

# Conducting Pairwise and Network Meta-analyses in Updated and Living Systematic Reviews: a Scoping Review Protocol

Menelaos Konstantinidis<sup>1,2</sup> • Catherine Stratton<sup>1,3</sup> • Sofia Tsokani<sup>4</sup> • Julian Elliott<sup>5</sup> • Mark Simmonds<sup>6</sup> • Jessie McGowan<sup>7</sup> • David Moher<sup>7,8</sup> • Andrea C. Tricco<sup>1,2,3</sup> • Areti-Angeliki Veroniki Ph.D.<sup>1,2</sup>

<sup>1</sup>Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada, <sup>2</sup>Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada, <sup>3</sup>Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada, <sup>4</sup>Department of Primary Education, University of Ioannina, Ioannina, Epirus, Greece, <sup>5</sup>School of Public Health and Preventative Medicine, Monash University, Melbourne, Vic, Australia, <sup>6</sup>Centre for Reviews and Dissemination, University of York, York, United Kingdom, <sup>7</sup>School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada, and <sup>8</sup>Centre for Journalology, Clinical Epidemiology Program, the Ottawa Hospital Research Institute, Ottawa, ON, Canada

## ABSTRACT

**Objective:** The objective of this scoping review will be to describe existing guidance documents or studies reporting on the conduct of meta-analyses in updated systematic reviews (USRs) or living systematic reviews (LSRs).

**Introduction:** The rapid increase in the medical literature poses a substantial challenge in keeping systematic reviews up to date. In LSRs, a review is updated with a pre-specified frequency or when some other signalling criterion is triggered. While the LSR framework is well-established, there is uncertainty regarding the most appropriate methods for conducting repeated meta-analyses over time, which may result in sub-optimal decision-making.

**Inclusion criteria:** Studies of any design (including commentaries, books, manuals) providing guidance on conducting meta-analysis in USRs or LSRs.

**Methods:** We will use the JBI methodology for scoping reviews. We will search multiple medical bibliographic databases (Cochrane Library, Embase, ERIC, MEDLINE, *JBI Evidence Synthesis*, and PsycINFO), statistical and mathematics databases (COBRA, Current Index to Statistics, MathSciNet, Project Euclid Complete, and zbMATH), pre-print archives (Arxiv, BioRxiv, and MedRxiv), as well as difficult to locate/unpublished (or gray) literature. Two reviewers will independently screen titles, abstracts, and full-text documents, and extract data. Characteristics of recommendations for meta-analysis in USRs and LSRs will be presented using descriptive statistics and categorized concepts.

Details of this review project can be found in Open Science Framework: <https://osf.io/9c27g>

**Keywords:** living systematic review; meta-analysis; network meta-analysis; updated systematic review

*JBI Evid Synth* 2024; 00(0):1–9.

## Introduction

### Overview

Systematic reviews (SRs) often focus on collating the highest level of evidence from multiple sources, and provide a rigorous and transparent approach to inform decision-making and policy development.<sup>1,2</sup> By

synthesizing the available literature, an SR aims to provide comprehensive and unbiased findings on a given question.<sup>3,4</sup> In turn, SRs lead to more reliable conclusions and recommendations for practice and policy decision-making, ultimately resulting in better outcomes and more efficient allocation of resources.<sup>4,5</sup>

It is crucial for knowledge users, such as health care providers, patients, and policy-makers, to have

**Correspondence:** Areti Angeliki Veroniki, [aret-angeliki.veroniki@unityhealth.to](mailto:aret-angeliki.veroniki@unityhealth.to)

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ACT is a member of the JBI Evidence Synthesis editorial board but was not involved in any decisions related to this paper. The remaining authors declare no conflicts of interest.

