[3] Optimized search strategies capture relevant articles that match the information needed to address the research questions and aim to balance precision and sensitivity to capture as many potentially relevant articles while maintaining constraints that limit discovery to what is practical and feasible.[4–6] This balance, however, is often elusive, as there is a slippery slope between achieving an optimized search strategy versus one that is more 'comprehensive' but may return a high number of irrelevant results that would be impractical and unnecessary for a high-quality synthesis.[7,8] Furthermore, the complexity of searching can lead to errors in accurately applying search strategies.[9,10]

When optimizing search strategies, evidence synthesis teams must make multiple decisions that impact the reliability and validity of synthesis findings.[11–13] These include decisions related to selecting information sources (e.g., academic databases, gray literature sources, platform indexes, clinical trial registries, organizational websites, and institutional repositories), determining search string architecture (e.g., selection of terms, use of controlled vocabularies, application of Boolean and proximity operators, truncation, and lemmatization search field selection), and considering which 'supplementary' methods to utilize (e.g., citation searching, application of artificial intelligence (AI) platforms, open web searching, contacting stakeholders and subject experts, co-citation searching, and hand searching). A growing body of literature demonstrates that such decisions can impact evidence synthesis findings.[14–17] One study of systematic reviews estimated that 60% of published systematic reviews did not retrieve 95% of relevant articles, largely because the evidence synthesis teams did not search the appropriate combinations of databases, which directly impacts synthesis findings.[18] Additionally, while supplementary search methods have been shown to yield studies of sufficient quality and relevance not discovered in bibliographic databases,[19–21] the resource-intensive nature of employing various methods can lead to differing opinions as to what methods are justified.[22]

Search strategies are informed primarily by the type of synthesis being conducted, the research question(s) and scope, and the intended use of the synthesis. Strategies are also shaped by external

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

factors, most importantly, time limitations and the availability of resources. Beyond these logistical constraints, the experience, knowledge, and expertise of the research team also play a significant role in shaping search strategies.[23] While it has been shown that including information specialists in the synthesis process can improve the quality due to their specialized knowledge of information sources and search techniques,[24–26] the information landscape remains challenging. Tacit knowledge can quickly become outdated due to shifts in indexing, industry mergers, and uncertainty around the future funding of government-supported databases and repositories,[27–31] To optimize search strategies and make informed methodological decisions as the landscape changes, information professionals and researchers alike require current and accurate data, specific to their research questions.

Despite the number and impact of decisions required in developing a search strategy, limited guidance exists for how to make these decisions using a data-driven approach. Conducting guidelines that provide detailed information on how to perform a variety of evidence synthesis steps, such as screening and quality assessment, with transparency and minimal bias, typically lack detailed guidance on search strategy development.[5,6] For example, conducting guidelines contain no prescribed way to determine the number of literature sources that should be searched or how to demonstrate that a search strategy has been optimized.[32,33] And, while valuable guidance does exist, including on general strategies,[34] which bibliographic databases may yield the most relevant articles,[35] and which citation indices are most appropriate for citation chasing,[36] such guidance is typically discipline-specific and can quickly become outdated as the information landscape shifts and new technology, resources, tools, and methods are developed.[37–40]

To help address the gap in quantitative search strategy guidance, we developed CiteSource, an R package and accompanying Shiny application, that supports data-driven search strategy development and reporting. Our goal was for CiteSource to serve as a tool for ensuring effective search methods, increasing the efficiency of search string development and testing, improving search strategy optimization, boosting search strategy transparency and reproducibility, and conducting analyses to assess outcomes of search strategy decisions. Overall, CiteSource provides real-time analysis functions, allowing evidence synthesizers to make data-informed decisions that boost the efficiency, rigor, and transparency of search strategies and associated reporting, which in turn facilitates comprehensive documentation consistent with established reporting standards (e.g., PRISMA-S and ROSES).

CiteSource adds to the growing suite of R packages designed to support search strategy development for evidence synthesis. For example, litsearchr is an R package and Shiny app that uses keyword co-occurrence networks to identify relevant search terms and phrases from a corpus of initial search results, aiding the term harvesting step of search strategy development.[38] It also provides an automated approach to the construction of a Boolean search strategy from a list of key terms and a tool to compare a set of search results to a list of known benchmark articles. searchbuildR is another, more recent example of a search strategy construction tool.[39] searchbuildR uses term frequency analysis to suggest keywords, including a view of their context within the abstracts, and PubMed identifiers to generate lists of medical subject headings (MeSH) terms. Users can construct a random set of records for search strategy development and a set for testing the search strategy, and can also test the consequences of widening or narrowing the distance between keywords when using proximity operators. While such tools support search term harvesting and search strategy construction, CiteSource is the first R package designed to test full search strategies across different sources and methods.

This paper introduces the CiteSource R package and accompanying Shiny application. We begin by outlining CiteSource's core functionalities and the features designed to empower researchers in their search strategy development. Subsequently, we present practical use cases to demonstrate CiteSource's utility in optimizing information source selection, refining search strings, and assessing the impact of various search methods. The paper concludes with a discussion of CiteSource's broader implications for enhancing evidence synthesis methodology, promoting reproducibility, and supporting the training of information specialists and researche rs (Box 1).

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

Box 1. Definitions of terms used throughout the paper are provided in Box 1.

Search strategy: The overall plan and application of search methods designed to identify and retrieve relevant evidence for an evidence synthesis project. Strategy encompasses the methods and techniques used to identify and retrieve relevant evidence.
Search method: An approach used to discover relevant records. This includes, but is not limited to, searching bibliographic databases/platforms, citation searching, use of AI platforms, open web searching, and contacting experts.

Search technique: A specific style/process applied within a search method (e.g., Searching title/abstract vs. title/abstract/keywords in bibliographic databases).

Search string: The combination of keywords, controlled vocabulary, Boolean operators, proximity operators, truncation/wildcards, and search fields, and any applied filters to query an information source (typically a bibliographic database/platform).

Syntax: The set of rules, commands, operators, symbols, and structure used to formulate a query that can be correctly interpreted by a particular search engine, database, or information source.

Record: A bibliographic entry, which represents a single citable work (e.g., a journal article, a book, a book chapter, a report, etc.).

Set of records: A strictly defined group of records (e.g., results from Web of Science, results from a variation of a string, and records included after screening). In CiteSource, the source, label, and string fields can be used to define sets of records.

Unique record: A record that has been found to have no duplication across all records being analyzed.

Duplicate record: A record that, when compared to other records, has been found to point to the same citable work.

Distinct record: A record that has been identified as unique within a set of records (e.g., when two search strings from one database are combined and deduplicated). Distinct Records may be found to be unique or duplicates when compared to a larger set of records (e.g., when comparing search results from Scopus and Web of Science).

2. Functionality

CiteSource is available as a downloadable R package, as well as a Shiny application. Here, we take a stepwise approach to describe the functionality of CiteSource using the Shiny Application. Detailed CiteSource vignettes that walk through use cases are listed under the 'articles' tab on the CiteSource website (https://www.eshackathon.org/CiteSource/)

1) **Upload citation files:** Users upload multiple citation files (.ris, .txt, or .bib) containing search results or records of included articles. The package also provides an option to re-upload .ris and .csv data that were previously exported from CiteSource, allowing users to reproduce all plots and tables with a single file.
2) **Assign custom metadata fields:** After uploading citation files, users assign custom metadata across three fields: 'cite source', 'cite label', and 'cite string' referred to in this paper and in the shiny app as *source*, *label*, and *string*.
   - ***Source*:** used to tag citations with metadata related to their source (e.g., Web of Science, and Scopus) or the method used to find them (e.g., citation searching, numbered search string). *Source* is a free-text field.
   - ***Label*:** used to tag citations with information related to their associated screening phase. The label field requires one of three terms: 'search', 'screened', or 'final'. All plots/tables require at least one file to be labeled as 'search', no other terms in the label field are permitted.
     - **Search**: indicates the citations are associated with initial search results (including benchmark articles).

