



Revista de Epistemología y Metaciencia en Medicina y Biología

REFLECTION ARTICLE

Evidence Redundancy in Umbrella Reviews: Epistemological Implications of Overlap

Redundancia de evidencia en revisiones paraguas: implicaciones epistemológicas del solapamiento

Alberto Guevara-Tirado, MD, MSc

ORCID: 0000-0001-7536-7884.

Corresponding author: Alberto Guevara-Tirado

Email: albertoguevara1986@gmail.com

Short title: Evidence redundancy in umbrella reviews

ABSTRACT

Introduction: Umbrella reviews synthesize multiple systematic reviews, but repetition of the same primary studies can create an appearance of accumulated evidence that does not represent independent information. **Objective:** To analyze the epistemological implications of overlap among systematic reviews included in umbrella reviews and to propose an interpretive framework that distinguishes redundancy from independent convergence. **Methods:** A metascientific narrative review of methodological studies, guidance, and tools related to overlap, corrected covered area, review appraisal, and evidence synthesis was conducted. The analysis focused on information independence, precision, concordance, and inferential credibility. **Results:** Three levels of redundancy were distinguished - bibliographic, informational, and inferential - and the main strategies for identifying and managing overlap were systematized. A conceptual overlap-concordance matrix is also proposed, differentiating relatively independent convergence, concordant repetition, methodological divergence, and genuine heterogeneity. **Conclusions:** A larger number of reviews does not necessarily imply a larger amount of independent evidence. Overlap should be interpreted as a dependency property between syntheses rather than as an isolated marker of quality. Explicit assessment can prevent an apparent multiplication of certainty and improve interpretation of umbrella reviews.

Keywords: Umbrella review; Metascience; Epistemology; Overlap; Systematic review; Evidence synthesis.

RESUMEN

Introducción: Las revisiones paraguas sintetizan múltiples revisiones sistemáticas, pero la repetición de los mismos estudios primarios puede producir una apariencia de acumulación de evidencia que no representa información independiente. **Objetivo:** Analizar las implicaciones epistemológicas del solapamiento entre revisiones sistemáticas incluidas en revisiones paraguas y proponer un marco interpretativo para distinguir redundancia de convergencia independiente. **Métodos:** Se realizó una revisión narrativa metacientífica de trabajos metodológicos, guías de conducta y herramientas relacionadas con solapamiento, área cubierta corregida, evaluación de revisiones y síntesis de evidencia. El análisis se organizó alrededor de la independencia de la información, la precisión, la concordancia y la credibilidad inferencial. **Resultados:** Se distinguieron tres planos de redundancia - bibliográfica, informativa e inferencial - y se sistematizaron las principales estrategias para identificar y gestionar el solapamiento. Además, se propone una matriz de solapamiento-concordancia que diferencia cuatro escenarios: convergencia relativamente independiente, repetición concordante, divergencia metodológica y heterogeneidad genuina. **Conclusiones:** Un mayor número de revisiones no implica necesariamente mayor cantidad de evidencia independiente. El solapamiento debe interpretarse como una propiedad de dependencia entre síntesis, no como un marcador aislado de calidad. Su evaluación explícita puede evitar la multiplicación aparente de la certeza y mejorar la interpretación de las revisiones paraguas.

Palabras clave: Revisión paraguas; Metaciencia; Epistemología; Solapamiento; Revisión sistemática; Síntesis de evidencia.

INTRODUCTION

Systematic reviews emerged in response to the difficulty of interpreting an expanding primary literature in isolation. Subsequently, the proliferation of systematic reviews addressing the same question created a new need: to synthesize the syntheses. Umbrella reviews - also referred to as overviews of reviews, reviews of reviews, or meta-reviews - include systematic reviews as their unit of inclusion and allow a broad body of evidence to be presented in an integrated manner. This design, however, introduces problems that do not arise with the same intensity when the unit of analysis is the primary study. Early methodology for reviews of reviews already emphasized the need to explicitly identify included reviews and consider their methodological differences.⁽¹⁾ Later descriptive studies showed substantial variability in how these syntheses are conducted and reported.⁽²⁾

The expansion of systematic reviews has also created a problem of scientific redundancy. Repeated production of meta-analyses addressing very similar questions may consume resources, fragment knowledge, and give disproportionate visibility to datasets that have already been examined.⁽³⁾ In an umbrella review, this redundancy acquires an additional dimension: different systematic reviews may contain the same primary studies. Therefore, a collection of ten reviews does not necessarily represent ten independent sources of corroboration. To a large extent, it may represent ten reorganizations of the same set of studies.

