



Deduplication methods for literature citations in systematic evidence reviews: practical insights to guide decision-making

Nicole Askin¹ · George Okoli¹

Received: 12 April 2024 / Revised: 7 July 2024 / Accepted: 8 August 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

When deciding on an appropriate software and/or methodology for deduplication of retrieved literature citations in systematic evidence synthesis, several factors ought to be considered, including the cost of a deduplication software, familiarity with its use, and the time necessary to conduct deduplication using a particular method. To guide researchers with and without knowledge of systematic evidence synthesis methods in making informed decisions, we sought to compare different commonly used deduplication software/methodologies—Covidence, Rayyan, the Automated Systematic Search Deduplicator, the Systematic Review Accelerator (SRA) Deduplicator, Deduklick, and the Bramer method in Endnote—from the perspectives of an experienced information specialist and a biomedical researcher. We summarized the cost, time, and effort required from each deduplication software/method and reported our findings descriptively. The findings suggest a properly balanced consideration between cost-effectiveness, ease of application of a method, context and researcher’s needs when deciding on deduplication method for systematic evidence synthesis.

Keywords Evidence synthesis · Systematic reviewing · Deduplication · Methodology

1 Introduction

Systematic evidence review/synthesis refers to systematically identifying, selecting, appraising, and synthesizing results on a specific topic from multiple studies. This includes methods such as the scoping/mapping review, rapid review, and systematic review with or without meta-analysis. Irrespective of the method of systematic evidence synthesis, it is a time- and resource-intensive process, involving development of a sensitive search strat-

✉ Nicole Askin
nicole.askin@umanitoba.ca

George Okoli
George.Okoli@umanitoba.ca

¹ University of Manitoba, Winnipeg, MB, Canada

egy for an electronic bibliographic database adapted to different bibliographic data sources with overlapping coverage, and often resulting in a wide retrieval of literature citations, typically in the thousands depending on the topic of review. This usually includes a significant number of duplicate literature citation records: Bramer (2015) identified a median of 43% duplicates in systematically retrieved citations. Detection of duplicate records among retrieved literature citations is often difficult due to varying metadata between data sources, such as differences in journal or author name presentation, or absence of metadata such as digital object identifiers (DOIs) for particular types of publications. The inconsistencies in metadata negatively impact the efficiency of deduplication functions in traditional reference management software. Nevertheless, removing obvious duplicates prior to the commencement of title/abstract screening reduces reviewer workload substantially, by limiting duplication of effort. Thus, deduplication (detection and removal of duplicate records) is a standard component of evidence synthesis methodologies.

In a survey of health technology assessments and guideline developments, Scott and colleagues concluded that automation tools are highly necessary and of immense importance in systematic evidence synthesis, with deduplication tools found to be the fourth most requested specific automation tools (Scott et al. 2021). Cleo and colleagues (2019) conducted a usability and acceptability review of four systematic review automation tools, including Covidence and Rayyan. While participants generally found these software easy to learn, they had concerns about response time and software “glitches” (Cleo et al. 2019).

Previous comparisons of specific duplication softwares and methods (e.g. McKeown and Mir 2021; Guimaraes et al., 2022) have focused on accuracy, but there is a lack of literature on other factors that impact method selection, in particular ease of use. For more insights to guide researchers with and without knowledge of systematic evidence synthesis methods, we sought to compare the cost, both in terms of the required time and subscription/purchase fees, as well as the usability of six different approaches to deduplication, from the perspective of an information specialist with extensive experience in deduplication as well as a biomedical researcher with expertise in systematic reviewing but limited deduplication experience.

There are several available methods for identifying duplicates among retrieved literature citations following a systematic literature search, ranging from a manual inspection (particularly for more limited searches) to purpose-built tools (Adam et al. 2022); we sought to evaluate a range of commonly used approaches. The Bramer method in Endnote is used by many information professionals; it employs an iterative technique based on changing the parameters used in Endnote to assess potential duplicates. Commonly used screening softwares Covidence and Rayyan incorporate duplication detection. There are also multiple specific app solutions; for this study we evaluated the Automated Systematic Search Deduplicator, the Systematic Review Accelerator (SRA) Deduplicator, and Deduklick.

