Article:
Konstantinidis, Menelaos, Kontouli, Katerina-Maria, Zhelnov, Pavel et al. (2026) Characterizing updated and living meta-analyses Part 1: A methodological scoping review on methods for conducting meta-analyses in updated and living systematic reviews. Journal of Clinical Epidemiology. 112389. ISSN: 0895-4356

https://doi.org/10.1016/j.jclinepi.2026.112389

Journal of Clinical Epidemiology 197 (2026) 112389

ORIGINAL RESEARCH

Characterizing updated and living meta-analyses part 1: a methodological scoping review on methods for conducting meta-analyses in updated and living systematic reviews

Menelaos Konstantinidis[a,b], Katerina-Maria Kontouli[c], Pavel Zhelnov[a,b], Catherine Stratton[d], Jessie McGowan[e], Mark Simmonds[f], David Moher[b,e,g], Andrea C. Tricco[a,b,d], Areti-Angeliki Veroniki[a,b,h,*]

[a]*Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, Ontario, Canada*
[b]*Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada*
[c]*Department of Primary Education, University of Ioannina, Ioannina, Epirus, Greece*
[d]*Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada*
[e]*School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada*
[f]*Centre for Reviews and Dissemination, University of York, York, UK*
[g]*Centre for Journalology, Methodological and Implementation Research Program, The Ottawa Hospital Research Institute, Ottawa, Ontario, Canada*
[h]*Center for Evidence Synthesis in Health, Department of Health Services, Policy, and Practice, Brown University School of Public Health, Providence, Rhode Island, USA*

Accepted 16 June 2026; Published online 22 June 2026

Abstract

Objectives: Systematic reviews (SRs) provide the highest level of evidence for clinical and policy decision-making but quickly become outdated as new studies emerge. Updated systematic reviews (USRs) and living systematic reviews (LSRs) offer 2 frameworks to maintain currency; however, statistical approaches for repeated updates of pairwise meta-analyses (PMAs) and network meta-analyses (NMA) remain heterogeneous. This scoping review aimed to identify, categorize, and summarize existing methods for conducting updated meta-analyses (primarily PMAs and, subject to available evidence, NMAs) in the context of USRs and LSRs, highlighting their properties, as well as their reported advantages and limitations.

Methods: A scoping review was conducted to identify studies that described or compared methods for updating PMAs or NMAs within USRs or LSRs. We followed the 2020 Joanna Briggs Institute methodology and searched across major bibliographic databases (MEDLINE, Embase, the Cochrane Library, Cochrane CENTRAL, the Cochrane Methodology Register, ERIC, and PsycINFO) from inception to March 6, 2024. Gray literature and expert recommendations supplemented our literature search. Two reviewers independently screened records, extracted data, and categorized methods.

Results: Of 2710 records identified, 51 studies and 1 companion report met the inclusion criteria. The studies were grouped into 9 methodological categories: Traditional/Cumulative Meta-analysis, Law of Iterated Logarithm, Shuster Sequential Method, Sequential Meta-Analysis, Trial Sequential Analysis, Stochastic Curtailment, Adaptive Design, Likelihood Ratio Test, and Bayesian meta-analysis. Most studies focused on USRs (72.5%) and PMAs (98%). Trial Sequential Analysis was the most frequently encountered approach (46/51), while only 1 study described an extension to NMA. Methods varied widely in error control and ability to adjust for heterogeneity. While most approaches addressed type I error, fewer considered type II error or heterogeneity.

Funding: This work is supported by the Canadian Institutes of Health Research Project Grant (No. 507309). M. K. and C.S. are funded by the Canadian Institutes of Health Research's Canada Graduate Scholarship Doctoral Program. C. S. is also funded by the Pierre Elliott Trudeau Foundation Doctoral Scholarship. A.C.T. is funded by a Tier 1 Canada Research Chair in Knowledge Synthesis for Knowledge Users.

* Corresponding author. Institute for Health Policy, Management, and Evaluation, University of Toronto; Center for Evidence Synthesis in Health, Department of Health Services, Policy, and Practice, Brown University; Affiliate Scientist, Knowledge Translation Program, 121 S Main St, Providence, RI 02903, USA.

E-mail address: argie_veroniki@brown.edu (A.-A. Veroniki).

https://doi.org/10.1016/j.jclinepi.2026.112389

Conclusion: This scoping review provides the first comprehensive mapping of methods for updated and living meta-analyses. Despite a range of available approaches, no single method demonstrated superiority across contexts. Further research should compare methods empirically, develop accessible software to support reliable, continuously updated evidence synthesis, and extend frameworks to NMA, especially as only a single methodological study was found for NMA.

Keywords: Living systematic review; Updated systematic review; Updated meta-analysis; Network meta-analysis; Trial sequential analysis; Type I error

1. Background

Systematic reviews (SRs) are commonly used to synthesize the highest level of evidence from multiple sources [1,2]. They follow a structured and transparent methodology to support evidence-based decision-making and policy development [1,2]. SRs aim to deliver thorough and objective answers to specific scientific questions by combining the existing body of research [3,4]. As a result, they provide a robust foundation for practice and policy recommendations, improving outcomes and ensuring more effective resource use [4,5].

Up-to-date evidence is vital for informed decision-making in medicine and health care, including health care providers, patients, and policymakers. However, the ever-growing volume of medical literature makes it increasingly difficult for SRs to stay current, as they can quickly lose relevance as new evidence emerges [6,7]. The urgency of this challenge became even more evident during the COVID-19 pandemic, where decision-makers needed to rely on rapidly changing evidence to make critical decisions, thereby highlighting the importance of timely access to high-quality, accurate, and relevant information [8]. To address the issue of outdated SRs, 2 main approaches have been proposed: updated systematic reviews (USRs) and living systematic reviews (LSRs) [4].

The term USR is often misused to refer to types of SRs that extend beyond the remit of USRs [9,10]. For example, Runjic et al [9] found that 64% of SRs used the term ''update'' to denote that the SR contained recent data. As such, for the present scoping review, we refer to the stricter definition of USR proposed by Runjic et al [9], namely, as SR in which ''*Authors clearly indicated in the manuscript which previous SR they were updating, backed up with a literature reference to the previous SR version.*'' In the context of USRs, updates are often scheduled at set intervals or triggered by the availability of substantial new findings, which may be identified through a prioritization process [11,12]. By updating and incorporating new evidence, USRs maintain the relevance of the SR, thereby ensuring subsequent recommendations or policies informed by the USR remain evidence based and reflect the current evidence [4].

In contrast to USRs, LSRs are a more agile strategy, where SRs are continuously revised as new evidence emerges [13]. This approach involves regular monitoring and updating of the literature and should continue until 1 of 3 conditions is met [13−15]: (1) the SR is no longer a priority for decision-makers, (2) the certainty of the evidence is deemed high enough (eg, using the Grading of Recommendations, Assessment, Development and Evaluations framework [16]), or (3) it is unlikely that further research will be conducted on the topic [13]. By maintaining a constantly updated synthesis of the evidence, LSRs aim to respond quickly to decision-makers' needs [13].