Pieper and colleagues formalized the use of the corrected covered area (CCA) as a measure to quantify overlap of primary studies across systematic reviews.⁽⁴⁾ Other authors have noted that substantial overlap may compromise some functions of umbrella reviews and that the usefulness of a synthesis also depends on thematic correspondence, methodological quality, and the currency of the included reviews.⁽⁵⁾ Practical literature on umbrella reviews has established this problem as one of the central considerations in their conduct.⁽⁶⁾

From a statistical perspective, reusing the same information several times may produce an appearance of greater precision when duplicated data repeatedly enter a synthesis. From an epistemological perspective, the problem is broader: documentary repetition of an observation does not, by itself, increase its evidentiary independence. Lunny and colleagues described multiple approaches to managing overlapping data and emphasized that duplication may give undue influence to certain studies.⁽⁷⁾ The Cochrane Handbook for overviews likewise recognizes that double counting can give excessive weight to particular results.⁽⁸⁾

This article begins from a simple proposition: the number of documents expressing a conclusion and the amount of independent evidence supporting that conclusion are different quantities. Its objective is to analyze the epistemological implications of overlap in umbrella reviews, distinguish the main forms of redundancy, and propose a practical framework for jointly interpreting the degree of overlap and the concordance among reviews.

METHODS

A methodological and metascientific narrative review was conducted. The search aimed to identify foundational work on overviews of reviews, studies on overlap, tools for its quantification and representation, reporting guidance, and frameworks for assessing the quality and certainty of syntheses. PubMed, documents from the Cochrane Handbook, and the JBI Manual for Evidence Synthesis were consulted, together with related methodological references identified through targeted citation tracking. The search was updated through October 8, 2026. Because this was a methodological reflection article rather than a systematic review intended to estimate an effect, no meta-analysis was performed and no PRISMA selection process was applied.

The analysis was organized around four questions: 1) what is repeated when overlap occurs; 2) under what conditions that repetition may alter the interpretation of evidence; 3) which tools can describe the phenomenon; and 4) how documentary concordance should be distinguished from independent convergence. For reporting, the PRIOR statement, developed specifically for overviews of reviews of healthcare interventions, was considered.⁽⁹⁾ For representation of overlap, evidence matrices and tools such as GROOVE, which allow examination of overall and pairwise CCA, were considered.⁽¹⁰⁾ The ccaR package, developed to facilitate analysis of primary-study overlap across reviews, was also examined.⁽¹¹⁾

The interpretation did not assume that CCA constitutes a quality score. Empirical findings showing that overall CCA may conceal pairs of reviews with very high overlap and, conversely, that a context with high overall overlap may contain pairs with little overlap were considered.⁽¹²⁾ Methodological cautions regarding the use of pairwise CCA when reviews contain very different numbers of primary publications were also incorporated.⁽¹³⁾ Finally, the recent development of sample-size weighted corrected covered area (wCCA) was considered as an approach to distinguishing bibliographic overlap from informational overlap.⁽¹⁴⁾

RESULTS

The analysis distinguished three levels that are often grouped under the term overlap. Bibliographic redundancy describes the reappearance of the same primary study in different reviews; informational redundancy refers to reuse of a relevant proportion of participants, events, or statistical weight; and inferential redundancy arises when several apparently independent conclusions rest on the same empirical core. This distinction prevents the number of reviews from being interpreted as a substitute for the number of independent sources of evidence (Table 1).

Table 1. Levels of redundancy in umbrella reviews

Level	What is repeated	Main risk	Epistemological interpretation
Bibliographic	The same publication or primary study appears in two or more reviews.	Repeated counting of studies and an appearance of greater documentary volume.	Indicates dependency among reviews, but by itself does not show how much weight the repeated study contributes.
Informational	The same participants, events, or studies with a large contribution to the effect are reused.	Inflated precision or disproportionate influence of specific evidence.	Duplication affects the effective amount of information available.

