

Science/Education Portraits IX-Reproducibility and Transparency of Systematic Reviews

Raphael YH Hon¹ and Maurice HT Ling^{2*}

¹Yong Loo Lin School of Medicine, National University of Singapore, Singapore

²HOHY Pte Ltd, Singapore

*Corresponding Author: Maurice HT Ling, HOHY Pte Ltd, Singapore.

Received: July 06, 2024; Published: July 23, 2024

DOI: 10.55162/MCMS.07.231

Abstract

The main advantage of systematic review (SR) over other review types, such as narrative review, is its transparent and rigorous approach to minimize bias and ensure future replicability. Reproducibility, comprehensiveness, and transparency are the 3 keys to minimize bias. Reproducibility and transparency are deemed more important than comprehensiveness as missing literature may be identifiable from a reproducible and transparent SR lacking comprehensiveness but missing literature cannot be identified from a comprehensive SR lacking reproducibility and/or transparency. Hence, a SR lacking reproducibility and/or transparency is a narrative review. Several studies on published SRs report poor reproducibility and transparency. Here, we examine a sample of 100 SRs published in 2023 and indexed within PubMed for reproducibility and transparency on 10 criteria - 7 for reproducibility and 3 for transparency. In terms of reproducibility, our results show that only 31 SRs gave sufficient information for reconstructed searches in PubMed to yield search results within 10% of the hits reported. In terms of transparency, none of the SRs provided the full list of search results and only 5 SRs provided sufficient information to determine which study is excluded by which exclusion criterion. Hence, our results support previous studies reporting poor reproducibility and transparency in SRs. We propose that search URLs should be given whenever possible, and the full list of search results with inclusion / exclusion analysis should be given in all cases.

Introduction

With increasing research papers published, reviews are necessary to summarize and synthesize the field periodically. The process of selecting source materials for inclusion results in 2 major classes of reviews - narrative review, and systematic review (SR). A narrative review is less formal than a SR as it does not require the presentation of reporting methodology, search terms, databases used, and inclusion and exclusion criteria [1]. The first SR is likely a 1753 paper by James Lind reviewed all the previous publications about scurvy [2]. The aim of a SR [3] is "to collate all the empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made (Antman et al 1992, Oxman and Guyatt 1993). Systematic review methodology, pioneered and developed by Cochrane, sets out a highly structured, transparent and reproducible methodology (Chandler and Hopewell 2013)." Hence, SRs are less prone to bias compared to narrative reviews [4], a view echoed by Mallett et al [5]. However, each has their own purpose and are complementary to each other [6]. The mindset between SR and narrative review is explained by Siddaway et al. [7] quoting Baumeister [8] - a SR adopts the mindset of a judge and jury who evaluate the evidence to render the fairest judgment possible, whereas a narrative review adopts the lawyer's approach to make the best case for one side of the argument.

Hence, introduction section of an article is usually narrative rather than systematic.

The 3 keys to minimize bias are reproducibility, comprehensiveness, and transparency; as the quality of a SR is only as good as its included studies [9]. Reproducibility is the ability to reproduce the search results within 10% of the number of results from the original search using the search descriptions [10, 11]. Comprehensiveness is the completeness of all relevant literature on a specific research question or topic [12-14]. Transparency is to define inclusion / exclusion criteria unambiguously, and the reason for each article's inclusion or exclusion [7, 15-17]. Reproducibility and transparency are more important than comprehensiveness - missing literature may be identifiable from a reproducible and transparent SR lacking comprehensiveness but missing literature cannot be identified from a comprehensive SR lacking reproducibility and/or transparency. A SR lacking reproducibility and/or transparency is a narrative review. Several studies on published SRs report poor reproducibility [10, 11] and transparency [16, 17]. Therefore, in this study, we examine a sample of SRs published in 2023 and indexed within PubMed for reproducibility and transparency. Our results show that only 31% of the PubMed reconstructed searches in PubMed were within 10% of the hits reported, and only 5% of the SRs show which articles are excluded under which exclusion criteria. More importantly, none of the SRs provided the full list of search results and only 5 SRs provided sufficient information to determine which study is excluded by which exclusion criterion.

Methods

A search on PubMed was performed on January 26, 2024; using "systematic review[title]" as search term¹ with 3 filters; namely, (a) date range between January 1 to December 31 of 2023, (b) SR as article type, and (c) free full text availability; and returning the first 9999² hits as the maximum number of hits returned using Entrez ESearch function [18, 19]. 100 SRs were randomly sampled for evaluation on a set on 10 criteria. Reproducibility of search for each article [10, 11] was determined by 7 criteria; namely, (a) whether the date of search is given, (b) whether the database of search was given, (c) whether the full search term(s) including fields and/or Boolean operators were given, (d) whether the date range of the search was given, (e) whether the search URL was given, (f) whether it was possible to repeat the search based on description of search strategy or search URL, and (g) whether the number of hits obtained was within 10% of that reported in the original article. Transparency [7, 15-17] was determined by 3 criteria; namely, (a) whether the full list of hits was given, (b) whether the inclusion / exclusion criteria were unambiguous, and (c) whether it was possible to determine which article was excluded by which exclusion criterion. 95% confidence interval of percentages were calculated using Wilson score interval method [20], which is an improvement over normal approximation [21, 22].