DOI: 10.11124/JBIES-24-00279

up-to-date evidence in order to make well-informed decisions. However, the rapid increase in the medical literature poses a substantial challenge in maintaining the timeliness of SRs, as they may quickly become outdated.<sup>6,7</sup> The need for up-to-date evidence has been further demonstrated by the COVID-19 pandemic, during which decision-makers and policy-makers were required to make critical decisions based on rapidly evolving evidence. This has highlighted the importance of having high-quality and relevant information.<sup>8</sup> To address these issues, regardless of the clinical context, two principal frameworks for updating SRs have been proposed: updated systematic reviews (USRs) and living systematic reviews (LSRs).<sup>4</sup>

USRs are updates or revisions of previously published SRs that incorporate new evidence to maintain their relevance over time.<sup>4</sup> Undertaking USRs ensures that the conclusions and recommendations of these reviews are relatively up-to-date. USRs are often updated at pre-specified intervals once a substantial amount of new evidence has become available (although the definition of substantial new evidence may vary) or when indicated through a priority-setting exercise.<sup>9,10</sup>

In contrast, LSRs are a more dynamic approach to updating SRs, where SRs are continuously updated as new evidence becomes available.<sup>11</sup> Specifically, a LSR is continuously updated, with regular monitoring and screening of the literature until at least 1 of the following conditions is met:<sup>11-13</sup> (1) the SR is no longer a priority for decision-makers, (2) the certainty of the evidence is sufficiently high (eg, as assessed through the Grading of Recommendations, Assessment, Development and Evaluations framework<sup>14</sup>), or (3) there is likely to be no new research in the future.<sup>11</sup> LSRs aim to facilitate timelier responses for decision-makers by providing a constantly updated synthesis of the evidence.

### Methodological challenges

Regardless of which of the 2 frameworks is used, both offer valuable strategies for keeping SRs up to date. However, neither is without methodological challenges. This review will consider the challenges associated with conducting meta-analyses—either pairwise meta-analysis (PMA) or network meta-analysis (NMA)—within the context of USRs and LSRs. Briefly, PMA refers to the statistical techniques used to combine the results of multiple studies comparing 2 interventions.<sup>3,4</sup> In contrast, an NMA extends this

approach to simultaneously synthesizing evidence across multiple interventions.<sup>15,16</sup>

Repeatedly updating an SR by conducting a PMA at each update could create multiple problems, the most common of which is the inflation of the global type I error rate (ie, the probability of rejecting a null hypothesis when it is, in fact, true, or concluding that the effect estimate is statistically significant when, in truth, this can be attributed to chance or other unrelated reasons).<sup>17</sup> In addition to the type I error rate, in PMAs and NMAs, challenges around heterogeneity may arise. One approach to deal with heterogeneity would be to re-estimate it at each update. However, while the number of studies included in the meta-analysis remains small (eg, at early updates), the estimation of heterogeneity parameter would be poor. Even if the number of studies is not an issue for precisely estimating heterogeneity, the magnitude of heterogeneity between updates may increase or decrease. This may, in turn, impact the effect estimates between updates<sup>18</sup> as well as the required information size (ie, the amount of information, for example, in terms of the number of patients or studies required to mitigate the risk of a chance finding).<sup>17</sup> Such errors can have serious implications for the validity and reliability of findings from USRs and LSRs. This may lead to erroneous conclusions, and in turn, sub-optimal decision-making when decisions are based on statistical significance.<sup>18</sup> Decision-making based on statistical significance can also affect non-updated PMAs;<sup>4</sup> however, the inflation of type I and type II error rates in updated PMAs makes over-reliance on statistical significance particularly problematic.

In the context of NMAs, the issues relating to error rate inflation persist; however, there are additional complexities. Firstly, heterogeneity may be further complicated by the increasing number of interventions and loops in the network,<sup>18</sup> which may result in an increase or decrease in the network's heterogeneity.<sup>18</sup> Additionally, NMA requires that the distribution of the studies between each comparison should be similar with respect to key effect modifiers (transitivity assumption).<sup>4</sup> The statistical manifestation of transitivity is consistency, which requires that the estimates from direct evidence (i.e. estimates from head-to-head studies) in the network is in agreement with estimates from indirect evidence (ie, estimates inferred by the network geometry via a common comparator) in the network.<sup>4</sup> In the context of updated NMAs, the transitivity assumption would have to be assessed at each update. The problem with this is that in updating

a network, new interventions may be introduced, thereby possibly changing the transitivity assessment from being violated to being satisfied or vice versa. Similarly, although the consistency assumption can be tested at each update provided that there is at least one closed loop,<sup>19</sup> inference of inconsistency, especially at early updates, is likely to be problematic due to an insufficient number of studies for adequate power.<sup>18</sup> Furthermore, other aspects of the NMA such as reporting bias, including small-study effects, would need to be assessed at each update.