While USRs and LSRs provide effective frameworks for ensuring SRs remain current [17−19], challenges arise in the context of updated pairwise meta-analyses (PMAs) and network meta-analyses (NMAs). PMA refers to the statistical technique used to combine the results of multiple studies comparing 2 interventions [3,4]. In contrast, NMA extends this approach to simultaneously synthesize evidence across multiple interventions [20,21]. As the number of PMA or NMA updates increases, repeated significance testing (or inference) inflates the cumulative (across multiple updates) type I error rates (probability of rejecting the null hypothesis when it is in fact true) and type II error rates (probability of failing to reject the null hypothesis when it is in fact false) [17−19]. Moreover, increased heterogeneity inflates uncertainty and consequently increases the required information size (RIS; eg, number of patients or events) needed to achieve adequate statistical power [22−24].

Although multiple methods have been proposed to address these challenges, there is no consensus on the optimal approaches for conducting PMA and NMA in the context of USRs and LSRs or even whether or not the problem should be addressed at all (despite evidence to indicate the need to do so when making inferences from repeatedly updated evidence) [15,22,23,25−29]. For example, Cochrane does not endorse the use of sequential methods for meta-analysis [30]; however, trial sequential analysis (TSA) and other sequential methods are commonly advocated for by others [22,31−34]. Moreover, certain methods are better suited to specific purposes. For example, sequential methods are intended to support early stopping decisions rather than inferential validity across an indefinite number of updates [15,22,24]. In contrast, other methods such as Bayesian meta-analysis, while adaptable to such an objective, are not designed to inform early stopping [35].

To address these challenges, the present scoping review aims to identify and categorize existing methods for conducting updated PMA or NMA in the context of USRs

What is new?

Key Findings

- We synthesized methodological documents for conducting meta-analyses in updated and living systematic reviews. 51 articles and 1 companion report where identified. We organized methods into 9 broad methods, spanning a total of 16 methods and submethods.
- Most identified methods focus on controlling type I error, while fewer explicitly address type II error or heterogeneity.
- Trial Sequential Analysis was the most frequently encountered approach, whereas extensions to network meta-analysis were rare.

What this adds to what is known?

- Existing guidance on updating meta-analyses remains fragmented, with emphasis typically placed on type I error inflation. This review highlights a wider range of methods and their characteristics, including reported advantages and disadvantages.
- Sequential meta-analysis, is the only method which has been extended to NMAs.

What is the implication and what should change now?

- This synthesis may serve as a consolidated methodological evidence base against which to compare any newly developed methods and can be used as the basis for future methodological research on the topic of updated meta-analyses. Methodological work should now focus on characterizing meta-analyses of existing updated and living systematic reviews, comparing the statistical properties of the most promising methods identified in this review (eg, through an empirical study of existing updated meta-analyses and simulations). The eventual aim should be to develop a consensus-based recommendation for how best to update meta-analyses.

and LSRs and to summarize the evidence relating to their observed characteristics.

2. Methods

The protocol for this scoping review was developed in accordance with the updated 2020 Joanna Briggs Institute methodology for scoping reviews [36] and was registered on the Open Science Framework (https://osf.io/9c27g) [37,38]. Reporting follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidance [39] (Appendix A).

In this review, we included studies based on the following participants, concept, context criteria:

- Participants: studies reporting on methods or comparison of methods to conduct PMAs and NMAs in USRs and LSRs.
- Concept: methods used to conduct PMAs and NMAs in USRs and LSRs.
- Context: no limits to context, setting (eg, geography or clinical context), field (eg, economics, law, environmental sciences, engineering, and social sciences), and type of USR or LSR (eg, including review of interventions, prognosis, diagnostic test accuracy) were applied.

Full details of the methods have been described in the protocol registration and the published protocol [37,38]. A detailed summary is also provided in Appendix B along with the full search strategy (Appendices D and E) and screening and abstraction templates (Appendices F–H). A list of minor deviations from the protocol can be found in Appendix C.

3. Results

3.1. Literature search

From a total of 2710 records, 628 were duplicates, resulting in 2082 titles and abstracts to be screened. We excluded 1814 titles and abstracts, yielding 268 full-text reports. Of the 268 full-text reports, we excluded 140 because they were not relevant to the conduct of USRs or LSR (ie, they were applied reviews), 75 because they did not report on methods for updating a PMA, and 1 because we could not retrieve the full text (Appendix J for the full list of excluded studies). Subsequently, 51 original studies [11,14,15,22,23,25,26,31–35,40–78] and 1 companion report [79] were included (see Appendix I). The study selection process is summarized in the PRISMA-ScR diagram (Fig 1).

3.2. Study characteristics

Among the 51 studies, the earliest identified publication was in 1997 (Fig 2). The median number of studies published per year was 3 (IQR: 2-3.5), ranging from 0 to 7. Most studies (with respect to the first author's primary affiliation) were conducted in Europe (30/51, 58.8%), followed by North America (18/51, 35.3%), and a minority in Asia (1/51, 2.0%), Australia (1/51, 2.0%), and South America (1/51, 2.0%). Of the 51 included studies, 39 (76.5%) were published in journals, 6 (11.8%) in books, 3 (5.9%) in conference proceedings, and 3 (5.9%) in theses. Most studies