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

Table 1. *CiteSource can produce multiple visualizations to help analyze search strategy outcomes.*

Plots	Function
Heatmaps	Visualize the number and percent of overlap between two sources, methods, or techniques
Upset plot	Visualize the number of unique and overlapping records across multiple sources, methods, or techniques
Bar chart	Visualize the number of unique and overlapping records for each source, method, or technique across screening phases

Table 2. *CiteSource can produce a number of tables to help analyze and report search strategy outcomes.*

Tables	Function
Record summary table	Displays the number of unique/overlapping records, percent of unique records contributed, and percent of records unique to each source, method, or technique
Precision sensitivity table	Overview of the number of records that were included across screening phases for each source/method. Includes each source/method precision and sensitivity/recall calculation
Record level table	Interactive table that allows users to search full citations, sort based on source/method association, and link directly to records via DOI or URL

    - **Screened**: indicates that the citations were included after title/abstract screening.
    - **Final**: indicates that the citations were included after being screened at the full-text level and included in the final synthesis.
  - ***String*****:** used to further differentiate sets of records. While the source/label fields alone can handle most use cases, the string field can be used to record other supplementary information a user may want to retain for analysis, or to further differentiate string variations (e.g., String1-narrow, String1-broad, String2-narrow, etc.) *String* is a free-text field.
3) **Deduplicate citations:** After tagging uploaded datasets, users must initialize the automated deduplication process. At the core of its functionality, CiteSource utilizes the Automated Systematic Search Deduplicator (ASySD) package to identify duplicates and automate the deduplication process.[41] ASySD identifies duplicates by running multiple rounds of comparisons to find potential matches based on shared metadata. Once potential pairs are identified, Jaro–Winkler similarity scores are calculated for each field and used together in over twenty combinations to classify pairs as 'true duplicates' when they meet specific thresholds (e.g., pages > 0.8 + volume > 0.8 + abstract > 0.9 + author > 0.5 + Journal > 0.6). Once automated deduplication is complete, users have the option to manually review records that were flagged due to their similarity, but which did not meet the threshold for automatic classification as duplicates.
4) **Analyze data:** Once deduplication is complete, users can generate a number of plots and tables to analyze the data (Tables 1 and 2). For a variety of use cases, examples of these are covered in the section on use cases.
5) **Export data:** Finally, users can export record metadata as .csv, .ris, and .bib files. Custom metadata is retained as part of the exported data. Exported data can later be reimported into CiteSource, these data can also be made available when reporting synthesis findings for added transparency and to provide an opportunity for aggregated analysis when combined with data from other syntheses.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

3. Use cases

CiteSource has a number of use cases that generally fall under two areas: optimizing search strategies and analyzing impact. Optimizing search strategies can include source selection and string development, as well as some aspects of search method selection, such as whether or not to include certain 'supplementary' methods like citation searching. Using the built-in visualizations and tables, typically iterative multi-step processes can be compressed into a single analysis step, saving users time in testing and validating strategies. For example, differences in search results from changes in the structure of Boolean operators, the use of stemming, lemmatization, near-term operators, or the use of different search fields can be compared and contrasted in a single plot. When using CiteSource to optimize a search, the package is applied during protocol development. Conversely, when used to determine the impact of a specific information source or method, data are analyzed only after the final inclusion/exclusion decisions are made, as it is necessary to see the overlap with the included records. By providing researchers with an understanding of the contribution of sources and methods to a given review, this type of analysis and reporting can provide much-needed data for evidence-based methodological approaches in updates and future syntheses on related topics.

3.1. Search strategy optimization

Search strategies are informed primarily by the type of synthesis being conducted, the research question(s), and the intended use of the synthesis. Strategies are also shaped by key external factors, such as time constraints, the availability of information resources and screening tools, as well as the experience, knowledge, and expertise of the research team.[23,33] While librarians and information specialists typically provide guidance on source selection and best practices for searching, the information landscape constantly changes with shifts in publishing and indexing practices, merging of information aggregators and publishers, and uncertainty around federally sponsored databases and repositories; all pose significant challenges for search strategy development.[27–31]

3.1.1. #1 Information source selection

A primary application of CiteSource involves facilitating data-driven information source selection within the review process. Through the systematic tagging of search results with 'source' metadata, CiteSource enables researchers to determine the degree of citation overlap and uniqueness among information sources. This analysis illustrates which sources have the greatest potential for review versus those that may be considered excessive or potentially insignificant due to substantial duplication. We broadly consider 'Information sources' to encompass contributions from distinct methodologies, such as citation chasing, in effect, treating a single paper as a source of information. Consequently, by integrating the assessment of both information sources and methodological approaches, researchers can attain a holistic visualization of the collective performance of their comprehensive search strategy.

As an example of how CiteSource can be applied to guide the selection of information sources, we can consider the following scenario. A research team is investigating the ecological consequences of acoustical disturbances originating from pile driving activities in marine settings. The selection of appropriate information repositories may require the inclusion of discipline-specific databases, for instance, ProQuest's Aquatic Sciences and Fisheries Abstracts (ASFA) and Oceanic Abstracts (OA), as well as multidisciplinary platforms, including Clarivate's Web of Science (specifically, the Science Citation Index Expanded—SCIE), Elsevier's Scopus, and Digital Science's Dimensions. Furthermore, the team believes that academic theses and dissertations may be valuable and decides to include specialized databases such as ProQuest's Open Access Theses and Dissertations (OATD). To initially evaluate the potential coverage and overlap of these sources, a relatively broad 'naive' search string is developed to capture a wide array of potentially relevant records. This string (Figure 1; Web of Science syntax) was subsequently adapted and applied across the six information sources. The results were then uploaded to CiteSource for comparative analysis, with each set of records distinguished by their assigned 'source' metadata.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

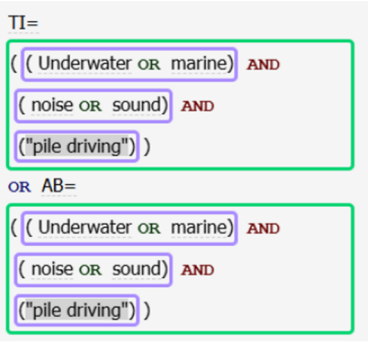


Figure 1. *Example Search String is a simpler approach combining the three PEO concepts within each search field. This figure shows strings in Web of Science syntax (visualized in https://www.medsyntax.org).*[42] *String syntax was adapted to meet the syntax requirements for each database. Full strings can be found in the supplementary file.*

Once automated deduplication is complete, the team can quickly assess the overlap between each source. In this example, the heatmap (Figure 2) clearly shows us that all 77 records that were found in Oceanic Abstracts (OA) were also found in Aquatic Sciences and Fisheries Abstracts (ASFA). This overlap is explained by the fact that Proquest's Oceanic Abstracts is included within the ASFA database while maintaining its own scope and the ability to search the database independently. While this overlap was included by design in this example, it provides an example of how CiteSource can be used to identify sources that may not provide unique content. In this case, the team (who may or may not have been knowledgeable about the ASFA/OA relationship) can use these data to make a data-informed decision to drop Oceanic Abstracts as a source. Simultaneously, recognizing a highly unique source like OATD might prompt consideration of exploring additional specialized repositories for theses and dissertations to ensure comprehensive coverage of this document type.

This example highlights CiteSource's capacity to support data-driven information source selection, as illustrated by the potential findings of redundancy between ASFA and Oceanic Abstracts and the novel records contributed by OATD. CiteSource facilitates the strategic inclusion of sources offering unique value and the exclusion of those providing largely redundant content, thereby optimizing the efficiency and effectiveness of evidence-gathering efforts from the outset.