A commonly advocated method to deal with error inflation issues in USRs and LSRs is trial sequential analysis (TSA).<sup>18</sup> This adapts the ideas of sequential clinical trial design to meta-analysis.<sup>18</sup> It requires specification of a required information size (eg, the required number of participants to have a sufficiently powered PMA) and alpha spending function; that is, a function that dictates what proportion of a nominal significance level (eg, alpha of 0.05) can be used (as a function of the required information size) to establish statistical significance for benefit or safety. It can be similarly applied for establishing statistically significant futility. In the original formulation of the method, this method does not deal with heterogeneity; however, more recent extensions of the approach have been proposed to adjust the required information size by the heterogeneity. Moreover, in its application to NMAs, it does not account for transitivity or inconsistency.<sup>18,20,21</sup> All in all, establishing a threshold for sufficient power in PMAs remains a challenge.

Given the potential impact of repeating meta-analyses in USRs and LSRs, it is essential to develop a robust understanding of the methodological challenges associated with conducting PMAs and NMAs in these contexts, and how existing methods account or fail to account for them. Further, despite the existence of multiple approaches to deal with the underlying issues arising from repeated updates, there does not exist (to our knowledge) any consensus on what methods are appropriate for conducting PMA and NMA in USRs and LSRs. However, it is worth noting that Cochrane does not support the use of sequential methods for meta-analysis.<sup>22</sup>

A preliminary search of PROSPERO, MEDLINE, the Cochrane Database of Systematic Reviews, Open Science Framework, and *JB1 Evidence Synthesis* was conducted and no current or in-progress scoping reviews or systematic reviews on the topic were identified.

## Review question

What methods guidance has been described or disseminated for PMA and NMA in USRs and LSRs?

## Inclusion criteria

### Participants

This is not a relevant domain as we are interested in methodological guidance.

### Concept

The concepts of interest are the specific methods used to conduct PMA and NMA in USRs and LSRs, and how corresponding methods control issues arising from the updating process. Such methods may include TSA or Bayesian methods, with specific details (eg, selection of the alpha spending function and required information size).

### Context

Although this work is motivated by the clinical and health care contexts, we will not apply any limits to the context or setting (eg, geography or setting of specific interventions), nor discipline or field (eg, including the health sciences, economics, law, public health, environmental sciences, engineering, and social sciences) so as to be as inclusive as possible of any relevant studies. However, we expect that most documents will be concentrated in the health/medical sciences and social sciences. Moreover, we will not apply any limits on the type of LSR or USR (eg, SRs of interventions, diagnostic test accuracy).

### Types of sources

We will include any type of report, including meta-epidemiological studies, simulation studies, commentaries, discussion papers, books, editorials, handbooks, manuals, tutorials, or formal guidance from any organization as long as they are relevant to conducting PMA or NMA in the context of USRs or LSRs. However, we will exclude any applied USRs or LSRs, in which the primary objective is to answer a specific applied question (eg, a specific population, intervention, comparison, outcome), as such studies are unlikely to discuss the strengths and weaknesses of the selected methods. Moreover, the number of studies required for screening would likely result in an unmanageable number of titles to be screened, with little expected value beyond what is already provided in methodological reports.

## Methods

The proposed review will be conducted in accordance with the JBI methodology for scoping reviews.<sup>23</sup> Reporting of this protocol follows the guidance set forth by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for Protocols (PRISMA-P), while the final scoping review will be reported using the PRISMA extension for Scoping Reviews (PRISMA-ScR).<sup>24,25</sup>

### Search strategy

The search strategy (Appendix I) was developed by an experienced librarian (JMG) and peer-reviewed by another librarian using the Peer Review of Electronic Search Strategies (PRESS) checklist.<sup>26</sup> The search strategy will aim to identify published and unpublished guidance for PMA and NMA in USRs and LSRs. An initial exploratory search of MEDLINE was conducted to identify key articles and to refine the search strategy (ie, ensure that the search strategy captured the articles in question).