2 Methods

A total of 25,729 citations were retrieved using a strategy adapted from a published systematic review on interventions targeted to care providers to improve seasonal influenza vaccination rates (Okoli et al. 2021). The searches were performed on 31 May 2023. Four databases across three different platforms were searched: Medline on Ovid, Embase on

Ovid, CINAHL with Full Text on EBSCOhost, and Scopus (Elsevier). Animal studies and commentary were excluded. A date limit of 2000 to present was used, but no language limit was included. Results were exported as RIS files. A search log for Medline is available as Appendix A.

Records were considered duplicates only if they represented the same bibliographic record. Multiple reports on the same study (for example, a conference abstract versus a journal article) were not considered to be duplicates. Repeated information of this type is typically addressed at a later stage of an evidence synthesis project. The time necessary for importing citations, deduplicating, and exporting deduplicated results were recorded for each method. Additionally, for the Bramer method, the time to download custom filters and optimize setup was also recorded; this time would be necessary for the initial review only and would not need to be repeated for subsequent projects. Time spent purchasing or downloading software was not included in time recordings. Default settings were used for software other than Endnote, except for SRA Deduplicator for which the “relaxed” deduplicator algorithm was selected. For the paid software, existing personal/institutional subscriptions were used for Endnote and Covidence, and free trials were used for Deduplick and Rayyan. Comparisons of the above deduplication processes were made between an experienced information specialist and a biomedical researcher, and the findings reported descriptively. In addition to recording the time required and the number of results following deduplication, the researchers recorded narrative observations of the deduplication process with each method.

3 Results

Table 1 is a description of the software compared in this study and the associated costs. The final record counts for each researcher according to method following deduplication are presented as Table 2. Table 3 is the time (in minutes) for each stage of the deduplication process according to researcher and method. There were variations, both between the methods and researchers, even from the most automated methods. Costs also vary depending on researcher needs, although some researchers may benefit from institutional subscriptions.

BR=biomedical researcher; IS=information specialist; N/A=not applicable; With import=deduplication and import happen at the same time and are therefore not exclusive; Not done=reviewer was unable to complete the process.

For both researchers the Bramer method was by far the most time intensive. As acknowledged by the creator, the Bramer method is complex and involves a significant learning curve (Bramer et al. 2016). While the method is well-explained, it is potentially cumbersome and time-intensive to put into practice. It also relies on use of Endnote, requiring that researchers purchase a license. While the custom import and export filters address some metadata inconsistencies between databases, there are others that remain unaddressed. For example, non-English records imported from Medline as RIS files often display the original non-English title in the Journal field, which complicates correct identification of duplicates using this field.

ASySD is accessible online and does not require downloading of software or a subscription purchase for use. However, the online version does not adapt well to large datasets (the documentation notes an upper limit of 50,000 citations). The dataset also disappears if the

Table 1 Description of the software options used for this study

Software	Provider	Costs	Description
EndNote 21 (Bramer method)	Clarivate	\$275 USD (full license) \$150 (student license)	The Bramer method (Bramer et al. 2016) uses the Endnote software. Adopts an iterative approach of changing the fields in a record which are used by the software to identify duplicates. Incorporates custom import and export filters to address some known inconsistencies in metadata formatting.
Automated Systematic Search Deduplicator (ASySD)	CAMA-RADES (University of Edinburgh)	Free	Combines several packages in the R software to identify duplicate citations (ASySD, n.d.).
Systematic Review Accelerator (SRA) Deduplicator	Institute for Evidence-Based Healthcare (Bond University)	Free	Deduplication module uses fuzzy matching to merge references deemed identical (SR-Accelerator, n.d.).
Deduklick	Risklick	\$54/year USD (single use)	Uses a “multistep algorithm of data normalization” to identify duplicates based on similarity scores (Borissov et al. 2022; Deduklick, n.d.).
Covidence	Covidence	\$289/year USD (single use)	Includes support for deduplication, screening, and data extraction (Covidence, n.d.). Primary screening tool used for authorship of Cochrane reviews.
Rayyan	Rayyan Systems	\$8.25/month USD (professional) \$4/month USD (student) *Free version exists but lacks auto-resolver features	Initially launched as a free tool but now largely subscription-based; its Auto-Resolver tool appears to only be available to subscribers (Rayyan 2021).