3.1.2. #2 Search string development

Search strings can range in complexity based on several factors, including the research topic, scope, database/platforms being searched, etc. Testing and iterating search strings can be time-consuming, due to differences in syntax across platforms and the need to compare the results of various techniques. CiteSource assists users in systematically evaluating and refining search strings. This process involves not only optimizing retrieval to enhance sensitivity and precision but also identifying and rectifying potential critical errors in syntax or logic,[9] all while speeding up the iterative analysis. By enabling the distinct tagging of search result sets (e.g., by string version or database) and the subsequent analysis of overlapping and unique records, users can quickly assess multiple iterations of a search string. Examples of such assessments include evaluating:

- the impact of including or excluding terms
- variations in syntax
- different combinations of strings
- different approaches to Boolean architecture
- the use of different search fields
- the effectiveness of capturing benchmark articles

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

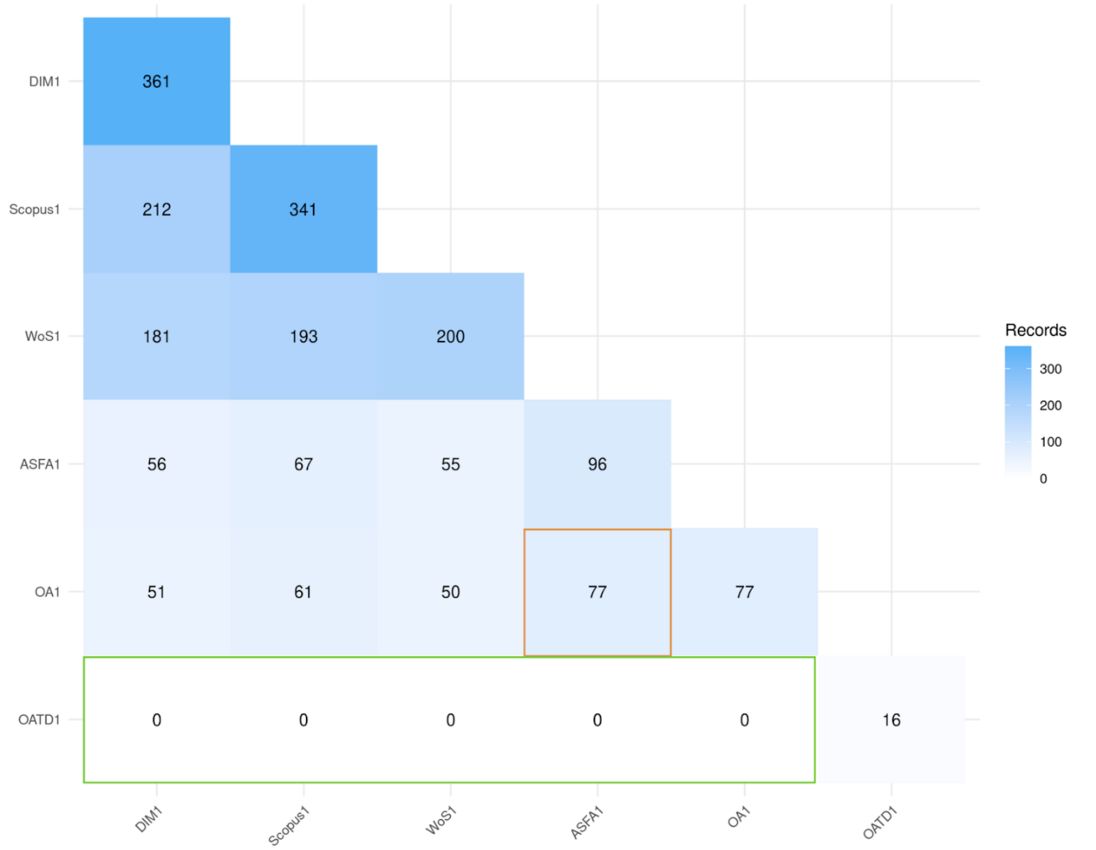


Figure 2. *A heatmap showing the six sets of records uploaded and the number of overlapping records between each set. In this example, all 77 records from Ocean Abstracts (OA) are found in Aquatic Sciences and Fisheries Abstracts (ASFA), highlighted here in orange. The plot also shows that none of the 16 records from ProQuest's Open Access Theses and Dissertations (OATD) were found in any of the other sources, highlighted here in green.*

To demonstrate CiteSource's application in search string development and optimization, this example examines different approaches to Boolean architecture for a simple search string. Due to differences in the way database platforms function, it is often useful to compare search techniques, including variations in Boolean nesting (i.e., the construction of parentheticals), metadata field selection, and verifying correct use of syntax. For example, some database platforms do not allow users to combine the results of multiple strings, nest Boolean logic, or combine field codes to search across multiple fields at once; consequently, simplified search strategies are often required to accommodate these limitations. In this example, our team is once again examining the impact of pile driving on marine life, utilizing a search string that represents three concepts: marine/underwater environment (population), pile driving (exposure), and noise/sound (outcome). Two search strings (Figure 3) were developed with their Boolean architecture intentionally designed to yield different results, illustrating CiteSource's capability to facilitate analysis of such variations. The first version, used in the prior scenario, is designed for broad compatibility across most platforms, reflecting a common necessity due to database constraints. The second string involves nested Boolean logic, representing a more robust and strategic approach that can be used in databases like Web of Science. Here, Web of Science, Dimensions, and Scopus have been selected to demonstrate how CiteSource can compare search results across different string approaches and database platforms; while observing string variations might not always necessitate

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

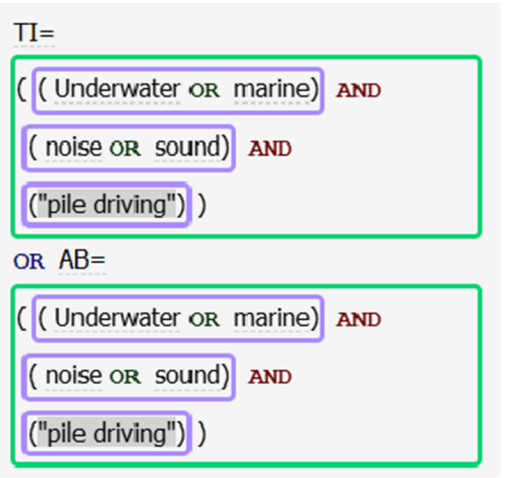
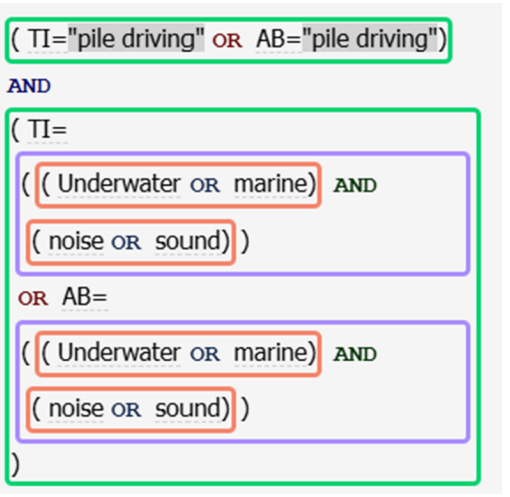
String Version #1	String Version #2
	

Figure 3. *Two different search strings using different approaches to Boolean architecture are shown. String Version #1 is a simpler approach combining the three PEO concepts within each search field. String Version #2 shows a more complex approach with nested Boolean phrases for a more comprehensive search. This figure shows strings in Web of Science syntax (visualized in https://www.medsyntax.org). String syntax was adapted to meet the syntax requirements for each database. Full strings can be found in the supplementary file.*

multiple platforms, applying this method here serves to further showcase CiteSource's capabilities in cross-database analysis.

After importing the search result files for each string version and database, metadata tags are assigned within CiteSource to differentiate each distinct set of results (Figure 4). CiteSource streamlines tagging and helps reduce errors by optionally auto-filling the 'source' field based on the file name; it also defaults the 'label' field to 'search', the most common initial designation. Both the source and label fields can be easily edited as necessary. Once citation files are tagged, we perform automated and manual deduplication (Figure 5). To efficiently resolve these flagged pairs, CiteSource's 'Individual Record Table' (Figure 6) is particularly useful, as it displays comprehensive bibliographic details for every record, and critically, provides direct links (e.g., via DOI or URL) to the original source documents when available. By examining the manual deduplication table in combination with the full details, users can quickly access the source articles and accurately determine if the flagged items are true duplicates, even in cases of inconsistent or incomplete metadata.