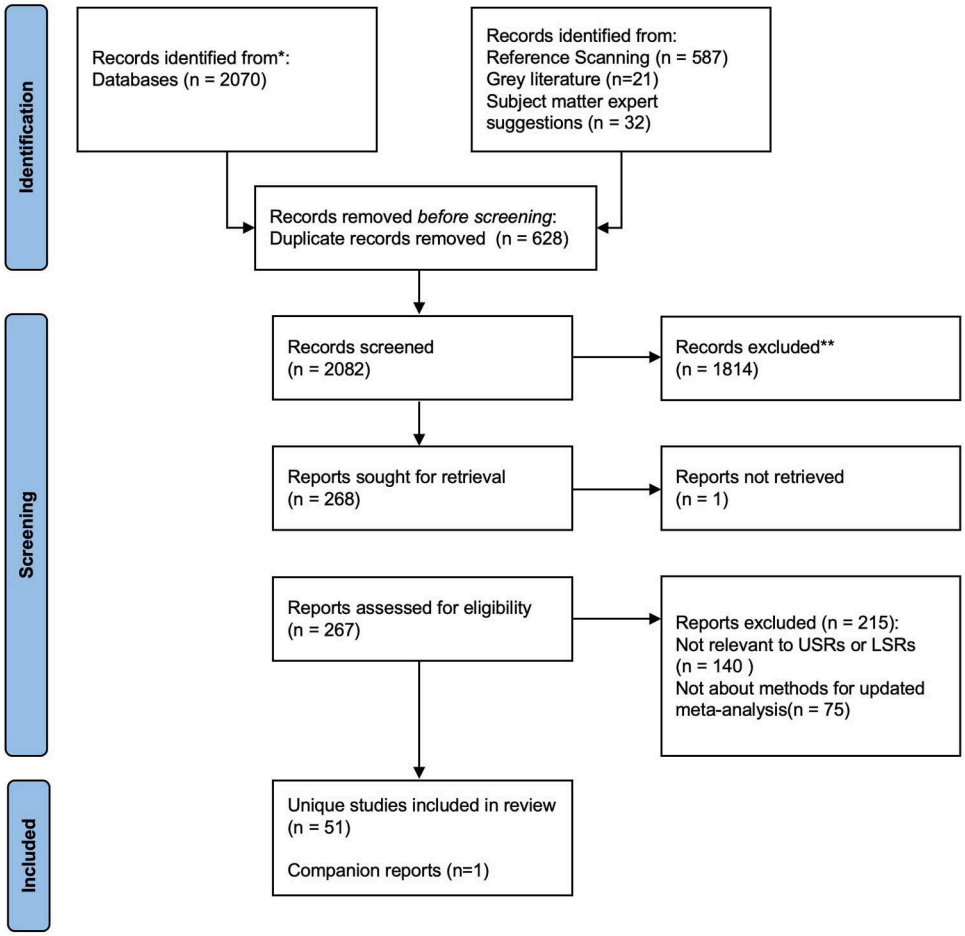

Figure 1. PRISMA diagram.

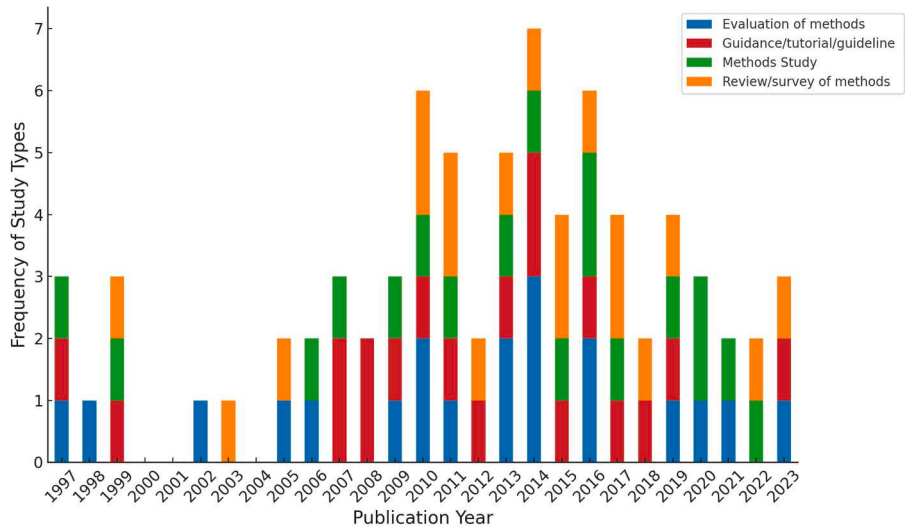

Figure 2. Bar chart showing the number of studies each year stratified by study type.

Table 1. Characteristics of the identified methods

Method	Adaptive design	Bayesian methods	LIL	LRT	Shuster	SMA	Stochastic curtailment	Traditional MA/CMA	TSA
Frequency	4	13	15	2	5	26	2	18	46
Review setting									
USR	4 (100%)	11 (84.6%)	11 (73.3%)	1 (50%)	2 (20%)	15 (68%)	2 (100%)	16 (88.8%)	34 (74%)
LSR	0 (0%)	2 (15.4%)	3 (20%)	0 (0%)	3 (60%)	6 (23%)	0 (0%)	2 (11.2%)	6 (13%)
Both	0 (0%)	0 (0%)	1 (6.6%)	1 (50%)	0 (0%)	5 (19%)	0 (0%)	0 (0%)	6 (13%)
Type of meta-analysis									
PMA	4 (100%)	10 (76.9%)	10 (66.6%)	2 (100%)	3 (60%)	20 (77%)	2 (100%)	16 (88.8%)	35 (76%)
NMA	0 (0%)	1 (7.7%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)	0 (0%)
Both	0 (0%)	2 (15.4%)	5 (33.3%)	0 (0%)	2 (40%)	5 (19%)	0 (0%)	2 (11.2%)	11 (24%)
Model structure									
Common effect	0 (0%)	0 (0%)	1 (6.6%)	1 (50%)	0 (0%)	1 (4%)	2 (100%)	0 (0%)	5 (11%)
Random effects	1 (25%)	1 (7.7%)	4 (26.6%)	0 (0%)	0 (0%)	7 (27%)	0 (0%)	6 (33.3%)	7 (15%)
Both	3 (75%)	92.3%(12)	10 (66.6%)	1 (50%)	5 (100%)	18 (69%)	0 (0%)	12 (66.7%)	34 (74%)
Outcome type									
Binary	0 (0%)	0 (0%)	2 (13.3%)	1 (50%)	0 (0%)	2 (8%)	0 (0%)	2 (11.1%)	8 (17%)
Continuous	0 (0%)	1 (7.7%)	2 (13.3%)	0 (0%)	0 (0%)	3 (12%)	0 (0%)	7 (38.9%)	3 (7%)
Any	4 (100%)	12 (92.3%)	11 (73.3%)	1 (50%)	5 (100%)	21 (80%)	2 (100%)	9 (50.0%)	35 (76%)
Effect size considerations[a,b]									
Yes	4 (100%)	11 (84.6%)	12 (80%)	2 (100%)	4 (80%)	20 (77%)	2 (100%)	8 (44.4%)	39 (85%)
No/NA	0 (0%)	2 (15.4%)	3 (20%)	0 (0%)	1 (20%)	6 (23%)	0 (0%)	10 (55.6%)	7 (15%)
Heterogeneity considerations[a]									
Yes	1 (25%)	3 (23.1%)	5 (33.3%)	1 (50%)	1 (20%)	14 (54%)	2 (100%)	4 (22.2%)	20 (44%)
No/NA	3 (75%)	10 (76.9%)	10 (66.6%)	1 (50%)	4 (80%)	12 (46%)	0 (0%)	14 (77.8%)	26 (56%)
Transitivity considerations[a]									
Yes	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (12%)	0 (0%)	0 (0%)	0 (0%)
No/NA	4 (100%)	13 (100%)	15 (100%)	2 (100%)	5 (100%)	23 (88%)	2 (100%)	18 (100%)	46 (100%)
Consistency considerations[a]									
Yes	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (12%)	0 (0%)	0 (0%)	0 (0%)
No/NA	4 (100%)	13 (100%)	15 (100%)	2 (100%)	5 (100%)	23 (88%)	2 (100%)	17 (100%)	46 (100%)

(*Continued*)

Table 1. Continued

Method	Adaptive design	Bayesian methods	LIL	LRT	Shuster	SMA	Stochastic curtailment	Traditional MA/CMA	TSA
Method of evaluation									
Simulation	3 (75%)	1 (7.7%)	5 (33.3%)	0 (0%)	2 (40%)	2 (8%)	0 (0%)	7 (38.9%)	1 (2.1%)
Empirical study	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)	3 (6.4%)
Example	4 (100%)	4 (30.8%)	6 (40%)	1 (50%)	2 (40%)	6 (24%)	1 (50%)	11 (61.1%)	16 (30%)
Citations	[65,66,80]	[14,31,34,39,40,49,51,53,55,65,66,72,78,80]	[33,40,42,44,50,54,55,65,66,68,69,72,75,76,79,80]	[45,65,70]	[44,54,67,68,76]	[14,15,22,23,34,42–44,49–51,54,57,58,65–71,75–78,80]	[44,63]	[11,22,39,41,42,44,54–57,59,61,73,74]	[11,15,23,24,31,32,34,40–43,47,48,50–55,58,60–72,74–76,78–80]

LIL, Law of Iterated Logarithm; LRT, likelihood ratio test; LSR, living systematic reviews; NA, not applicable; NMA, network meta-analyses; PMA, pairwise meta-analyses; RCI, repeated CI; SMA, sequential meta-analysis; TSA, trial sequential analysis; USR, updated systematic reviews.