Table 2 Retained record counts for each researcher and according to software type following deduplication

Method	Biomedical researcher results	Information specialist results
Bramer method	12,337	12,623
ASySD	12,641	12,712
SRA Deduplicator	12,503	12,929
Deduklick	12,843	12,858
Covidence	13,077	13,103
Rayyan	13,300	13,230

Table 3 Time (in minutes) for each stage of the deduplication process, by researcher and software type

Methods	Customization		Reference import		Deduplication		Export		Total	
	BR	IS	BR	IS	BR	IS	BR	IS	BR	IS
Bramer	5	2	45	18	80	70	7	1	137	91
ASySD	N/A	N/A	9	4	56	25	4	1	69	30
SRA Deduplicator	N/A	N/A	4	2	55	40	2	1	61	43
Deduklick	N/A	N/A	5	5	With import	With import	10	1	15	6
Covidence	N/A	N/A	81	5	With import	With import	3	1	84	46
Rayyan	N/A	N/A	10	5	55	45	Not done	2	65+	52

webpage is reloaded, which can present another technical problem. Further, the preferred upload format is Endnote XML, which requires either that this is selected as the download format (where available) from the data source, or that the files are preprocessed through a reference management software.

Like ASySD, SRA Deduplicator is available online without download or purchase. One researcher (the biomedical researcher) reported that the use of the tool was not clearly explained in its documentation. The other researcher (the information specialist) appreciated the grouping of references according to likelihood of duplicates, as well as the availability of a split group functionality to simplify addressing misidentification of potential duplicates; however, it did not appear to be possible to address true duplicates that were sorted to the non-duplicates grouping.

For both researchers, Deduplicator was the fastest option for the deduplication itself. However, both researchers encountered significant technical difficulties in accessing the deduplication functionality, necessitating intervention (resolution) by the software's support team. The actual deduplication process was relatively user-friendly. One researcher (the biomedical researcher) particularly appreciated that it provided numbers for pre- and post-deduplication according to record source. However, it was not possible to manually intervene in the case of an incorrectly assessed duplicate or non-duplicate.

Covidence's import and deduplication processes happen simultaneously – records are automatically deduplicated at point of upload. However, limits on file size necessitated uploading in batches, which increased time investment. Using Covidence for deduplication simplifies later processing when the same software is used for record screening.

Rayyan offers different levels of deduplication depending on subscription or software version. The free version is extremely manual; it proved infeasible to rely solely on that method given the size of the sample dataset, so the researchers relied instead on a paid version. The auto-resolver feature offered to subscribers significantly reduced the time necessary for manual processing – for one researcher (the information specialist) it cut the unresolved potential duplicates count from over 19,000 to only 6320. The beta version had the potential to decrease this yet further since it allowed for auto-resolution of potential duplicates at a lower level of certainty (the non-beta version allowed only as low as 95%). The beta version also had a much more user-friendly interface for duplicate resolution. However, even at a high level of certainty, the auto-resolver identified non-duplicates as duplicates where the only difference between titles was a number; for example, several *Morbidity and Mortality Weekly Report* entries were incorrectly flagged as duplicates because they reported on succeeding years. Additionally, one researcher (the biomedical researcher) was unable to successfully export a deduplicated file.

4 Discussion

In considering what software or methodology to use for deduplication in systematic evidence synthesis, there are several factors to consider, including the cost of a software, familiarity with the use, the time necessary to conduct deduplication using a particular method, and the effectiveness of a method in correctly identifying duplicate results. Time needed for deduplication must be balanced against effectiveness, as more effective deduplication may take longer initially but reduce time spent in later stages of the analysis process.