CiteSource utilizes the UpsetR package[43] to visualize overlap across sources (Figure 7). This plot offers a powerful visualization for comparing the retrieval of different search string versions across multiple databases. Each row represents an information source, defined by the user with the 'source' field. Columns show the number of records that overlap across these sources. In this example, examining the intersections reveals that String Version #2 identified 17 records not captured by String Version #1(blue boxes indicate records found exclusively by String Version #2). The Detailed Record Table (Figure 8) is another way users can evaluate this overlap. The table shows the breakdown of each source, its overall record contribution, and the percentage of unique record contributions. In this case, there were only 11 unique records, 5 each from Dimensions and Scopus, and a single unique record from Web of Science.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

file.name	records	source	label
DIM1.ris	375	DIM1	search
DIM2.ris	383	DIM2	search
Scopus1.ris	342	Scopus1	search
Scopus2.ris	354	Scopus2	search
WoS1.ris	201	WoS1	search
WoS2.ris	208	WoS2	search

Figure 4. *After uploading files, users have the ability to tag each accordingly. In this case, six uploaded .ris files are tagged according to the information source and string. For efficiency, CiteSource auto-fills the source field according to the file name and automatically adds 'search' for each label, as this is the most common tag.*

	author1	author2	title1	title2	year1	year2	journal1	journal2	abstract1	abstract2	doi1	doi2	pages1	pages2	source1	source2
1	Reyff, James	Reyff, J.	Underwater Sounds From Un...	Underwater sounds from un...	2012	2012	The Effects of Noise on A...	Advances in experimental ...	In 2009, the California D...			10.1007/978-1-4419-7311-5...	439	439	DIM1	Scopus1
2	MacGillivray, Alexander	MacGillivray, A.O.	A model for underwater so...	A model for underwater so...	2013	2015		Proceedings of Meetings o...	Marine impact pile drivin...	Marine impact pile drivin...	10.1121/2.0000030	10.1121/2.0000030	045008		DIM1	Scopus1
3	Von Pein, J. and Lippert,...	von Pein, J. and Lippert,...	Scaling offshore pile dri...	Scaling offshore pile dri...	2022	2022	Proceedings of the Instit...	Proceedings of the Intern...	For the installation of o...	Underwater noise due to c...	10.1121/2.0001622				Scopus1	Scopus1
4	Lee, K.M. and Wochner, M....	Lee, K.M. and Wochner, M....	Mitigation of low-frequen...	Mitigation of low-frequen...	2012	2012	Proceedings of Meetings o...	11th European Conference ...	Collections of bubbles ca...	Large air cavities with e...	10.1121/1.4767960			1134	Scopus1	Scopus1
5	Weizhengrong, Zhou and Xi...	Zhou, WZR and Xu, XM	Preliminary Exploration f...	Preliminary Exploration f...	2016	2016		2016 IEEE/OES	As one of the most	As one of the most	10.1109/coa.2016.7535775		1		DIM1	WoS1

Figure 5. *Manual deduplication table. The three rows highlighted in blue have been selected as duplicates. Columns show metadata for each record identified as a potential duplicate. Green boxes are added here to show records in pairs #3 and #4 are separate conference papers. Red boxes are added here to highlight incorrect metadata.*

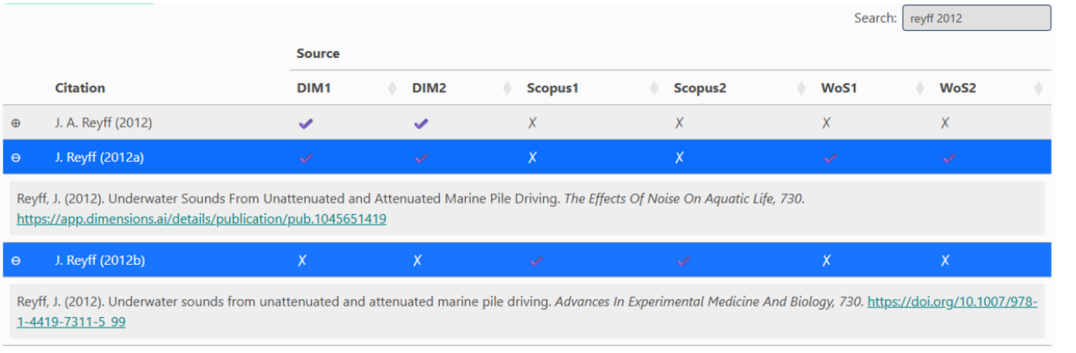


Figure 6. *Individual record table is shown with two records that are identified in pair #1 from Figure 5. Used in conjunction with the manual deduplication table, records can be quickly checked by linking directly to the item. In this case, record 'Reyff 2012a' is correct, while metadata from 2012b incorrectly uses the book series title.*

3.1.3. #3 Benchmark testing

One way to understand a search string's effectiveness is through benchmark testing. By starting with a set of articles that meet the stated inclusion criteria, researchers can gauge the sensitivity and effectiveness of search strings, identify potential changes to a search string to improve sensitivity, and

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

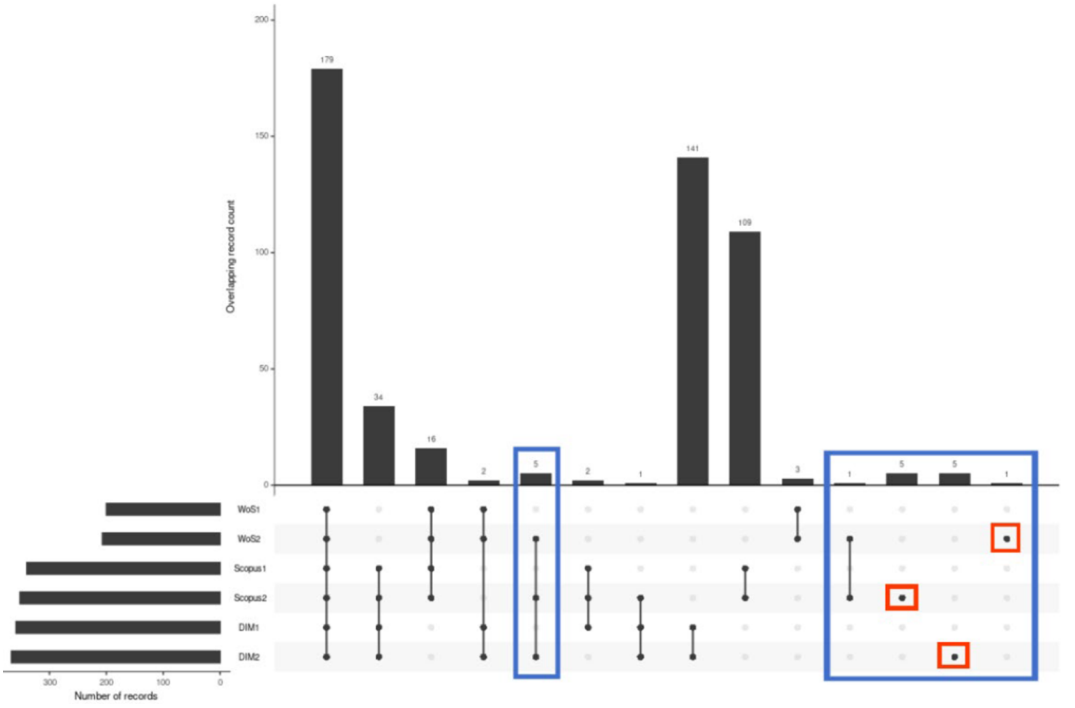


Figure 7. *The upset plot visualization can be used to assess the number of overlapping and unique records across each search string version and across databases (WoS=Web of Science; DM=Dimensions). This upset plot shows 17 additional records being found using Search String Version #2; blue boxes are added to highlight these records. Of the 17 additional records, 11 are unique to one of the three sources; red boxes highlight these records.*

Record Summary

	Records Imported[1]	Distinct Records[2]	Unique Records[3]	Non-unique Records[4]	Source Contribution %[5]	Source Unique Contribution %[6]	Source Unique %[7]
DIM1	375	359	0	359	19.7%	0.0%	0.0%
DIM2	383	367	5	362	20.1%	45.5%	1.4%
Scopus1	342	340	0	340	18.6%	0.0%	0.0%
Scopus2	354	352	5	347	19.3%	45.5%	1.4%
WoS1	201	200	0	200	11.0%	0.0%	0.0%
WoS2	208	207	1	206	11.3%	9.1%	0.5%
Total	1863	[8] 504	11	1814	100.0%	100.0%	2.2%

[1] Number of raw records imported from each database.

[2] Number of records after internal source deduplication.

[3] Number of records not found in another source.

[4] Number of records found in at least one other source.

[5] Percent distinct records contributed to the total number of distinct records.

[6] Percent of unique records contributed to the total unique records.

[7] Percentage of records that were unique from each source.

[8] Total citations discovered (after internal and cross-source deduplication).