[a] Important considerations of methods described by authors.

[b] This variable was abstracted as any information relating to biases in the effect size resulting from the updating process. Examples include the inflation of the type I and type II error rates.

failed to report their source of funding (26/51, 51.0%), followed by being publicly sponsored (23/51, 45.1%); and a minority reported no funding for the research (2/51, 3.9%).

3.3. Methods characteristics

From the 51 studies, through discussion, the identified methods were categorized into 9 broad categories for updating PMAs and NMAs: Traditional Meta-Analysis/Cumulative Meta-Analysis [11,22,40,42,43,45,55–58,60,62,72,73], Law of Iterated Logarithm [25,34,41,43,45,51,55,56,66,67,69,71,74,75,78,79], Shuster Sequential Method [25,55,68,75], Sequential Meta-Analysis (SMA) [14,15,22,23,25,26,35,43–45,50–52,55,58,59,66–70,74–77,79], TSA [11,15,23,25,26,31–33,35,41–44,48,49,51–56,59,61–71,73–75,77–79], Stochastic Curtailment [45,64], Adaptive Design [66,67,79], Likelihood ratio test (LRT) [46,70], and Bayesian meta-analysis [14,31,35,40,41,50,52,54,56,66,67,71,77,79] (Table 1). In addition, repeated CIs (a way of presenting SMA and TSA) were reported in 3 studies [46,66,70]. The median frequency each method was encountered in the included studies was 13 (IQR: 4-18), with a minimum of 2 (Stochastic Curtailment) and a maximum of 46 (TSA) studies (Table 1).

Most of the studies were reported in the context of USRs (37/51, 72.5%) and PMAs (50/51, 98%). Most methods (6/9, 66.7%) were described for both common effect and random effects analyses. Furthermore, all methods (9/9, 100%) were developed for any type of outcome (or were agnostic of the outcome type) and reported some aspect of bias or challenges in estimating the effect size arising from repeated analysis.

Of the studies reporting on Adaptive design, 75% assessed the method in a simulation study. The rest of the methods had a relatively low proportion of studies assessing them in a simulation study, ranging from 0% (LRT, Stochastic Curtailment) to 40% (Shuster) (Table 1).

3.4. Description of methods

Among the 9 broad methodological categories, we identified 16 distinct methods and submethods (Table 2). Traditional PMA reestimates the pooled effect and between-study heterogeneity each time new studies are added, treating each update as a separate meta-analysis. In contrast, cumulative meta-analysis (CMA) orders studies chronologically and sequentially incorporates them one by one, recalculating the pooled estimate and heterogeneity at each step to illustrate the evolution of evidence over time [11]. Two CMA submethods were identified: the stricter alpha method, which applies a lower significance threshold to mitigate inflation of the type I error rate, and the two-stage CMA [42], which fixes the heterogeneity estimate after an initial stage and uses that stage's pooled effect as the null hypothesis for subsequent updates [50].

Table 2. Detailed description of the methods and submethods identified in this review for conducting meta-analyses in USRs and LSRs

Method	Submethod	Description
Adaptive design	*P value combination*	Combines stage-specific *P*-values from successive cumulative meta-analyses using a predefined combination rule (eg, Fisher's method). This approach permits design modifications between updates—such as adding studies or altering information size—while preserving the overall type I error rate [65,66,80].
	Proschan & Hunsberger	Extends the *P* value combination framework by using a *conditional error function* to recalculate critical boundaries after interim analyses. This quantifies the remaining type I error under the null hypothesis, ensuring valid inference following adaptive modifications [80].
	Weighted combination	Generalizes the *P*-value combination method by combining stage-specific *z*-scores rather than *P*-values, with prespecified information-based weights (eg, proportional to sample size or number of events). The resulting weighted sum, normalized by the square root of the sum of squared weights, follows an approximately standard normal distribution under the null hypothesis and preserves the overall type I error while improving efficiency when information per update differs [65].
Bayesian	—	Incorporates prior distributions for the effect and heterogeneity parameters, updating posterior distributions with each new study. The posterior from each update becomes the prior for the next, allowing continuous evidence updating [34].
Law of Iterated Logarithm	—	Adjusts the conventional *z*-statistic with a penalty term to maintain the nominal type I error rate across repeated updates, ensuring that the statistic remains within predefined limits (eg, ± 1.96 for 95% CIs) [33,50].
Likelihood ratio test	—	Uses likelihood ratios instead of *P*-values to evaluate hypotheses. Likelihood ratios from each study are multiplied at each hypothesized effect size to derive a summary test statistic [45,70].
Shuster sequential method	—	Uses prespecified stopping times based on the number of included trials. Stopping occurs if either the maximum planned number of trials is reached or the cumulative *z*-score exceeds the boundary at a given update [67].
Sequential meta-analysis	*Basic Approach*	Aims to provide valid inference when accrued information is less than or equal to the required information size. Not designed for indefinite updating of a meta-analysis. At each update, the pooled treatment effect, its variance, and the corresponding *z*-score are recalculated. These statistics are compared against prespecified stopping boundaries defined by a spending function that allocates the allowable type I and type II errors (commonly 0.05 and 0.20). The required information size (RIS) determines when the full error allowance is reached. If the *z*-score and inverse variance cross a boundary before the RIS is met, the analysis can be stopped early, indicating that additional evidence is unlikely to change the conclusion. Boundaries identified in the included studies included O'Brien—Fleming, triangular and inverse triangular, double triangular, and Bolland safety boundaries, each suited to different contexts (eg, superiority, inferiority, equivalence, or harm).
	Semi-Bayes	Places a prior on the between-study heterogeneity parameter (eg, based on prior PMAs). The Bayesian estimate of heterogeneity is updated at each iteration and then used as a fixed value in a frequentist PMA [22]. The authors of the approach also derive an approximate semi-Bayes approach, which avoids the need for numerical integration [22].
	NMA-SMA	Extends the SMA framework to network meta-analysis by defining a spending function and monitoring boundaries for each pairwise comparison, determining significance when a comparison-specific *z*-score crosses its boundary [62].
Stochastic curtailment	—	Estimates the probability that an apparent trend (eg, treatment benefit) will reverse with additional data, assuming no true difference between treatment groups. If the probability of reversal is low, further updating may be curtailed [63].