Other considerations which may impact decision-making are existing institutional subscriptions as well as plans for further analysis; for example, if a research team intends to use a software for reference screening, it may be simpler to use the same software for deduplication. Selection of an appropriate deduplication method should be cost-effective because it allows a research team to limit the time and resources necessary to complete this task. Finally, user-friendliness and tool stability are important considerations in initial selection, as an investment of time is needed to become proficient with a tool and troubleshoot technical issues that it may present, and that investment may be lost if a tool proves unstable or otherwise unsuitable.

Several studies have compared systematic review tools against generic reference management software. Clark and colleagues (2020) explored citation deduplication using Systematic Review Accelerator (SRA) Deduplicator and were able to remove duplicate records in 16 min, a significant reduction over our findings. A similar study also by Clark and colleagues (2021) found that manual deduplication took over two hours while deduplication with SRA Deduplicator took only 36 min, with both times including the time needed to learn the task; in this context their findings were more comparable to ours. They noted a slightly higher error rate using this tool rather than manual deduplication. McKeown and Mir (2021) found Covidence and Rayyan to significantly improve on deduplication performance of the default settings of several reference management softwares, with an accuracy of 96% and 97% respectively compared to manual deduplication. Forbes and colleagues (2022) compared SRA Deduplicator to the Bramer method of deduplication utilising citations retrieved for Cochrane reviews. They found that while both methods had similar accuracy, use of SRA Deduplicator was 330% faster than the Bramer method, with a median deduplication time of 6.5 min compared to 25 min with the Bramer method; our results showed a similar time preference, although at a reduced ratio (Forbes et al. 2022). Borissov and colleagues (2022) found that Deduklick demonstrated improved recall and took an average time of less than a minute compared with 70 min for manual deduplication. Finally, Guimaraes et al. (2022) found that Systematic Review Accelerator and Rayyan had high sensitivity compared to the default settings of Zotero and Endnote.

5 Limitations

This study adapted the methods of a previous systematic review to generate the citation dataset, and as such, while our study represents a real-world application of the deduplication methodologies of interest, it is also limited in terms of topic area and databases used. It is possible that a search on a different topic or using different data sources may have exhibited differences in the proportion of duplicates and the time necessary to identify them using the evaluated methodologies. Due to the size of the result set we did not pursue accuracy assessment of the deduplication processes; this consideration has been well described in existing literature. The differences in deduplication times between the researchers could be due to differing levels of experience and approach to the task of deduplication, but also due to other technical factors such as differences in Internet speed or computing power. The order in which deduplication methods were tested may also have impacted results due to practice bias.

The software evaluated in this study continue to evolve over time. For example, Covidence implemented changes to its deduplication process specifically in response to McKewon and Mir's 2021 study (McLoughlin 2022). This study did not consider potential future methodological updates, such as the proposed adaptation of the Bramer method by Main (2023). Noteworthy is that this study identified a nearly 50% reduction in the time required for use of that method, which is a promising outcome given our findings.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10742-024-00335-4>.

Author contributions All authors contributed equally to the conception, conduct, and reporting of this work.

Funding No funding was provided.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