Results and Discussion

Our PubMed search yield 10243 hits; of which, 9999 hits were collected. Of these 9999 hits, 100 were randomly selected for analysis using 10 criteria of reproducibility [10, 11] and transparency [7, 15-17]. The results are summarized in Table 1.

In terms of reproducibility, our results show that all 100 (95% CI: 96.30% to 100.00%) SRs gave the database searched but only 91 (95% CI: 83.77% to 95.19%) SRs gave actual search terms; and only 76 (95% CI: 66.77% to 83.31%) SRs gave the date range for search, with only 23% (95% CI: 15.84% to 32.15%) showing the date when the search was performed. Nevertheless, we are able to repeat all PubMed searches based on the description given. Only when it is possible to re-perform the search by either using search URL or reconstructing the search from description of search strategy, then the accuracy of search (in terms of the number of returned hits) is meaningful. Our results show that only 31 (95% CI: 22.78% to 40.63%) of the re-performed search yield results within 10% of number of hits reported. This result is consistent with poor search reproducibility reported by other studies [11, 23, 24]. An important reason for lack of search reproducibility is that the search URL is not given in all 100 (95% CI: 0.00% to 3.70%) of the examined SRs,

1. <https://pubmed.ncbi.nlm.nih.gov/?term=%22systematic+review%22%5btittle%5d+AND+meta-analysis%5btittle%5d&filter=simsearch2.fff&filter=pubt.systematicreview&filter=dates.2023/1/1-2023/12/31>
2. [https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=%22systematic+review%22%5btittle%5d%20AND%20meta-analysis%5btittle%5d%20AND%20%22free%20full%20text%22%5bfilter%5d%20%22systematic%20review%22%5bfilter%5d%20AND%20\(%222023/1/1%22%5bPDAT%5d%20:%20%222023/12/31%22%5bPDAT%5d\)&retmax=9999](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=%22systematic+review%22%5btittle%5d%20AND%20meta-analysis%5btittle%5d%20AND%20%22free%20full%20text%22%5bfilter%5d%20%22systematic%20review%22%5bfilter%5d%20AND%20(%222023/1/1%22%5bPDAT%5d%20:%20%222023/12/31%22%5bPDAT%5d)&retmax=9999)

leading to potential discrepancies in the interpretation of search strategy which eventuates to different implementation of the search. This is supported by Chin et al [24] in 2022 suggesting that only 22% reported the reviews gave the full Boolean search logic. Hence, we propose that search URLs should be given whenever possible, as search URL represents the actual implementation of search strategy.

<i>Criteria Type</i>	<i>Criteria</i>	<i>Percentage</i>	<i>Upper 95% CI</i>	<i>Lower 95% CI</i>
Reproducibility	Date of Search Given	23%	15.84%	32.15%
	Database Given	100%	96.30%	100.00%
	Search Terms Given	91%	83.77%	95.19%
	Date Range Given	76%	66.77%	83.31%
	Search URL Given	0%	0.00%	3.70%
	Ability to Repeat Search	100%	96.30%	100.00%
	Number of hits within 10% of Reported (PubMed)	31%	22.78%	40.63%
Transparency	Full List of Hits Given	0%	0.00%	3.70%
	Unambiguous Inclusion / Exclusion Criteria	95%	88.82%	97.85%
	Determine Which Article is Excluded by Which Exclusion Criterion	5%	2.15%	11.18%

Table 1: Summary of Results.

After obtaining the search results and in view of transparency, the full list of search results should be provided for both verification and as baseline for inclusion / exclusion as the reason for including or excluding a study is important [16, 25]. However, none of the 100 (95% CI: 0.00% to 3.70%) examined SRs provided the full list of search results. Hence, the search results cannot be objectively verified. Next, a set of unambiguous inclusion / exclusion criteria must be applied to all studies [26], and our result shows that 95 SRs (95% CI: 88.82% to 97.85%) meet this criteria. Nevertheless, given that none of the examined SRs provide a full list of search results, it is not unexpected that only a minority of the examined SRs - 5 SRs (95% CI: 2.15% to 11.18%) - provided sufficient information to determine which study is excluded by which exclusion criterion [16, 25]. Taken together, our result suggests a severe lack of transparency in the inclusion and exclusion of identified studies; hereby, propose that the full list of search results with inclusion / exclusion analysis should be given in all cases.

Conclusion

Our results support previous studies reporting poor reproducibility and transparency in SRs. Of the 100 examined SRs, only 31 SRs gave sufficient information for reconstructed PubMed searches yield search results within 10% of the hits reported, and only 5 SRs provided sufficient information to determine which study is excluded by which exclusion criterion. We propose that search URLs should be given whenever possible, and the full list of search results with inclusion / exclusion analysis should be given in all cases.