(*Continued*)

Table 2. Continued

Method	Submethod	Description
Traditional PMA/cumulative meta-analysis	*Basic approach and CMA*	Aims to provide valid inference when accrued information is less than or equal to the required information size. Not designed for indefinite updating of a meta-analysis. Traditional PMA reestimates pooled effects and heterogeneity with each update. CMA adds studies sequentially in chronological order, recalculating summary effects to visualize evidence accumulation [57].
	Stricter alpha	Applies a more conservative alpha threshold at each update to control the type I error rate during repeated testing [41].
	Two-stage cumulative meta-analysis	Estimates the effect estimate and heterogeneity using the first (chronologically) 5-10 studies (stage 1)[52]. Any updates after stage 1 use a fixed heterogeneity (ie, as estimated in stage 1). Moreover, the summary effect size estimated in stage 1 is used as the null hypothesis for hypothesis testing at later stages.
TSA	*Pogue and Yusuf*	Defines monitoring boundaries using the O'Brien–Fleming function and assumes the required information size equals that of a large, adequately powered trial for the same interventions and assumed effect [63,64].
	Contemporary TSA	Extends the Pogue and Yusuf approach by adjusting the required information size and monitoring boundaries for heterogeneity using I^2 or D^2. D^2 quantifies the proportional reduction in variance when switching from a random-effects to a common-effects model [24].

Several methods explicitly address repeated testing. The Law of Iterated Logarithm (LIL) modifies the conventional z-statistic with a penalty term to maintain the nominal type I error rate across repeated updates [34,51]. The Shuster Sequential Method defines stopping boundaries a priori based on the number of planned updates, terminating the analysis if the cumulative z-score exceeds the boundary or the maximum number of updates is reached [7]. SMA distributes the overall type I and II error rates across updates using a spending function and compares cumulative z-scores against prespecified stopping boundaries until the RIS is achieved [22,74,76,77]. Four key boundary types were identified: O'Brien–Fleming (superiority), triangular and inverse triangular (superiority or inferiority), double triangular (superiority, inferiority, or equivalence), and Bolland safety (harm detection; Table 3 for details regarding the 4 boundaries) [77]. Extensions include the semi-Bayes approach [22], which stabilizes heterogeneity estimation using an informative prior, and sequential network meta-analysis (SNMA), which applies an analogous framework across multiple pairwise comparisons within a treatment network [63].

On the other hand, TSA uses the Lan–DeMets O'Brien–Fleming spending function to establish monitoring boundaries [31,64,65]. The Pogue and Yusuf variant estimates the RIS based on assumptions from large, adequately powered trials [64,65], while contemporary TSA adjusts for between-study heterogeneity using I^2 or D^2 (Diversity) statistics to refine boundary estimation (see Table 2 for a description of D^2) [31,32]. Both SMA and TSA can reduce the necessary number of patients required when futility boundaries are incorporated, as updating of meta-analyses can stop when there is a statistically significant indication of futility [24], thereby potentially reducing costs and the number of patients unnecessarily subjected to inefficacious interventions.

Another method designed to manage repeated testing includes stochastic curtailment, which evaluates the probability that an observed trend would change with further updates and permits early stopping if that probability is low [64].

Within the broader adaptive design category [66,67,79], 3 submethods were identified: the *P*-value combination, Proschan and Hunsberger, and weighted combination approaches. All 3 share a common foundation in adaptive combination testing, which allows evidence from successive cumulative meta-analyses to be combined while preserving overall type I error control, even when design modifications are made between updates. The *P*-value combination method aggregates update-specific *P*-values using a predefined combination rule. The Proschan and Hunsberger method extends this framework by using a conditional error function (Table 3) to recalculate critical boundaries following interim analyses, and the weighted combination method generalizes the approach by combining z-scores with information-based weights (ie, weights proportional to the inverse of the variance of the meta-analysis effect estimate).

Finally, the LRT approach combines likelihood ratios across studies to derive a cumulative test statistic without relying on *P*-values [46,70], while Bayesian meta-analysis iteratively updates prior distributions for the treatment effect and heterogeneity with each new study, allowing posterior estimates to evolve dynamically as evidence accumulates [35].

Collectively, these methods reflect a continuum of strategies aimed at preserving statistical power control,

Table 3. Summary of monitoring boundaries identified for use in sequential meta-analysis based on Whitehead et al, 2002 [78]

Boundary type	Description
Restricted design	This stops early only if one treatment is substantially superior to the other; it does not stop for futility. It is suitable when the *full sample size* is required to assess other outcomes, such as safety, while still permitting early stopping for strong benefit.
Triangular boundary	Allows early stopping when there is sufficient evidence that the new treatment is significantly better than control, or when it is unlikely that a difference will be shown (futility). It is most appropriate for *efficacy* outcomes, when interest lies in demonstrating superiority.
Reverse triangular boundary	Same as triangular boundary but for inferiority.
Double triangular boundary	Formed by combining a triangular and reverse triangular test, this design has high power to detect both *superiority and inferiority* and can also be used to assess *equivalence* when the estimated treatment effect lies within a prespecified range. It is efficient for efficacy outcomes when both benefit and harm are of interest.
Bolland safety boundary	Intended for *safety* variables, this design recommends stopping as soon as there is sufficient evidence that the new treatment is worse than control. If no harm is shown, data collection continues to meet the power requirement for the primary efficacy variable.

validity, and interpretability in the context of repeatedly updated evidence synthesis. Specific details regarding each of the methods and submethods can be found in Table 2.

3.5. Properties of the methods

In Figure 3, we summarize the central properties of each method, focusing on the advantages and disadvantages as reported across the included studies. We summarize the findings here.

3.5.1. Error control

All the methods and submethods except for traditional PMA and stochastic curtailment address type I error. However, as heterogeneity increases, it becomes difficult to control for type I error using adaptive designs [79]. Furthermore, using the LIL method, when the number of updates is large, the method may have a limited impact on controlling the type I error [34,51].

In contrast, type II error was reported less frequently in the literature—it was not accounted for in the LIL, Shuster, Stochastic curtailment, traditional MA/CMA methods, or the Pogue & Yusuf TSA approach.

3.5.2. Validity and adjustment

The 2-stage CMA partially controls for type I and type II error rates and does not require an RIS, thereby providing unbiased summary effect estimates at each update [58]. However, 2-stage CMA has poor control of error rates when heterogeneity is large [58].

The LIL, Shuster, traditional PMA/CMA, traditional SMA, and stochastic curtailment fail to account for heterogeneity. In the case of SNMA, several approaches have been proposed for incorporating heterogeneity, including a semi-Bayes approach with external heterogeneity estimates [63]; however, further work is needed to determine what the optimal approach is. A related issue arises regarding robustness against the misestimation of heterogeneity, which can be particularly problematic at early updates (ie, when few studies are included) [22]. The only methods that explicitly deal with heterogeneity misestimation completely or partially are the Bayesian method, Shuster, semi-Bayes SMA, and SNMA.

The last essential characteristic relates to whether the information fraction increases as new studies are added. Ordinarily, this is expected as more studies and patients would usually increase the precision. However, when information-based sequential approaches are applied, increased heterogeneity can cause the information fraction to decrease [15,25].