- Adam, G.P., Wallace, B.C., Trikalinos, T.A.: Semi-automated Tools for Systematic Searches. In E. Evangelou & A. A. Veroniki (Eds.), *Meta-Research: Methods and Protocols* (pp. 17–40). Springer US. (2022). https://doi.org/10.1007/978-1-0716-1566-9_2
- ASySD. (n.d.). Retrieved September 9: from (2023). <https://camarades.shinyapps.io/RDedup/>
- Borissov, N., Haas, Q., Minder, B., Kopp-Heim, D., von Gernler, M., Janka, H., Teodoro, D., Amini, P.: Reducing systematic review burden using deduklick: A novel, automated, reliable, and explainable deduplication algorithm to foster medical research. *Syst. Reviews*. **11**(1), 172 (2022). <https://doi.org/10.1186/s13643-022-02045-9>
- Bramer, W.M.: Removing duplicates in retrieval sets from electronic databases. (2015). <https://www.bmi-online.nl/wp-content/uploads/2015/01/bramer-presentation-Removing-duplicates-in-retrieval-sets-from-electronic-databases-zonder-demo.pdf>
- Bramer, W.M., Giustini, D., de Jonge, G.B., Holland, L., Bekhuis, T.: De-duplication of database search results for systematic reviews in EndNote. *J. Med. Libr. Association*. **104**(3), 240–243 (2016). <https://doi.org/10.3163/1536-5050.104.3.014>
- Clark, J., Glasziou, P., Del Mar, C., Bannach-Brown, A., Stehlik, P., Scott, A.M.: A full systematic review was completed in 2 weeks using automation tools: A case study. *J. Clin. Epidemiol*. **121**, 81–90 (2020). <https://doi.org/10.1016/j.jclinepi.2020.01.008>
- Clark, J., McFarlane, C., Cleo, G., Ramos, C.I., Marshall, S.: The Impact of Systematic Review Automation Tools on Methodological Quality and Time Taken to complete systematic review tasks: Case Study. *JMIR Med. Educ*. **7**(2), e24418 (2021). <https://doi.org/10.2196/24418>
- Cleo, G., Scott, A.M., Islam, F., Julien, B., Beller, E.: Usability and acceptability of four systematic review automation software packages: A mixed method design. *Syst. Reviews*. **8**(1), 145 (2019). <https://doi.org/10.1186/s13643-019-1069-6>
- Covidence. (n.d.). Covidence. Retrieved September 9: from (2023). <https://www.covidence.org/>
- Deduklick. (n.d.). Retrieved September 9: from (2023). <https://www.risklick.ch/products/deduklick/>
- Forbes, C., Clark, J., Greenwood, H.: Automation of Duplicate Detection for Systematic Reviews. *ICML* (2022)
- Guimarães, N.S., Ferreira, A.J.F., Silva, R., de Paula, R., Lisboa, A.A., Magno, C.S., Ichiara, L., M. Y., Barreto, M.L.: Deduplicating records in systematic reviews: There are free, accurate automated ways to do so. *J. Clin. Epidemiol*. **152**, 110–115 (2022). <https://doi.org/10.1016/j.jclinepi.2022.10.009>
- Main, E.: Extending the Bramer Method with EndNote 20: Deduplicating with DOI. *The J. Can. Health Libr. Association*. **44**(2), 48–63 (2023). <https://doi.org/10.29173/jchla29720>

- McKeown, S., Mir, Z.M.: Considerations for conducting systematic reviews: Evaluating the performance of different methods for de-duplicating references. *Syst. Reviews.* **10**(1), 38 (2021). <https://doi.org/10.1186/s13643-021-01583-y>
- McLoughlin, R.: Improving our deduplication process. *Covidence.* (2022)., March 7 <https://www.covidence.org/blog/improving-our-deduplication-process/>
- Okoli, G.N., Reddy, V.K., Lam, O.L.T., Abdulwahid, T., Askin, N., Thommes, E., Chit, A., Abou-Setta, A.M., Mahmud, S.M.: Interventions on health care providers to improve seasonal influenza vaccination rates among patients: A systematic review and meta-analysis of the evidence since 2000. *Fam. Pract.* **38**(4), 524–536 (2021). <https://doi.org/10.1093/fampra/cmaa149>
- Rayyan: (2021)., November 8 <https://www.rayyan.ai/>
- Scott, A.M., Forbes, C., Clark, J., Carter, M., Glasziou, P., Munn, Z.: Systematic review automation tools improve efficiency but lack of knowledge impedes their adoption: A survey. *J. Clin. Epidemiol.* **138**, 80–94 (2021). <https://doi.org/10.1016/j.jclinepi.2021.06.030>
- SR-Accelerator. (n.d.). Retrieved September 9: from (2023). <https://sr-accelerator.com/#/deduplicator>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.