Inferential	Several reviews reach conclusions supported by a shared empirical core.	Confusing repetition of an inference with independent corroboration.	Several syntheses may constitute a single line of evidence reformulated in different ways.
-------------	---	--	--

Available strategies serve different purposes, and no single measure resolves the entire problem. Evidence matrices identify which studies recur; CCA summarizes the density of overlap; pairwise analyses and GROOVE identify local concentrations that may be hidden within an overall estimate; and selection or reconstruction strategies at the level of primary studies can prevent double counting when quantitative synthesis is required. Methodological quality tools such as AMSTAR 2 address a different question and should not be used as substitutes for assessing overlap.⁽¹⁵⁾ Likewise, a reporting guideline such as PRISMA 2020 improves transparency, but compliance does not guarantee independence among the synthesized sources.⁽¹⁶⁾ (Table 2).

Table 2. Strategies to identify or manage overlap

Strategy	Main utility	Limitation	Recommended use
Evidence matrix	Shows which primary studies are contained in each review.	May become difficult to read with very large matrices.	Basis for understanding the structure of overlap and calculating subsequent measures.
Overall CCA	Summarizes overlap across the entire matrix as a percentage.	May conceal specific pairs with extremely high or minimal overlap.	Global description; should not be used as a quality score.
Pairwise CCA / GROOVE	Identifies which pairs of reviews share the highest proportion of studies.	Interpretation depends on the size and composition of each review.	Complement overall CCA and explore local patterns.
wCCA	Weights overlap by considering sample size.	Recent method; does not capture all forms of informational weight.	Useful complement when repeated studies differ substantially in size.
Selection of one review	Avoids duplication by choosing the most relevant, recent, or rigorous review.	May exclude unique studies present in discarded reviews.	Useful when several reviews essentially address the same question.
Reconstruction from primary studies	Allows synthesis without double counting at the review level.	Substantially increases workload and complexity.	Preferable when quantitative reanalysis is required and reviews overlap extensively.

The relationship between overlap and concordance provides a more informative interpretation than either element alone. High concordance with low overlap suggests relatively independent convergence; high concordance with high overlap may represent consistent repetition but not necessarily additional evidence; discordance with high overlap should prompt examination of analytical decisions, eligibility criteria, search dates, or synthesis methods; and discordance with low overlap may reflect genuine heterogeneity between empirical bases. This overlap-concordance matrix is proposed as a conceptual tool rather than a validated certainty scale (Table 3).

Table 3. Conceptual overlap-concordance matrix

Overlap	Concordance	Epistemological interpretation	Interpretive action
Low	High	Relatively independent convergence.	Agreement among syntheses has greater corroborative value, provided design and quality are adequate.
High	High	Concordant repetition of the same empirical core.	Do not count the number of reviews as independent evidence; describe the dependency.
High	Low	Methodological divergence over similar evidence.	Examine PICO, searches, extraction, risk of bias, statistical model, and updating.
Low	Low	Uncertainty or heterogeneity across distinct empirical bases.	Avoid a premature single conclusion and explore modifiers or population/exposure differences.

DISCUSSION

The main epistemological implication of overlap is that it requires separation of two concepts that are often confused: documentary repetition and independent corroboration. If five systematic reviews include essentially the same primary studies, agreement in their conclusions may be reassuring regarding the stability of the synthesis process, but it is not equivalent to five independent replications of the phenomenon. The epistemologically relevant unit is not always the publication; it is the source of information supporting the inference.

This point may be expressed as a principle of epistemic non-multiplication: repetition of the same empirical base across multiple synthesis products does not proportionally increase the amount of independent evidence. The principle does not imply that redundant reviews are useless. A new review may update searches, correct errors, incorporate more rigorous methods, examine different outcomes, or address another population. What is not valid is to treat their mere coexistence as though each provided a new observation of the world.