Figure 8. *The Detailed Record Table provides information on the number of imported records, distinct records, unique records, duplicate records, the percentage of records contributed by each source, the percentage of unique records contributed by each source, and the percentage of each source that was unique.*

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

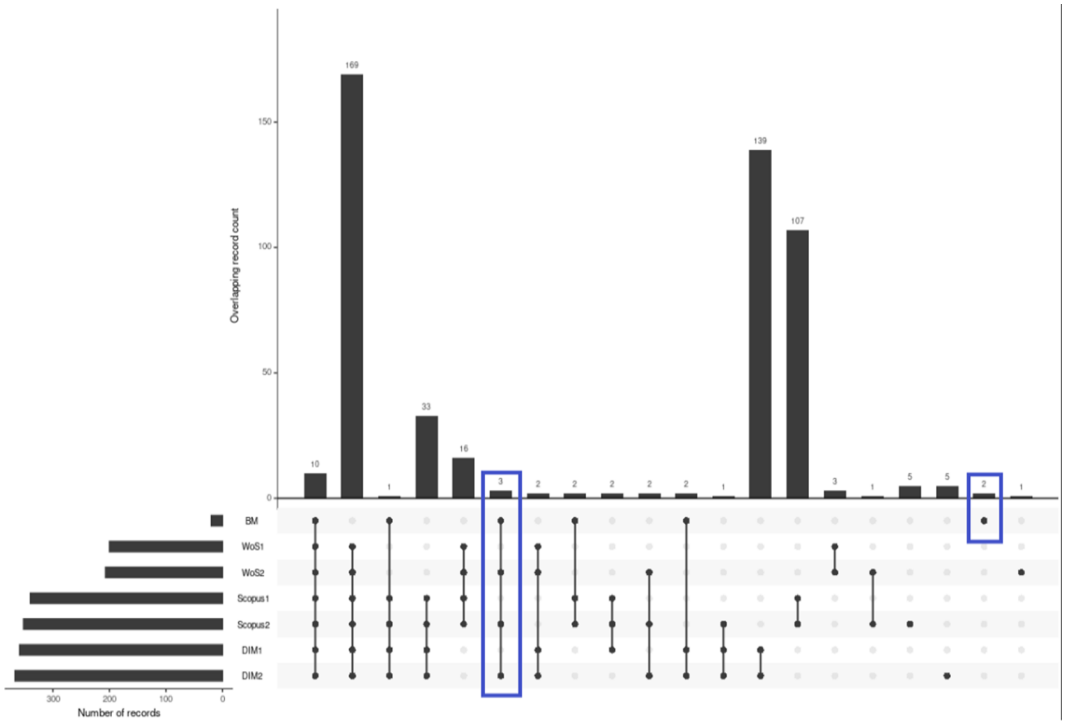


Figure 9. *This upset plot shows an additional three benchmarking articles being found in Version #2 and two benchmarking articles that are not found. Blue boxes are added to highlight these records.*

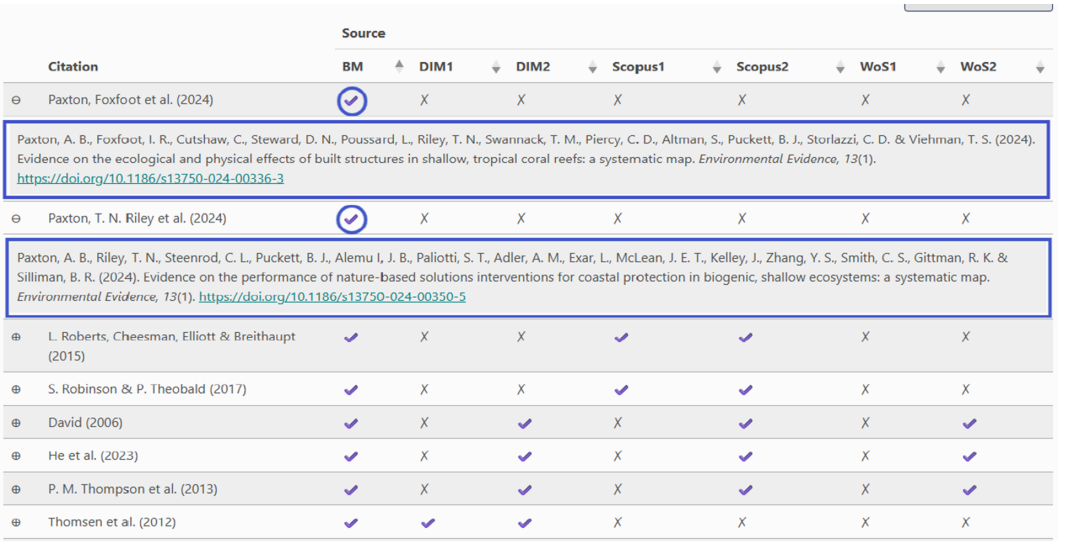
understand the limitations of the search.[44] These benchmarking functionalities are particularly useful in fields where certain key studies are considered foundational. Ensuring these foundational studies are captured by search strategies is critical for the credibility and completeness of a review. Here, we demonstrate how CiteSource can assist researchers when benchmark testing the results of searches.

Continuing with the example above, we have hypothetically selected 20 articles as a benchmark set (designated 'BM') based on prior knowledge of the field and examining previous related reviews. In this case, the Upset plot (Figure 9 – similar to Figure 7 but now including the benchmark set as a distinct set) shows us the performance of our search strategies against these key articles. The plot reveals that ten of the 20 benchmarking articles (designated BM) were captured across the three databases by both String Version #1 and String Version #2. Looking at the remaining overlap, we can see that String Version #2 found three additional benchmark articles that String Version #1 missed, and that two benchmarking articles were not found in any database with either string. This information, used in combination with the record-level table (Figure 10), can be used to quickly review why these benchmark articles were missed, whether due to missing/incorrect terms/syntax, or whether these records were not indexed in the information sources.

3.2. Impact assessment

Transparent reporting of methods and results is a fundamental requirement once an evidence synthesis is complete, mandated by reporting guidelines and checklists.[45–47] This includes details on search methods, information sources consulted, specific search strings utilized, and the flow of records through screening (often visualized in a PRISMA diagram). Beyond its utility in the iterative optimization of search strategies, CiteSource enables researchers to retrospectively analyze and quantify the contributions of different information sources and search methods to the resulting body of evidence.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press



	Citation	Source: BM	DIM1	DIM2	Scopus1	Scopus2	WoS1	WoS2
⊖	Paxton, Foxfoot et al. (2024)	✓	✗	✗	✗	✗	✗	✗
	Paxton, A. B., Foxfoot, I. R., Cutshaw, C., Steward, D. N., Poussard, L., Riley, T. N., Swannack, T. M., Piercy, C. D., Altman, S., Puckett, B. J., Storlazzi, C. D. & Viehman, T. S. (2024). Evidence on the ecological and physical effects of built structures in shallow, tropical coral reefs: a systematic map. *Environmental Evidence, 13*(1). https://doi.org/10.1186/s13750-024-00336-3							
⊖	Paxton, T. N. Riley et al. (2024)	✓	✗	✗	✗	✗	✗	✗
	Paxton, A. B., Riley, T. N., Steenrod, C. L., Puckett, B. J., Alemu I, J. B., Paliotti, S. T., Adler, A. M., Exar, L., McLean, J. E. T., Kelley, J., Zhang, Y. S., Smith, C. S., Gittman, R. K. & Silliman, B. R. (2024). Evidence on the performance of nature-based solutions interventions for coastal protection in biogenic, shallow ecosystems: a systematic map. *Environmental Evidence, 13*(1). https://doi.org/10.1186/s13750-024-00350-5							
⊕	L. Roberts, Cheesman, Elliott & Breithaupt (2015)	✓	✗	✗	✓	✓	✗	✗
⊕	S. Robinson & P. Theobald (2017)	✓	✗	✗	✓	✓	✗	✗
⊕	David (2006)	✓	✗	✓	✗	✓	✗	✓
⊕	He et al. (2023)	✓	✗	✓	✗	✓	✗	✓
⊕	P. M. Thompson et al. (2013)	✓	✗	✓	✗	✓	✗	✓
⊕	Thomsen et al. (2012)	✓	✓	✓	✗	✗	✗	✗

Figure 10. *The individual record table facilitates the quick review of findings. In this example, we see the two benchmarking articles that were not found across the three sources using both string versions.*

With minimal effort, this impact assessment analysis can increase transparency, empower future researchers to build upon documented effective practices, and contribute to refining methodological guidance within the evidence synthesis community.