There will be no exclusions based on date, language, or publication status. All publications not available in English will be translated using DeepL (DeepL, Cologne, Germany) and evaluated for potential inclusion. Documents translated by DeepL will be reviewed by a person fluent in the original language to ensure that the translation is accurate. Bibliographic databases and unpublished and difficult to locate literature sources will be search from inception, and the date of search will be reported in the final review.

The bibliographic databases included in the search strategy via Ovid are as follows:

- MEDLINE (primary database)
- Cochrane Database of Systematic Reviews
- Cochrane CENTRAL
- Cochrane Methodology Register
- Embase
- ERIC
- *JBİ Evidence Synthesis*
- PsycINFO

We will search the following statistical and mathematical databases:

- COBRA: Collection of Biostatistics Research Archive
- Current Index to Statistics
- MathSciNet Database

- Project Euclid Complete
- zbMATH

Furthermore, we will search the following pre-print archives:

- arXiv
- BioRxiv
- MedRxiv

Lastly, we will conduct searches of unpublished and difficult to locate (or gray) literature across multiple sources:

- TRIP database
- Google
- Google Scholar
- Canadian Agency for Drugs and Technologies in Health (CADTH) Grey Matters tool
- Theses via the Center for Research Libraries Foreign Dissertation, DART-Europe E-theses Portal, Electronic Theses Online Service (ETHOS) | British Library, Networked Digital Library of Theses and Dissertations, open access dissertations, and Thesis Canada Portal, WorldCat.
- Known SR producer websites, for example, those of the Agency for Healthcare Research and Quality, CADTH, Campbell, Canadian Institutes of Health Research (CIHR), Centre for Reviews and Dissemination (CRD), Cochrane, L'Institut national d'excellence en santé et en services sociaux (INESSS), *JBİ Evidence Synthesis*, National Institute for Health and Care Excellence (NICE), and Sign.
- Other relevant sources will include the Center for EBM, EQUATOR Network, Health Quality Ontario Publications and OHTAC Recommendations, iloveevidence.com, and the libraries at Unity Health Toronto and the University of Toronto.

In addition, we will use our professional networks to identify additional eligible reports. Our systematic search process will incorporate a forward citation approach, expanding our exploration beyond primary databases to include relevant studies identified through citation tracking and reference lists.

### Study/Source of evidence selection

Citations identified through the search strategy will be imported into the Synthesi.SR software for

screening. Prior to title and abstract screening, all reviewers will undergo a training exercise via Zoom and proceed to a pilot exercise of 50 articles in which they will screen the titles/abstracts independently. Additional pilots, training, and discussion will be provided until the team of reviewers achieves  $\geq 70\%$  agreement. Once the pilot phase is complete, each identified record will be independently screened by 2 reviewers (from a wider set of reviewers).

The full text of potentially relevant documents identified in title and abstract screening will be retrieved. A second pilot training and exercise will be conducted for full-text screening on a set of 25 full-text studies. Similarly to the above pilot, additional piloting, training, and discussion will be provided until reviewers achieve  $\geq 70\%$  agreement on the 25 full-text articles. Once the pilot is complete, each article will be screened independently against the inclusion and exclusion criteria by 2 reviewers.

Throughout the screening process, disagreements between reviewers will be resolved through discussion between the reviewers, and where necessary, a third reviewer will arbitrate the decision for inclusion. The result of the search and screening will be summarized in a PRISMA 2020 flow diagram.<sup>27</sup> Author name, year of publication, citation, and reason for exclusion of records excluded at the full-text stage will be provided in the supplementary material of the final review manuscript. The number of studies excluded for a given reason will be included in the PRISMA 2020 flow diagram.

### Data extraction

Prior to undertaking the extraction, a pilot of 10 articles will be conducted to assess and calibrate the form. Two review authors will independently extract relevant information (see list below) from each publication using a standardized form (see Appendix II). The final form and guidance will be provided as an appendix to the final scoping review. Throughout the extraction process, disagreements between extractors will be resolved through discussion between the extractors, and where necessary, a third extractor will arbitrate the decision for inclusion.