CCA is useful because it makes visible a dependency that might otherwise remain hidden. However, its conventional categories should be read as descriptors of an inclusion structure, not as automatic levels of confidence. A high CCA does not demonstrate that a conclusion is false, just as a low CCA does not guarantee that it is correct. In addition, empirical analysis has shown discrepancies between overall CCA and pairwise overlap.⁽¹²⁾ Therefore, when the number of reviews permits, the overall estimate should be accompanied by an evidence matrix or pairwise assessment.⁽¹⁰⁾

The emergence of measures such as wCCA is conceptually relevant because it recognizes a limitation of bibliographic counting: two repeated studies may contribute very different amounts of information.⁽¹⁴⁾ A trial with thousands of participants and a small pilot study both count as one publication in a traditional matrix, although their statistical contribution to a review may differ substantially. Weighting by sample size brings the analysis closer to the concept of informational dependency, although it still does not replace assessment of the actual weight of each study within each meta-analysis.

Another important consequence is that methodological quality assessment and dependency assessment should remain separate. AMSTAR 2 allows critical weaknesses in systematic reviews to be appraised, but a high-quality review may overlap almost completely with another high-quality review.⁽¹⁵⁾ Conversely, two reviews with little overlap may both be methodologically weak. Quality addresses how rigorously a synthesis was produced; overlap addresses how much of its empirical base is shared with other syntheses.

A similar issue arises with certainty of evidence. GRADE frameworks organize judgments across domains such as risk of bias, inconsistency, indirectness, imprecision, and publication bias.⁽¹⁷⁾ Overlap should not mechanically become an additional downgrading domain. Its effect depends on how reviews are used within the overview. If an author presents several reviews narratively and makes their dependency explicit, the problem is different from quantitatively combining estimates that repeatedly contain the same participants.

Metascience has shown that the quality of a synthesis depends on decisions made at every stage: question formulation, searching, selection, extraction, analysis, and interpretation.⁽¹⁸⁾ Umbrella reviews add a second layer of decisions because the researcher inherits choices made by the authors of each systematic review. This two-level architecture explains why two reviews may share almost all of their studies and still reach different results: they may use different outcomes, select different follow-up times, exclude subgroups, apply different statistical models, or assess risk of bias differently.

Current methodological recommendations recognize that umbrella reviews require explicit transparency regarding review selection, quality, and relationships among included reviews. The JBI Manual for Evidence Synthesis provides a specific framework for umbrella reviews and emphasizes the need for a systematic approach to included syntheses.⁽¹⁹⁾ The PRIOR statement complements this objective from a reporting perspective.⁽⁹⁾ From an editorial standpoint, requesting an evidence matrix or an explicit justification of how overlap was handled may be more informative than requiring only an overall score.

Redundancy should also be situated within the broader problem of research waste. Unnecessarily duplicated, poorly planned, or inadequately reported research may consume resources without proportionally increasing available knowledge.⁽²⁰⁾ Nevertheless, duplication is not synonymous with waste when there is a clear scientific rationale. Independent replication is essential to science; the problem arises when reuse of the same empirical material is presented as independence.

For authors of umbrella reviews, a reasonable practice is to first construct an evidence matrix; identify reviews that address essentially equivalent questions; quantify overall overlap and, when informative, pairwise overlap; separately assess methodological quality and currency; and decide before synthesis whether all reviews will be retained, one review will be selected for each question, or the evidence will be reconstructed at the primary-study level. The decision should be justified by the research question rather than by an automatic CCA threshold.

For readers and editors, the number of included reviews should no longer be interpreted as an automatic indicator of robustness. An umbrella review containing six partially independent reviews may have greater informational diversity than another containing thirty reviews that recycle the same set of trials. In this sense, reporting the number of reviews without reporting the number of unique primary studies and their pattern of overlap provides an incomplete picture of the true density of evidence.

LIMITATIONS

This article is a narrative review and conceptual analysis; therefore, it does not aim to estimate the global frequency of overlap or quantitatively compare the performance of different tools. The proposed overlap-concordance matrix is an interpretive framework and requires empirical evaluation before it can be used as a formal classification instrument.

In addition, independence among reviews cannot be fully summarized by matching publications. Dependency may arise from shared databases, partially overlapping cohorts, analytical decisions, or multiple reports of the same study. Therefore, even a well-constructed bibliographic matrix may underestimate deeper forms of informational dependency.

CONCLUSIONS

In umbrella reviews, more reviews do not necessarily mean more independent evidence. Overlap should be understood as a dependency structure among syntheses rather than as a simple bibliographic inconvenience or a quality score. Distinguishing bibliographic, informational, and inferential redundancy allows more precise interpretation of what is being repeated. Joint assessment of overlap and concordance may help distinguish independent convergence from apparent repetition. From an epistemological perspective, the credibility of a conclusion depends less on the number of documents supporting it than on the independence, quality, consistency, and relevance of the information those documents contain.