3.2.1. #4 Analyzing the impact of information sources & methods

To illustrate how CiteSource can be used to analyze the impact of various sources and methods, we can return to the research team investigating acoustical disturbances from pile driving. Once the team completed their searches using their optimized string across Web of Science, Scopus, and Dimensions, they supplemented their search with a call for papers from experts in the field, as well as by applying forward and backward citation chasing on two seminal review articles. At title/abstract screening, the team includes 100 articles, and after full-text review, the resulting final set contains 55 articles. To analyze the impact of each source and method, the team can now upload all sets of records and tag them accordingly using the source and label fields (Figure 11). After automated and manual deduplication is complete, the team can generate the Phase Analysis plot (Figure 12), which provides a visual breakdown of each source's unique/overlapping record count as well as a count of these records over the screening phases. Looking at the plot, it's easy to see the performance of each information source and method. Citation searching, for example, while initially providing a large number of unique articles, contributed no unique records after title and abstract screening. In fact, this method was by far the least precise, as only five of the 192 distinct records it produced were included in the final synthesis. The plot also shows that while the call for papers yielded a small number of results, this method of discovery contributed the largest number of unique records to the final synthesis. Finally, we can see that both Scopus and Dimensions outperformed Web of Science in the number of unique records that they contributed to the final synthesis. The data are further detailed in the Precision/Sensitivity Table (Figure 13), providing a full picture of performance. Looking at the table, we can see that despite the lower number of unique articles included in the final synthesis, Web of Science had a slightly higher precision.

In practical terms, this impact assessment functionality can lead to more strategic decision-making in resource selection and method application. This quantification of contribution and overlap builds upon methods such as the Search Summary Table (SST) devised by Alison Bethel, which offers a structured overview of search results, indicating which searches found included references and

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

file.name	records	source	label	string
DIM2.ris	383	DIM2	search	database
Final.ris	55		final	
PaperCall.ris	40	PaperCall	search	database
Scopus2.ris	354	Scopus2	search	database
TI_AB.ris	100		screened	
WoS2.ris	208	WoS2	search	database
PaperCites.ris	113	CiteSearch	search	Forward
PaperRefs.ris	83	CiteSearch	search	Backward

Figure 11. *After uploading files to assess impact, users remove source tags for any files that represent records that have been screened and included in the final synthesis. Labels can be changed from 'search' to 'screened' or 'final'. In this example, forward and backward citation chasing are combined as 'CiteSearch' in the 'source' field.*

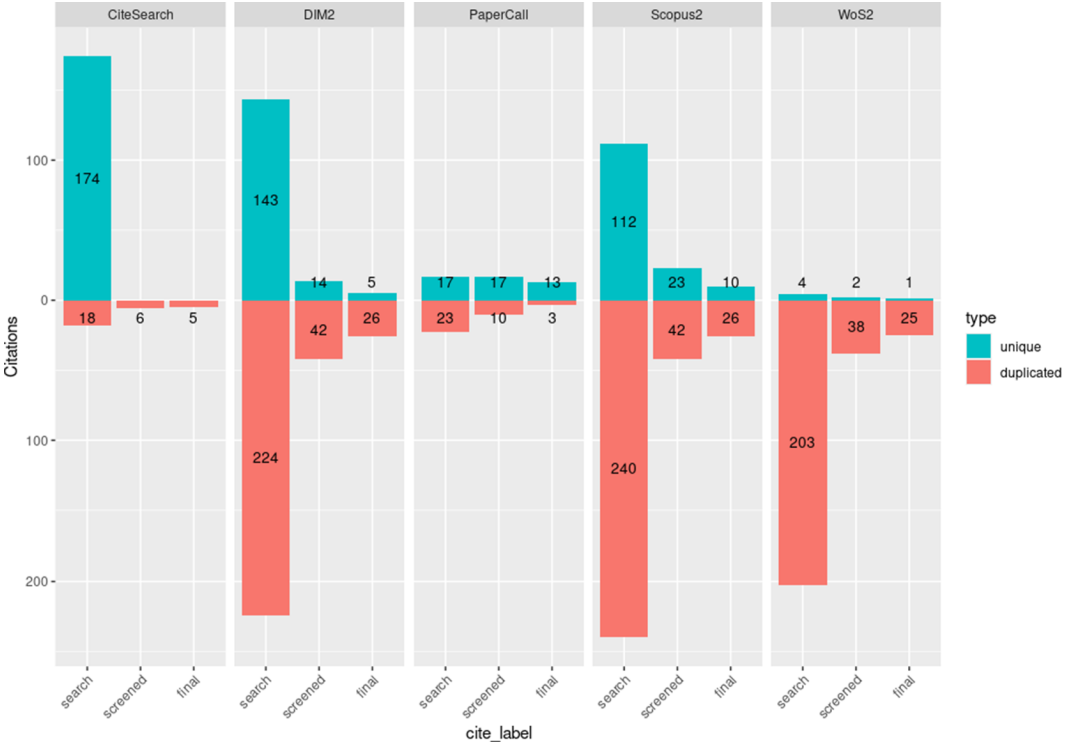


Figure 12. *The phase analysis bar plot shows the number of unique and duplicate records across each source and screening phase. Each facet contains the data for an individual source. In this example, the call for papers yielded the lowest number of records, but had a significant number of unique articles that were included in the final synthesis.*

providing sensitivity and precision calculations for search methods.[48] For instance, if certain methods consistently yield high-quality results that pass through all screening phases, they might be prioritized in future reviews or updates to the review. Conversely, methods or sources that contribute little value can be de-emphasized or excluded. This functionality enhances transparency about where and how

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

Record Counts & Precision/Sensitivity					
	Distinct Records[1]	Screened Included[2]	Final Included[3]	Precision[4]	Sensitivity/Recall[5]
CiteSearch	192	6	5	2.60	9.09
DIM2	367	56	31	8.45	56.36
PaperCall	40	27	16	40.00	29.09
Scopus2	352	65	36	10.23	65.45
WoS2	207	40	26	12.56	47.27
Total	[6] 696	[7] 100	[8] 55	[9] 7.90	NA

[1] Number of records after internal source deduplication.

[2] Number of citations included after title/abstract screening.

[3] Number of citations included after full text screening.

[4] Number of final included citations / Number of distinct records.

[5] Number of final included citations / Total number of final included citations.

[6] Total citations discovered (after internal and cross-source deduplication).

[7] Total citations included after Ti/Ab Screening.

[8] Total citations included after full text screening.

[9] Overall Precision = Number of final included citations / Total distinct records.

Figure 13. *The precision/sensitivity table shows each source/method's calculated precision and sensitivity. These calculations are based on the number of distinct records each source/method contributed, as well as the number of records from each source that were included in the final record set.*

included studies were found, facilitates a data-driven approach for future search method decisions, and deepens the understanding of the true impact of different resources and search methods. By providing insights into the performance and impact of each resource or method, CiteSource ensures that systematic reviews are both comprehensive and methodologically sound, supporting more robust and credible evidence synthesis.

4. Implications for evidence synthesis

CiteSource improves the rigor and effectiveness of evidence synthesis by providing data to inform critical decisions made during the search process. Traditionally, search strategy development often relies heavily on the researcher's experience, knowledge, and sometimes outdated guidance. CiteSource moves beyond this by allowing researchers to compare and optimize various aspects of their search strategy during the search strategy development phase.

Enhanced Efficiency and Time Savings: Evidence synthesis is time-consuming, especially the iterative process of developing and validating search strategies. CiteSource significantly compresses this cycle by offering rapid analysis and visualization (Heatmaps, Upset Plots, and Summary Tables) immediately after deduplication. This allows researchers to test variations in sources, strings, and methods much more quickly than analyzing results individually, saving time that can be redirected

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

to critical analysis and synthesis. For example, in the case of benchmarking, if a researcher wanted to analyze three variations of a string across three databases using ten benchmarking articles, each of the nine results would need to be cross-referenced against the set of benchmarking articles. This would require the user to search for each of the ten articles nine times and document the results. With CiteSource, the researcher can quickly analyze all nine sets of results alongside the benchmarking articles to see which articles were found using each string across each database. These results can also be exported in .csv or .ris for later reference and reporting.

Data-Driven Decisions for Optimization: CiteSource enables researchers to use data to make informed decisions about which information sources provide the best return on investment by quantifying their unique contributions and overlap. It also facilitates the testing and refinement of search strings by allowing rapid comparison of different variations and their impact on retrieval. This allows for a more strategic approach to balancing sensitivity and precision. Benchmark testing is also more feasible with CiteSource, as it provides a quantitative way to assess if key articles are captured.

Enhanced Reporting and Transparency: Reporting guidelines like PRISMA and ROSES require transparent and detailed documentation of the search process. CiteSource directly supports reporting by generating clear outputs and maintaining provenance. Plots and tables offer visual and quantitative summaries. Crucially, the tool embeds custom metadata into standard fields during export, providing a clear and reproducible audit trail of the methodology. This enhances the credibility and reproducibility of the synthesis.