3.5.3. Extensions and advanced considerations

SMA was the only method identified for updating an NMA [63]. It was noted that (1) heterogeneity in NMA may increase or decrease as more interventions and loops are added [25]; (2) SNMA assumes consistency between direct and indirect evidence, which may not hold in all updates, thereby rendering the analysis biased; (3) as new studies are added, a shift in the network may occur from satisfying assumptions to violating them; and (4) at early updates, small sample sizes often hinder the power needed for reliable inference about consistency and heterogeneity [25].

While most of the models can be defined analytically, the LIL methods, Bayesian method, and semi-Bayes SMA require simulation- and sampling-based approaches, thereby requiring more advanced software.

4. Discussion

This review identified and categorized existing methods for conducting pairwise and network meta-analyses in the context of USRs or LSRs. Across 51 studies, 9 broad methodological approaches were identified, ranging from

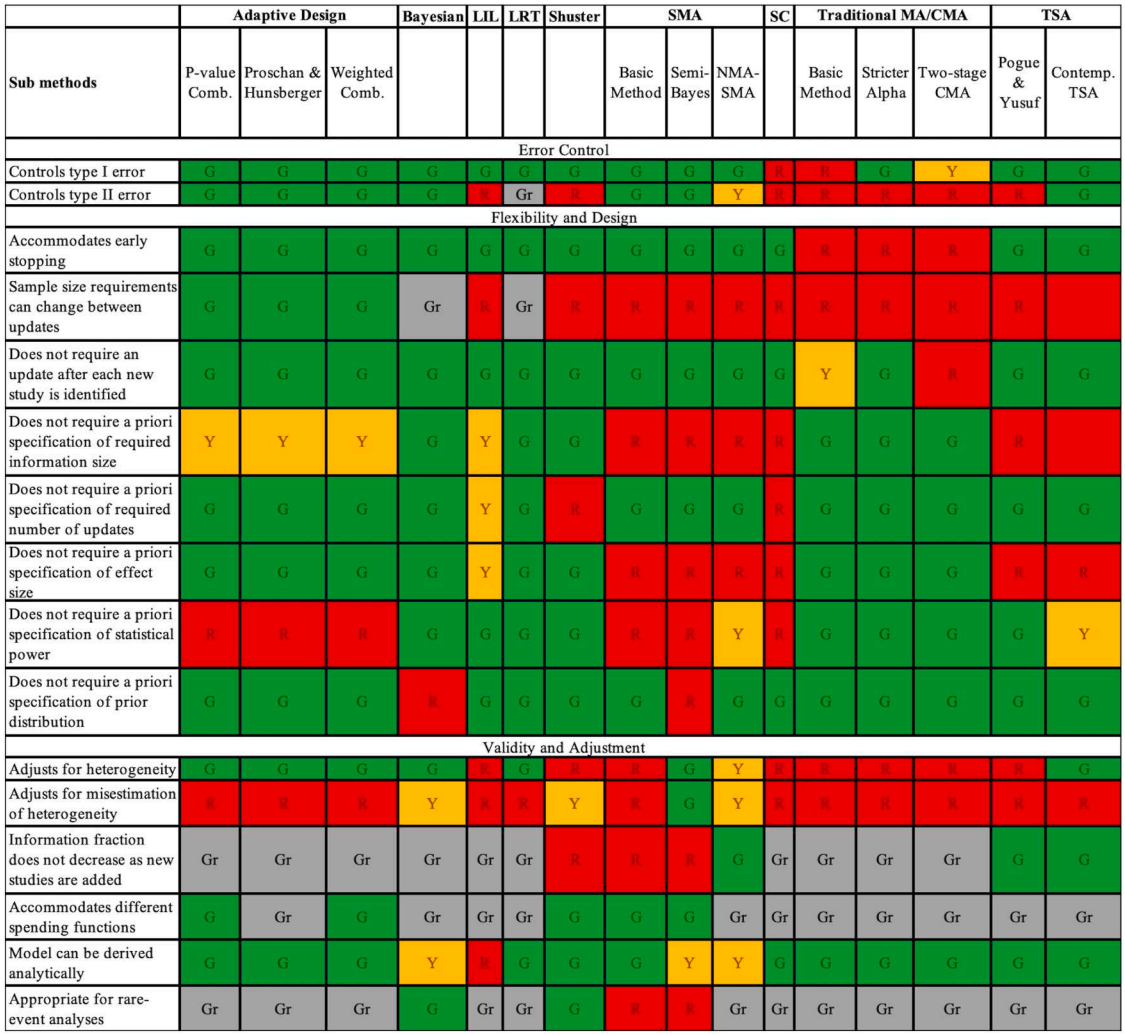

Figure 3. Figure comparing properties of each identified method and submethod as reported by articles referencing each method Green indicates that at least one included study reported this property for the specific method/sub-method. Yellow indicates that at least 1 included study partially reported this property for the specific method/submethod. Gray indicates that no study reported this property for the specific method/submethod. Red indicates that at least 1 included study reported that the method/submethod does not contain this property. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

traditional cumulative to more advanced sequential approaches and adaptive frameworks. Despite the number of approaches, consensus remains elusive regarding the optimal strategy for repeatedly updating the meta-analyses in USRs and LSRs. Nonetheless, many authors advocate for sequential approaches—especially TSA—to mitigate inflated type I and type II error rates. A striking result of this scoping review was that we only identified a single study developing the topic of updated NMAs. This remains an underexplored area, possibly because of the complexities discussed in Section 3.5.

It is also worth noting that Bayesian methods do not solve the problem explicitly. Rather, Bayesian meta-analysis redefines the inferential problem by incorporating prior specification, posterior uncertainty estimation, and decision thresholds, such that quantities such as type I error are not applicable. Beyond the Bayesian context, the availability of software for updating meta-analyses remains a challenge. The included studies provided software or code infrequently for the proposed or evaluated methods, with the most available being for TSA (through the Copenhagen Trial Unit) [80]. More recently, the RTSA package was also released for undertaking TSA in R [81]. Overall, and especially in light of the lack of an optimal strategy for updating, authors of SRs should clearly articulate their strategy for dealing with multiplicity at the protocol stage. Such guidance has already been provided by Asante et al [82] when TSA is the intended strategy; however, this degree of transparency and forethought should also be the default when other strategies are selected.

This review has several strengths. To our knowledge, it is the first to systematically map methodological approaches for updated and living meta-analyses. The review followed a registered protocol, used a peer-reviewed search strategy, and adhered to Joanna Briggs Institute methodology for scoping reviews and PRISMA-ScR reporting

guidelines. In addition, it offers a structured summary of methodological properties, strengths, and limitations.

Several limitations should also be acknowledged. First, some relevant studies may have been missed due to database indexing limitations, and our eligibility criteria were restricted to studies explicitly addressing updated or living meta-analyses, potentially excluding adaptable stand-alone methods. Second, we excluded applied USRs and LSRs; as such, it is possible that additional methods were described in a subset of those USRs or LSRs. However, we expect any resulting selection bias to be minimal. Third, few studies provided implementation guidance or software, limiting reproducibility and real-world uptake of the identified methods. Fourth, the date of our last search was over a year ago. However, to the best of our knowledge, no new methodological approaches for updated and living meta-analyses have been published since that time; as such, we consider the review to remain current.