ETHICAL CONSIDERATIONS

This article analyzes publicly available scientific literature and methodological concepts and does not include human participants, animals, or identifiable individual-level data. Therefore, approval by an ethics committee and informed consent were not required.

CONFLICT OF INTEREST

The author declares no conflicts of interest related to the content of this article.

FUNDING

Self-funded work. No external funding was received for its preparation.

REFERENCES

1. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol*. 2011;11:15. doi:10.1186/1471-2288-11-15.
2. Hartling L, Chisholm A, Thomson D, Dryden DM. A descriptive analysis of overviews of reviews published between 2000 and 2011. *PLoS One*. 2012;7(11):e49667. doi:10.1371/journal.pone.0049667.
3. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94(3):485-514. doi:10.1111/1468-0009.12210.
4. Pieper D, Antoine SL, Mathes T, Neugebauer EAM, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. *J Clin Epidemiol*. 2014;67(4):368-375. doi:10.1016/j.jclinepi.2013.11.007.
5. Ballard M, Montgomery P. Risk of bias in overviews of reviews: a scoping review of methodological guidance and four-item checklist. *Res Synth Methods*. 2017;8(1):92-108. doi:10.1002/jrsm.1229.
6. Fusar-Poli P, Rada J. Ten simple rules for conducting umbrella reviews. *Evid Based Ment Health*. 2018;21(3):95-100. doi:10.1136/ebmental-2018-300014.
7. Lunny C, Pieper D, Thabet P, Kanji S. Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews. *BMC Med Res Methodol*. 2021;21(1):140. doi:10.1186/s12874-021-01269-y.
8. Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: Overviews of Reviews. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.5. Cochrane; 2024.
9. Gates M, Gates A, Pieper D, Fernandes RM, Tricco AC, Moher D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. *BMJ*. 2022;378:e070849. doi:10.1136/bmj-2022-070849.
10. Bracchiglione J, Meza N, Bangdiwala SI, Niño de Guzmán E, Urrutia G, Bonfill X, Madrid E. Graphical Representation of Overlap for OVERviews: GROOVE tool. *Res Synth Methods*. 2022;13(3):381-388. doi:10.1002/jrsm.1557.
11. Bougioukas KI, Diakonidis T, Mavromanolis AC, Haidich AB. ccaR: a package for assessing primary study overlap across systematic reviews in overviews. *Res Synth Methods*. 2023;14(3):443-454. doi:10.1002/jrsm.1610.
12. Bracchiglione J, Meza N, Pérez-Carrasco I, Vergara-Merino L, Madrid E, Urrutia G, Bonfill X. A methodological review finds mismatch between overall and pairwise overlap analysis in a sample of overviews. *J Clin Epidemiol*. 2023;159:31-39. doi:10.1016/j.jclinepi.2023.05.006.
13. Kirvalidze M, Abbadi A, Dahlberg L, Sacco LB, Calderón-Larrañaga A, Morin L. Estimating pairwise overlap in umbrella reviews: considerations for using the corrected covered area (CCA) index methodology. *Res Synth Methods*. 2023;14(5):764-767. doi:10.1002/jrsm.1658.
14. Ying X, Bougioukas KI, Pieper D, Mayo-Wilson E. Weighted corrected covered area (wCCA): a measure of informational overlap among reviews. *Res Synth Methods*. 2025;16(4):701-708. doi:10.1017/rsm.2025.19.
15. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. doi:10.1136/bmj.j4008.
16. 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. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71.
17. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394. doi:10.1016/j.jclinepi.2010.04.026.
18. Johnson BT, Hennessy EA. Systematic reviews and meta-analyses in the health sciences: best practice methods for research syntheses. *Soc Sci Med*. 2019;233:237-251. doi:10.1016/j.socscimed.2019.05.035.
19. Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Umbrella Reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JB Manual for Evidence Synthesis*. JBI; 2024. doi:10.46658/JBIMES-24-08.
20. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374(9683):86-89. doi:10.1016/S0140-6736(09)60329-9.