Support for Methods Research: By providing a platform for quantifying the performance of different sources and methods (e.g., through precision/sensitivity calculations), CiteSource exposes underlying data such as detailed overlap statistics, source contribution breakdowns, and precision/sensitivity metrics across different search iterations, which can inform the future development of evidence-based search guidelines for the wider community. CiteSource also supports methodological research on searching itself, potentially facilitating 'Study Within A Review' (SWAR) focused on search methodology,[49] speed analysis of sources and methods,[17,20,50–55] and support initiatives and frameworks to improve search strategy development and validation, such as searchRxiv and CEEDER.[56–60]

5. Application beyond evidence synthesis

Beyond its direct contributions to the evidence synthesis workflow, CiteSource also offers valuable applications in related areas such as training and resource management.

Training and Education: CiteSource serves as an effective training tool for students and early-career researchers learning evidence synthesis methods. Its interactive visualizations offer a hands-on way to understand abstract concepts like Boolean logic, database coverage, and effects of string variations on search outcomes. Instructors can use it to demonstrate best practices and build practical skills in systematic searching. This makes the learning process more intuitive and data-informed.

Library Collection Development: Librarians, often tasked with making difficult resource allocation decisions, can leverage CiteSource to inform their collection development strategies. By analyzing search results, they can gather objective data on database coverage and overlap for specific research topics that are critical to their patrons. These data can be used to justify subscription costs, compare existing resources with potential new ones, and make recommendations based on demonstrated value and uniqueness.

6. Conclusion

The CiteSource R package and Shiny application provide a foundation for a new era of data-driven search strategy design in evidence synthesis. As the information landscape continues to evolve, openly available and accessible tools like CiteSource will be crucial for empowering researchers to make

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

informed decisions. CiteSource's unique ability to enhance transparency in reporting, support aggregate analysis of search performance, and expedite the iterative design and validation processes provides immediate, tangible benefits for improving the efficiency and rigor of evidence gathering. However, the true potential of CiteSource to transform the field will be realized through its widespread adoption and the collective contribution of its users. With regular reporting of these data, research teams have an opportunity to collectively contribute findings that can be further synthesized, leading to enhanced methodological practices and more robust, evidence-based guidelines for search strategy development. Whether integrated into the development of new syntheses to optimize evidence discovery or applied retrospectively to analyze and learn from past projects, CiteSource offers a powerful means to strengthen the quality, defensibility, and methodological integrity of evidence synthesis methods.

Acknowledgments. We thank the Evidence Synthesis Hackathon (https://www.eshackathon.org/) coordinators, especially Neal Haddaway and Chris Pritchard, for their contributions to initial development. We would also like to thank Alison Bethel for her inspirational work and review of table outputs. A special thanks to Kaitlyn Hair for her continued work on the Automated Systematic Search Deduplicator (ASySD) R package and for integrating critical deduplication functionality into CiteSource. Finally, we thank the anonymous reviewers for their thoughtful reviews of the manuscript. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Government, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Author contributions. TR: conceptualization, software, investigation, validation, data curation, visualization, project administration, writing—original draft, writing—review & editing. SY: conceptualization, investigation, validation, writing—original draft, writing—review & editing, funding acquisition. AP: writing—original draft, writing—review & editing. KH: conceptualization, software, investigation, validation, writing—review & editing. LW: conceptualization, software, investigation, validation, visualization, data curation, writing—review & editing. MG: conceptualization, software, investigation, validation, writing—review & editing.

Competing interest statement. The authors declare that no competing interests exist.

Data availability statement. The source code of the software is publicly available on GitHub (https://github.com/ESHackathon/CiteSource). All citation data presented in this paper are available on OSF (https://osf.io/hxusf/overview).

Funding statement. The authors declare that no specific funding has been received for the software described or for this article.

References

[1] Pullin AS, Stewart GB. Guidelines for systematic review in conservation and environmental management. *Conserv Biol*. 2006;20(6): 1647–1656. https://doi.org/10.1111/j.1523-1739.2006.00485.x.

[2] Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Res Synth Methods*. 2020;11(2): 181–217. https://doi.org/10.1002/jrsm.1378.

[3] Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev*. 2021;10(1): 39. https://doi.org/10.1186/s13643-020-01542-z.

[4] Lefebvre C, Glanville J, Briscoe S, et al. Chapter 4: Searching for and selecting studies. In: *Cochrane Handbook for Systematic Reviews of Interventions Version 651*. Cochrane; 2024. https://environmentalevidence.org/information-for-authors/.

[5] AS Pullin, GK Frampton, B Livoreil, et al. (Eds). Guidelines and Standards for Evidence synthesis in Environmental Management. Version 5.0. Collaboration for Environmental Evidence; 2018. https://environmentalevidence.org/wp-content/uploads/2022/10/CEE-Guidelines-Version-5.0-051022.pdf.

[6] Klerings I, Robalino S, Booth A, et al. Rapid reviews methods series: guidance on literature search. *BMJ Evid-Based Med*. 2023;28(6): 412–417. https://doi.org/10.1136/bmjebm-2022-112079.

[7] Booth A. How much searching is enough? Comprehensive versus optimal retrieval for technology assessments. *Int J Technol Assess Health Care*. 2010;26(4): 431–435. https://doi.org/10.1017/S0266462310000966.

[8] Cooper C, Booth A, Husk K, et al. A tailored approach: a model for literature searching in complex systematic reviews. *J Inf Sci*. 2022;50(4): 1030–1062. https://doi.org/10.1177/01655515221114452.

[9] Salvador-Olivan JA, Marco-Cuenca G, Arquero-Aviles R. Errors in search strategies used in systematic reviews and their effects on information retrieval. *J Med Libr Assoc*. 2019;107(2): 210–221. https://doi.org/10.5195/jmla.2019.567.

[10] Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *J Clin Epidemiol*. 2006;59(10): 1057.e1–1057.e9. https://doi.org/10.1016/j.jclinepi.2006.01.007.

[11] Michael Clark J, Beller E, Glasziou P, Sanders S. The decisions and processes involved in a systematic search strategy: a hierarchical framework. *J Med Libr Assoc*. 2021;109(2): 201–211. https://doi.org/10.5195/jmla.2021.1086.