Future work should prioritize developing and validating methods for updated NMA, empirically comparing candidate approaches, and integrating these into accessible software packages. Simulation studies are also needed to assess the statistical performance of competing methods under realistic update scenarios (including the ability to manage type I and II error rates), alongside empirical evaluations of how these methods are applied in practice. These are ongoing studies that our team is presently working on. Moreover, a study evaluating the barriers and facilitators contributing to the wide adoption of methods for updating meta-analyses would be a valuable contribution to the literature. Collectively, these efforts will help establish best-practice standards for repeated evidence synthesis in dynamic evidence environments.

5. Conclusion

This scoping review highlights various methods for continually updating PMAs and NMAs within USRs and LSRs. Each approach offers distinct strengths and limitations, often striking different balances between controlling type I error, type II error, and evolving heterogeneity. No method was universally reported as superior based on the properties we described and according to what was reported in the original studies. As such, further work is needed to have a strong evidence base for a concrete recommendation. Future work should focus on head-to-head method comparisons, developing clearer guidance on selecting and reporting sequential approaches, and developing methods for the NMA context, all of which will enable more rigorous and timely evidence syntheses that inform policy and practice.

CRediT authorship contribution statement

Menelaos Konstantinidis: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Katerina-Maria Kontouli:** Writing – review & editing, Validation, Investigation, Formal analysis, Data curation. **Pavel Zhelnov:** Writing – review & editing, Software, Investigation, Data curation, Conceptualization. **Catherine Stratton:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis. **Jessie McGowan:** Writing – review & editing, Validation, Methodology, Funding acquisition, Conceptualization. **Mark Simmonds:** Writing – review & editing, Validation, Methodology, Funding acquisition, Formal analysis. **David Moher:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Andrea C. Tricco:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Areti-Angeliki Veroniki:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

A.-A.V. reports financial support was provided by Canadian Institutes of Health Research Project Grant (507309). M.K. reports financial support was provided by Canadian Institutes of Health Research Canada Graduate Scholarship Doctoral Program and relationship with Alexion Pharmaceuticals Inc that includes: employment. C.S. reports financial support was provided by Canadian Institutes of Health Research Canada Graduate Scholarship Doctoral Program and Pierre Elliott Trudeau Foundation Doctoral Scholarship. A.C.T. reports financial support was provided by Tier 1 Canada Research Chair in Knowledge Synthesis for Knowledge Users.

A.A.V. is a statistical editor for the Journal of Clinical Epidemiology but had no involvement with the peer review process or decision for publication. J.M. is an editor for the Journal of Clinical Epidemiology but had no involvement with the peer review process or decision for publication. D.M. is on the editorial board of the *Journal of Clinical Epidemiology* but had no involvement with the peer review process or decision for publication. A.C.T. is the co-Editor-in-Chief of the Journal of Clinical Epidemiology had no involvement with the peer review process or decision for publication. There are no competing interests for any author.

Acknowledgments

The authors would like to thank Tamara Rader, MLIS (Freelance Health Science Librarian), for peer reviewing the MEDLINE search strategy. In addition, the authors also thank Kristina Ferreira for helping retrieve several full-text PDFs.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2026.112389.

Data availability

All data can be found in the appendices.

References

[1] Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318:527−30.

[2] Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Developing guidelines. BMJ 1999;318:593−6.

[3] Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. Hoboken, PA: John Wiley & Sons; 2021.

[4] Higgins JP, Chandler J, Cumpston M, Li T, Page M, Welch V. Cochrane handbook for systematic reviews of interventions. London, UK: Cochrane; 2022. , 6.3. .

[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;312:171−9.

[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: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;147:224−33.

[8] Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. BMJ 2020;369:m1847.

[9] Runjic R, Kadic AJ, Runjic E, Gudelj K, Milić J, Pacheco RL, et al. Most systematic reviews that used the term ''update'' in title/abstract were not an updated version. J Clin Epidemiol 2023;154:1−7.

[10] Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? Lancet 2006;367:881−3.

[11] 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;2008:MR000023.

[12] 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:i3507.

[13] 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.

[14] 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:e1001603.

[15] Simmonds M, Elliott JH, Synnot A, Turner T. Living systematic reviews. Methods Mol Biol 2022;2345:121−34.

[16] 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;336:924−6.

[17] Murad MH, Wang Z, Chu H, Lin L, El Mikati IK, Khabsa J, et al. Proposed triggers for retiring a living systematic review. BMJ Evid Based Med 2023;28:348−52.

[18] Akl EA, El Khoury R, Khamis AM, El Mikati IK, Pardo-Hernandez H, Farran S, et al. The life and death of living systematic reviews: a methodological survey. J Clin Epidemiol 2023;156:11−21.

[19] Kamso MM, Pardo JP, Whittle SL, Buchbinder R, Wells G, Glennon V, et al. Crowdsourcing and automation facilitated the identification and classification of randomized controlled trials in a living review. J Clin Epidemiol 2023;164:1−8.

[20] White IR. Network meta-analysis. Stata J 2015;15:951−85.

[21] Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med 2002;21:2313−24.

[22] Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. Stat Med 2011;30:903−21.

[23] Imberger G. Multiplicity and sparse data in systematic reviews of anaesthesiological interventions cause increased risk of random errors and lack of reliability of conclusions. Copenhagen, Denmark: University of Copenhagen; 2014.

[24] 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;9:1−12.

[25] 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.

[26] Thorlund K, Imberger G, Wetterslev J, Brok J, Gluud C. Comments on 'Sequential meta-analysis: an efficient decision-making tool' by I van der Tweel and C Bollen. Clin Trials 2010;7:752−4.

[27] Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JPA, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol 2009;38:276−86.

[28] AlBalawi Z, McAlister FA, Thorlund K, Wong M, Wetterslev J. Random error in cardiovascular meta-analyses: how common are false positive and false negative results? Int J Cardiol 2013;168:1102−7.

[29] Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. BMJ open 2016;6:e011890.

[30] 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. London, England: Cochrane Methods Support Unit; 2019.

[31] Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 2017;17:39.

[32] 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;9:86.

[33] Brok J, Huusom LD, Thorlund K. Conclusive meta-analyses on antenatal magnesium may be inconclusive! are we underestimating the risk of random error? Acta Obstetricia Gynecol Scand 2012;91:1247−51.

[34] Lan KKG, Hu M, Cappelleri JC. Applying the law of iterated logarithm to cumulative meta-analysis of a continuous endpoint. Stat Sinica 2003;13:1135−45.

[35] Spence GT, Steinsaltz D, Fanshawe TR. A Bayesian approach to sequential meta-analysis. Stat Med 2016;35:5356−75.

[36] Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020;18:2119−26.

[37] Konstantinidis M, Veroniki AA, Stratton C, Tsokani S, Elliott JH, Simmonds M, et al. Conducting pairwise and network meta-analyses in Updated-and Living Systematic Reviews: a scoping review protocol. Philadelphia, PA: Wolters Kluwer; 2024.