We will extract the following data for each publication:

- Corresponding author name
- Corresponding author email address
- Year of publication
- Publishing source name

- Publication type (eg, book, report, commentary)
- Aim/objective (including what methodological issues the publication methods aim to solve)
- Field
- Type of SRs considered (ie, USR or LSR)
- Type of meta-analysis (eg, PMA or NMA)
- Type of methods considered (eg, TSA or Bayesian meta-analysis)
- Has an evaluation of the recommendations occurred (eg, empirical or simulation-based evaluation of how well the method or methods control the Type I error rate)?

We will extract the following details relating to the methodological guidance:

- Description of recommendation including specific methodological and implementation details (eg, selection of priors in Bayesian PMA or selection of the functional form for alpha spending function in trial sequential PMA)
- Advantages and disadvantages of the methods as reported by the author
- Intended impact on conducting USRs and LSRs
- Availability of code or other knowledge translation products
- Any additional notes.

### Data analysis and presentation

Synthesis of the extracted data will consist of a quantitative (descriptive) analysis and a high-level content analysis, as suggested by Pollock *et al.*<sup>28</sup> Briefly, in the case of the former, we will provide descriptive statistics for each characteristic that has been extracted (ie, mean and standard deviation for continuous variables, and frequencies and percentages for categorical variables). For the content analysis, 2 authors will independently categorize the following extracted items (also listed above):

- Type of method
- Description of recommendation, including specific methodological and implementation details
- Advantages and disadvantages of the methods, as reported by the author
- Description of evidence underpinning recommendation
- Intended impact on conducting USRs and LSRs.

We anticipate that qualitative items will be categorized using an inductive approach<sup>28</sup>; however, if



a deductive approach is taken, we will document this in the final review. In the end, we will have a well-defined categorization of the types of methods (and their characteristics) that are recommended for conducting PMA and NMA in USRs and LSRs. This will include their respective strengths and limitations, as reported in the included studies. We will describe the results textually and provide visuals where appropriate.

## Acknowledgments

Tamara Rader, Freelance Health Sciences Librarian, for peer-reviewing the MEDLINE search strategy.

## Funding

This work is supported by the Canadian Institutes of Health Research Project Grant (No. 507309). MK and CS are funded by the Canadian Institutes of Health Research's Canada Graduate Scholarship Doctoral Program. CS is also funded by the Pierre Elliott Trudeau Foundation Doctoral Scholarship. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis.

## Author contributions

MK, AAV, and ACT conceived and designed the study. MK wrote the first draft of the article, and edited it according to co-author comments. JM developed the search strategy. All authors critically reviewed, commented, and approved the final manuscript.

## References

1. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;318(7182):527-30.
2. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Developing guidelines. *BMJ* 1999;318(7183):593-96.
3. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to Meta-analysis*. Wiley; 2021.
4. Higgins JP, Chandler J, Cumpston M, Li T, Page M, Welch V. *Cochrane handbook for systematic reviews of interventions* version 6.3 (updated February 2022). Cochrane; 2022.
5. Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux P, Prasad K, *et al.* How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *Jama* 2014;312(2):171-79.
6. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010;7(9):e1000326.
7. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147(4):224-33.
8. Glasziou PP, Sanders S, Hoffmann T. Waste in COVID-19 research. *British Medical Journal Publishing Group*; 2020.
9. Moher D, Tsertsvadze A, Tricco A, Eccles M, Grimshaw J, Sampson M, *et al.* When and how to update systematic reviews. *Cochrane Database Syst Rev* 2008;1:MR000023.
10. Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J, *et al.* When and how to update systematic reviews: consensus and checklist. *BMJ* 2016;354.
11. Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, *et al.* Living systematic review: 1. Introduction—the why, what, when, and how. *J Clin Epidemiol* 2017;91:23-30.
12. Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JP, Mavergames C, *et al.* Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11(2):e1001603.
13. Simmonds M, Elliott JH, Synnot A, Turner T. Living systematic reviews. *Methods Mol Biol* 2022;2345:121-34.
14. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-26.
15. White IR. Network meta-analysis. *Stata J* 2015;15(4):951-85.
16. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21(16):2313-24.
17. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol* 2009;9(1):1-12.
18. Simmonds M, Salanti G, McKenzie J, Elliott J, Agoritsas T, Hilton J, *et al.* Living systematic reviews: 3. Statistical methods for updating meta-analyses. *J Clin Epidemiol* 2017;91:38-46.
19. Higgins JPT, Jackson D, Barrett J, Lu G, Ades A, White I. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Med* 2012;3(2):98-110.
20. Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Stats Med* 2011;30(9):903-21.
21. Nikolakopoulou A, Mavridis D, Egger M, Salanti G. Continuously updated network meta-analysis and statistical monitoring for timely decision-making. *Stat Methods Med Res* 2018;27(5):1312-30.
22. Brooker J, Synnot A, McDonald S, Elliott J, Turner T, Network LE. *Guidance for the Production and Publication of Cochrane Living Systematic Reviews: Cochrane Reviews in Living Mode*. Cochrane; 2019.
23. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, *et al.* Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid Synth* 2020;18(10):2119-26.