Supplementary Materials

Data file for this study can be downloaded at https://bit.ly/SR_Reproducibility_Transparency.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Jahan N., et al. "How to Conduct a Systematic Review: A Narrative Literature Review". Cureus 8.11 (2016): e864.

2. Clarke M and Chalmers I. "Reflections on the History of Systematic Reviews". *BMJ Evidence-Based Medicine* 23.4 (2018): 121-122.
3. Lasserson T, Thomas J and Higgins J. Chapter 1: Starting a review. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (Updated August 2023)*, eds Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V (Cochrane) (2023).
4. Faggion CM, Bakas NP and Wasiak J. "A Survey of Prevalence of Narrative and Systematic Reviews in Five Major Medical Journals". *BMC Medical Research Methodology* 17.1 (2017): 176.
5. Mallett R., et al. "The Benefits and Challenges of Using Systematic Reviews in International Development Research". *Journal of Development Effectiveness* 4.3 (2012): 445-455.
6. Furley P and Goldschmied N. "Systematic vs. Narrative Reviews in Sport and Exercise Psychology: Is Either Approach Superior to the Other?". *Frontiers in Psychology* 12 (2021): 685082.
7. Siddaway AP, Wood AM and Hedges LV. "How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses". *Annual Review of Psychology* 70.1 (2019): 747-770.
8. Baumeister R. *Writing a Literature Review. The Portable Mentor: Expert Guide to a Successful Career in Psychology*, eds Prinstein M, Patterson M (Springer, New York), 2nd Ed (2013): 119-32.
9. Powell JT and Koelemay MJW. "Systematic Reviews of the Literature Are Not Always Either Useful or the Best Way To Add To Science". *EJVES Vascular Forum* 54 (2022): 2-6.
10. Rethlefsen ML., et al. "Systematic Review Search Strategies are Poorly Described and Not Reproducible: A Cross-Sectional Meta-Research Study". *J Clin Epidemiol* 166 (2024): 111229.
11. Koffel JB and Rethlefsen ML. "Reproducibility of Search Strategies Is Poor in Systematic Reviews Published in High-Impact Pediatrics, Cardiology and Surgery Journals: A Cross-Sectional Study". *PLoS ONE* 11.9 (2016): e0163309.
12. Williams RI., et al. "Re-Examining Systematic Literature Review in Management Research: Additional Benefits and Execution Protocols". *European Management Journal* 39.4 (2021): 521-533.
13. Muka T., et al. "A 24-Step Guide on How to Design, Conduct, and Successfully Publish a Systematic Review and Meta-Analysis in Medical Research". *European Journal of Epidemiology* 35.1 (2020): 49-60.
14. Shaheen N., et al. "Appraising Systematic Reviews: A Comprehensive Guide to Ensuring Validity and Reliability". *Frontiers in Research Metrics and Analytics* 8 (2023): 1268045.
15. Greyson D., et al. "Systematic Review Searches Must Be Systematic, Comprehensive, and Transparent: A Critique of Perman et al". *BMC Public Health* 19.1 (2019): 153.
16. Hiebl MRW. "Sample Selection in Systematic Literature Reviews of Management Research". *Organizational Research Methods* 26.2 (2023): 229-261.
17. Uttley L., et al. "The Problems with Systematic Reviews: A Living Systematic Review". *Journal of Clinical Epidemiology* 156 (2023): 30-41.
18. Schuler GD., et al. "Entrez: Molecular Biology Database and Retrieval System". *Methods in Enzymology* 266 (1996): 141-162.
19. Sayers EW., et al. "Database Resources of the National Center for Biotechnology Information". *Nucleic Acids Research* 50.D1 (2022): D20-D26.
20. Wilson EB. "Probable Inference, the Law of Succession, and Statistical Inference". *Journal of the American Statistical Association* 22.158 (1927): 209-212.
21. Wallis S. "Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods". *Journal of Quantitative Linguistics* 20.3 (2013): 178-208.
22. Newcombe RG. "Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods". *Statistics in Medicine* 17.8 (1998): 857-872.
23. Rethlefsen ML., et al. "Systematic Review Search Strategies are Poorly reported and Not Reproducible: A Cross-Sectional Metaresearch Study". *Journal of Clinical Epidemiology* 166 (2024): 111229.

24. Chin JM, et al. "The Transparency and Reproducibility of Systematic Reviews in Forensic Science". *Forensic Science International* 340 (2022): 111472.
25. Rousseau DM, Manning J and Denyer D. "Evidence in Management and Organizational Science: Assembling the Field's Full Weight of Scientific Knowledge Through Syntheses". *Academy of Management Annals* 2.1 (2008): 475-515.
26. Meline T. "Selecting studies for systematic review: Inclusion and exclusion criteria". *Contemporary Issues in Communication Science and Disorders* 33.1 (2006): 21-27.

Volume 7 Issue 2 August 2024

© All rights are reserved by Maurice HT Ling, et al.