[38] Konstantinidis M, Stratton C, Tsokani S, Elliott J, Simmonds M, McGowan J, et al. Conducting pairwise and network meta-analyses in updated and living systematic reviews: a scoping review protocol. JBI Evid Synth 2025;10:11124.

[39] 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;169:467−73.

[40] Agapova M, Devine EB, Nguyen H, Wolf FM, Inoue LYT. Using indirect comparisons to compare interventions within a Cochrane review: a tool for comparative effectiveness research. J Comp Effectiveness Res 2014;3:345−57.

[41] Bender R, Bunce C, Clarke M, Gates S, Lange S, Pace NL, et al. Attention should be given to multiplicity issues in systematic reviews. J Clin Epidemiol 2008;61:857−65.
[42] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. New York, NY: Elsevier; 2009.
[43] Borm GF, Donders ART. Updating meta-analyses leads to larger type I errors than publication bias. J Clin Epidemiol 2009;62:825−30.
[44] Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019;10:ED000142.
[45] Dogo SH, Clark A, Kulinskaya E. Sequential change detection and monitoring of temporal trends in random-effects meta-analysis. Res Synth Methods 2017;8:220−35.
[46] Dormuth CR, Filion KB, Platt RW. Likelihood ratio meta-analysis: new motivation and approach for an old method. Contemp Clin trials 2016;47:259−65.
[47] Fanshawe TR, Shaw LF, Spence GT. A large-scale assessment of temporal trends in meta-analyses using systematic review reports from the cochrane library. Res Synth Methods 2017;8:404−15.
[48] Garattini S, Jakobsen JC, Wetterslev J, Bertele V, Banzi R, Rath A, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. Eur J Intern Med 2016;32:13−21.
[49] Garritty C. Updating systematic reviews: the policies and practices of health care organizations involved in evidence synthesis. New York, NY: Elsevier; 2009.
[50] Higgins JPT. Prevalence and problems of updated reviews: a survey and discussion. In: Second symposium on systematic reviews: beyond the basics. Toronto, Canada: TSpace; 1999.
[51] Hu M, Cappelleri JC, Lan KKG. Applying the law of iterated logarithm to control type I error in cumulative meta-analysis of binary outcomes. Clin Trials 2007;4:329−40.
[52] Iannizzi C, Akl EA, Anslinger E, Weibel S, Kahale LA, Aminat AM, et al. Methods and guidance on conducting, reporting, publishing, and appraising living systematic reviews: a scoping review. Syst Rev 2023;12:238.
[53] Ioannidis JPA, Contopoulos-Ioannidis DG, Lau J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. J Clin Epidemiol 1999;52:281−91.
[54] Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol 2014;14:120.
[55] Khamis AM, Kahale LA, Pardo-Hernandez H, Schünemann HJ, Akl EA. Methods of conduct and reporting of living systematic reviews: a protocol for a living methodological survey. F1000Research 2019;8:221.
[56] Koricheva J, Jennions MD, Lau J. Chapter 15: temporal trends in effect sizes: causes, detection, and implications. Handbook of meta-analysis in ecology and evolution. Princeton, NJ: Princeton University Press; 2013:237−54.
[57] Kulinskaya E, Huggins R, Dogo SH. Sequential biases in accumulating evidence. Res Synth Methods 2015;7:294−305.
[58] Kulinskaya E, Mah EY. Cumulative meta-analysis: what works. Res Synth Methods 2022;13:48−67.
[59] Kulinskaya E, Wood JN. Trial sequential methods for meta-analysis. Res Synth Methods 2013;5:212−20.
[60] Lansky A, Wethington H. Living systematic reviews and other approaches for updating evidence. Am J Public Health 2020;110:1687−8.
[61] Mascha EJ. Alpha, Beta, Meta: guidelines for assessing power and type I error in meta-analyses. Anesth Anal 2015;121:1430−3.
[62] Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. A systematic review identified few methods and strategies describing when and how to update systematic reviews. J Clin Epidemiol 2007;60:1095−104.
[63] 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;27:1312−30.
[64] Pogue J, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Controlled Clin trials 1997;18:580−93.
[65] Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. Lancet 1998;351:47−52.
[66] Quan H, Ma Y, Zheng Y, Cho M, Lorenzato C, Hecquet C. Adaptive and repeated cumulative meta-analyses of safety data during a new drug development process. Pharm Stat 2015;14:161−71.
[67] Roloff V, Higgins JPT. Making best use of existing evidence when planning trials. Trials 2011;12:A51.
[68] Shuster JJ, Neu J. A Pocock approach to sequential meta-analysis of clinical trials. Res Synth Methods 2013;4:269−79.
[69] Sutton AJ, Higgins JPT. Recent developments in meta-analysis. Stat Med 2007;27:625−50.
[70] Ter Schure J, Grunwald P. Accumulation bias in meta-analysis: the need to consider time in error control. F1000Research 2019;8:962.
[71] Trikalinos TA, Ioannidis JPA. Publication bias in meta-analysis: prevention, assessment and adjustments - assessing the evolution of effect sizes over time. New York, NY: Elsevier; 2005.
[72] Tsertsvadze A, Maglione M, Chou R, Garritty C, Coleman C, Lux L, et al. Updating comparative effectiveness reviews: current efforts in AHRQ's Effective Health Care Program. J Clin Epidemiol 2011;64:1208−15.
[73] Tsertsvadze A, Moher D, Tricco A, Eccles M, Grimshaw J, Sampson M, et al. When and how to update systematic reviews: a methodological systematic review [abstract]. Thousand Oaks, CA: XIV Cochrane Colloquium; 2006:61.
[74] van der Tweel I, Bollen CW. Sequential meta-analysis: an efficient decision-making tool. Clin Trials 2010;7:136−46.
[75] Vergara-Merino L, Verdejo C, Carrasco C, Vargas-Peirano M. Living systematic review: new inputs and challenges. Medwave 2020;20:e8092.
[76] Whitehead A. A prospectively planned cumulative meta-analysis applied to a series of concurrent clinical trials. Stat Med 1997;16:2901−13.
[77] Whitehead A. Meta-Analysis Of Controlled Clinical Trials - Meta-Analysis of Controlled Clinical Trials. John Wiley & Sons, Ltd; 2002.
[78] Ye F. Statistical methodologies for trial design with network meta-analysis. London, UK: Iowa State University; 2023.
[79] Roloff V. Statistical methods for using meta-analysis to plan future research. Cambridge, UK: University of Cambridge; 2012.
[80] Trial Sequential Analysis (TSA). Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital − Rigshospitalet. Copenhagen, Denmark: The Copenhagen Trial Unit; 2021.
[81] Soerensen AL, Olsen MH, Lange T, Gluud C. RTSA: 'Trial Sequential Analysis' for Error Control and Inference in Sequential Meta-Analyses. London, UK: BioMed Central; 2023.
[82] Asante MA, Wagner V, Hafliðadóttir SH, Kumburegama BWMB, Pedersen EB, Ribeiro JP, et al. Where did I leave my systematic review protocol, and what should it contain regarding trial sequential analysis? BMC Med Res Methodol 2026;26:44.