24. Moher D, Shamseer L, Clarke M, Ghera D, Liberati A, Petticrew M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4(1):1-9.
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, *et al.* PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169(7): 467-73.
26. McGowan J, Sampson M, Salzweid DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40-46.
27. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Int J Surg* 2021;88:105906.
28. Pollock D, Peters MD, Khalil H, McInerney P, Alexander L, Tricco AC, *et al.* Recommendations for the extraction, analysis, and presentation of results in scoping reviews. *JBI Evid Synth* 2023;21(3):520-32.

## Appendix I: Search strategy

*Ovid MEDLINE(R) ALL 1946 to July 06, 2023*

- 1 ((updat\* or up-to-date or maintain\* or living or statistical) adj3 (review\* or synthesis# or evidence)).tw,kf. (39 219)
- 2 (meta analys\* or meta-analys\* or metaanalys\* or meta regression or meta-regression or metaregression or nma or pw-ma).tw,kf. (275 635)
- 3 (meta analys\* or meta-analys\* or metaanalys\* or meta regression or meta-regression or metaregression or nma or pw-ma).tw,kf,mp. (308 713)
- 4 1 and (2 or 3) (6589)
- 5 (error or errors or trial sequential or “law of the iterated logarithm” or Shuster method or Tipton or false or false-negative or false-positive).tw,kf. (515 946)
- 6 method\*.ti. and method\*.ab. (324 454)
- 7 (checklist\* or consensus or approach\* or guide\* or guidance or handbook\* or toolkit\* or recommendation\*).tw,kf. (3 543 804)
- 8 exp \*statistics as topic/ or Computer Simulation/ or Data Interpretation, Statistical/ or (simulation or empirical or statistical analysis#s).tw,kf. or method\*.ab. /freq = 3 (1 380 313)
- 9 7 and 8 (347 007)
- 10 4 and (5 or 6 or 9) (620)



## Appendix II: Extraction form

Data item	Notes/options
1. Reviewer 1	Initials of the study extractor
2. Reviewer 2	Initials of the person verifying the data
3. Study	Study ID
4. Title	Manuscript title
5. Publishing source name	Eg, name of journal, website, or organization
6. Publication type	Book, report, etc.
7. Author	Last name of first author
8. Author email	
9. Country of study	Country of corresponding author
10. Year of publication	
11. Aim/objective of the publication	Include what methodological issues are aimed to be solved
12. Field	
13. Type of SR that the recommendations are for	USR, LSR, or both
14. Type of meta-analysis	PMA, NMA, or both
15. Type of method considered	Eg, text describing trial sequential analysis
16. Has an evaluation of the recommendation occurred? (Eg, empirical or simulation)	Yes/No/Not reported
17. If yes to Q16, provide text of description	
18. Description of recommendation, including specific methodological and implementation details	
19. Advantages and disadvantages, as described by the authors	
20. Intended impact on USRs or LSRs, as described by the authors	
21. Is any applicable code or data available?	Yes/No/Not applicable
22. Additional notes	