[12] MacFarlane A, Russell-Rose T, Shokraneh F. Search strategy formulation for systematic reviews: issues, challenges and opportunities. *Intell Syst Appl*. 2022;15. https://doi.org/10.1016/j.iswa.2022.200091.
[13] Livoreil B, Glanville J, Haddaway NR, et al. Systematic searching for environmental evidence using multiple tools and sources. *Environ Evid*. 2017;6(1). https://doi.org/10.1186/s13750-017-0099-6.
[14] Konno K, Pullin AS. Assessing the risk of bias in choice of search sources for environmental meta-analyses. *Res Synth Methods*. 2020;11(5): 698–713. https://doi.org/10.1002/jrsm.1433.
[15] 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. https://doi.org/10.1016/j.jclinepi.2021.03.013.
[16] Konno K, Akasaka M, Koshida C, et al. Ignoring non-English-language studies may bias ecological meta-analyses. *Ecol Evol*. 2020;10(13): 6373–6384. https://doi.org/10.1002/ece3.6368.
[17] Frandsen TF, Moos C, Marino C, Eriksen MB. Supplementary databases increased literature search coverage beyond PubMed and Embase. *J Clin Epidemiol*. 2025;181: 111704. https://doi.org/10.1016/j.jclinepi.2025.111704.
[18] Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev*. 2017;6(1): 245. https://doi.org/10.1186/s13643-017-0644-y.
[19] Frandsen TF, Eriksen MB. Supplementary strategies identified additional eligible studies in qualitative systematic reviews. *J Clin Epidemiol*. 2023;159: 85–91. https://doi.org/10.1016/j.jclinepi.2023.04.017.
[20] 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*. 2018;9(2): 195–223. https://doi.org/10.1002/jrsm.1286.
[21] 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 Inf Libr J*. 2010;27(2): 114–122. https://doi.org/10.1111/j.1471-1842.2009.00863.x.
[22] Briscoe S, Abbott R, Melendez-Torres GJ. Expert searchers identified time, team, technology and tension as challenges when carrying out supplementary searches for systematic reviews: a thematic network analysis. *Health Inf Libr J*. 2024;41(2): 182–194. https://doi.org/10.1111/hir.12468.
[23] Briscoe S, Abbott R, Melendez-Torres GJ. The phronesis of expert searchers on using forward citation searching and web searching to search for studies for systematic reviews: a hermeneutic phenomenological analysis. *J Inf Sci*. 2022;51(1): 96–110. https://doi.org/10.1177/01655515221130237.
[24] Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol*. 2015;68(6): 617–626. https://doi.org/10.1016/j.jclinepi.2014.11.025.
[25] Meert D, Torabi N, Costella J. Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. *J Med Libr Assoc: JMLA*. 2016;104(4): 267–277. https://doi.org/10.3163/1536-5050.104.4.004.
[26] Ghezzi-Kopel K, Ault J, Chimwaza G, et al. Making the case for librarian expertise to support evidence synthesis for the sustainable development goals. *Res Synth Methods*. 2022;13(1): 77–87. https://doi.org/10.1002/jrsm.1528.
[27] Kate Silverman W. MIT Press announces new initiative to increase the reach of academic scholarship. 2023. Accessed September 2025. https://news.mit.edu/2023/mit-press-announces-new-initiative-increase-reach-academic-scholarship-0207
[28] Clarivate. Clarivate successfully completes acquisition of ProQuest. 2021. Accessed September, 2025. https://clarivate.com/news/clarivate-successfully-completes-acquisition-of-proquest/
[29] Sage. Sage acquires the scientific and medical publisher Mary Ann Liebert, Inc. 2024. Accessed September, 2025. https://www.sagepub.com/explore-our-content/press-office/press-releases/2024/12/13/sage-acquires-the-scientific-and-medical-publisher-mary-ann-liebert-inc
[30] Brainard J. Fast-growing open-access journals lose impact factors. *Science*. 2023;379(6639): 1283–1284. https://doi.org/10.1126/science.adi0092.
[31] Barshay J. A treasure trove of education reports and studies is under threat. *The Hechinger Report*. https://hechingerreport.org/proof-points-eric-under-threat/
[32] Cooper C, Varley-Campbell J, Booth A, Britten N, Garside R. Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use. *J Clin Epidemiol*. 2018;99: 53–63. https://doi.org/10.1016/j.jclinepi.2018.02.025.
[33] Cooper C, Garside R, Varley-Campbell J, Talens-Bou J, Booth A, Britten N. "It has no meaning to me." how do researchers understand the effectiveness of literature searches? A qualitative analysis and preliminary typology of understandings. *Res Synth Methods* 2020;11(5): 627–640. https://doi.org/10.1002/jrsm.1426.
[34] MacDonald H, Comer C, Foster M, et al. Searching for studies: a guide to information retrieval for Campbell systematic reviews. *Campbell Syst Rev*. 2024;20(3): e1433. https://doi.org/10.1002/cl2.1433.
[35] Gusenbauer M. Search where you will find most: comparing the disciplinary coverage of 56 bibliographic databases. *Scientometrics*. 2022;127(5): 2683–2745. https://doi.org/10.1007/s11192-022-04289-7.
[36] Gusenbauer M. Beyond Google scholar, Scopus, and web of science: an evaluation of the backward and forward citation coverage of 59 databases' citation indices. *Res Synth Methods*. 2024;15(5): 802–817. https://doi.org/10.1002/jrsm.1729.
[37] Lefebvre C, Glanville J, Wieland LS, Coles B, Weightman AL. Methodological developments in searching for studies for systematic reviews: past, present and future? *Syst Rev* 2013;2(1): 78. https://doi.org/10.1186/2046-4053-2-78.

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press

[38] Grames EM, Stillman AN, Tingley MW, Elphick CS. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods Ecol Evol*. 2019;10(10): 1645–1654. https://doi.org/10.1111/2041-210x.13268.

[39] Kapp C, Fujita-Rohwerder N, Lilienthal J, Sieben W, Waffenschmidt S, Hausner E. The searchbuildR shiny app: a new implementation of the objective approach for search strategy development in systematic reviews. *Cochrane Evid Synth Methods*. 2024;2(6). https://doi.org/10.1002/cesm.12078.

[40] 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*. 2023;10: 101935. https://doi.org/10.1016/j.mex.2022.101935.

[41] Gremmen-Verleg M, Gremmen-Verleg BE. Medsyntax [Software]. 2023. https://www.medsyntax.org

[42] 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. *BMC Biol*. 2023;21(1): 189. https://doi.org/10.1186/s12915-023-01686-z.

[43] Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. *IEEE Trans Vis Comput Graph*. 2014;20(12): 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248.

[44] Lagisz M, Yang Y, Young S, Nakagawa S. A practical guide to evaluating sensitivity of literature search strings for systematic reviews using relative recall. *Res Synth Methods*. 2025;16(1): 1–14. https://doi.org/10.1017/rsm.2024.6.

[45] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372: n71. https://doi.org/10.1136/bmj.n71.

[46] O'Dea RE, Lagisz M, Jennions MD, et al. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biol Rev*. 2021. https://doi.org/10.1111/brv.12721.

[47] Haddaway NR, Macura B, Whaley P, Pullin AS. ROSES RepOrting standards for systematic evidence syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environ Evid* 2018;7(1): 1–8. https://doi.org/10.1186/s13750-018-0121-7.

[48] Bethel AC, Rogers M, Abbott R. Use of a search summary table to improve systematic review search methods, results, and efficiency. *J Med Libr Assoc*. 2021;109(1): 97–106. https://doi.org/10.5195/jmla.2021.809.

[49] Devane D, Burke NN, Treweek S, et al. Study within a review (SWAR). *J Evid Based Med*. 2022;15(4): 328–332. https://doi.org/10.1111/jebm.12505.

[50] Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Med Res Methodol*. 2016;16(1): 127. https://doi.org/10.1186/s12874-016-0232-1.

[51] Levay P, Heath A, Tuvey D. Efficient searching for NICE public health guidelines: would using fewer sources still find the evidence? *Res Synth Methods*. 2022;13(6): 760–789. https://doi.org/10.1002/jrsm.1577.

[52] Bethel AC, Shaw N, Abbot R, et al. Insights from search summary tables for evidence and gap maps: a case study on peer support interventions. *J Med Libr Assoc*. 2025;113(2): 177–183. https://doi.org/10.5195/jmla.2025.1831.

[53] Ewald H, Klerings I, Wagner G, et al. Searching two or more databases decreased the risk of missing relevant studies: a metaresearch study. *J Clin Epidemiol*. 2022;149: 154–164. https://doi.org/10.1016/j.jclinepi.2022.05.022.

[54] 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*. 2016;5: 39. https://doi.org/10.1186/s13643-016-0215-7.

[55] Gargon E, Williamson PR, Clarke M. Collating the knowledge base for core outcome set development: developing and appraising the search strategy for a systematic review. *BMC Med Res Methodol*. 2015;15: 26. https://doi.org/10.1186/s12874-015-0019-9.

[56] White J. searchRxiv: a resource for sharing database search strategies. *Med Ref Serv Q*. 2024;43(1): 72–79. https://doi.org/10.1080/02763869.2024.2286856.

[57] Konno K, Cheng SH, Eales J, et al. The CEEDER database of evidence reviews: an open-access evidence service for researchers and decision-makers. *Environ Sci Pol*. 2020;114: 256–262. https://doi.org/10.1016/j.envsci.2020.08.021.

[58] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75: 40–46. https://doi.org/10.1016/j.jclinepi.2016.01.021.

[59] Huynh Khanh Vi Tran JB, Ali NB, Unterkalmsteiner M. How good are my search strings? Reflections on using an existing review as a quasi-gold standard. *arXiv.* 2024; 1–20. https://doi.org/10.48550/arXiv.2402.11041.

[60] Haddaway NR, Rethlefsen ML, Davies M, et al. A suggested data structure for transparent and repeatable reporting of bibliographic searching. *Campbell Syst Rev*. 2022;18(4): e1288. https://doi.org/10.1002/cl2.1288.

Cite this article: Riley T, Young S, Paxton A, Wallrich L, Hair K, Grainger M. CiteSource: An R package for data-driven search strategy development and enhanced evidence synthesis reporting. *Research Synthesis Methods*. 2026;17: 1026–1044. https://doi.org/10.1017/rsm.2026.10084

https://doi.org/10.1017/rsm.2026.10084 Published online by Cambridge University Press