The Umbrella Collaboration®: An Innovative Tertiary Evidence Synthesis Methodology. Validation Study Protocol

Beltran Carrillo, Marta Rubinos-Cuadrado, Jazmin Parellada-Martin, Alejandra Palacios-López, Beltran Carrillo-Rubinos, Fernando Canillas-Del Rey, Juan Jose Baztán-Cortes, Javier Gómez-Pavon

Submitted to: JMIR Research Protocols
on: October 06, 2024

Table of Contents

Original Manuscript **5**
Supplementary Files **25**
..... **26**
..... **27**
Figures **28**
Figure 1 ..... **29**

The Umbrella Collaboration®: An Innovative Tertiary Evidence Synthesis Methodology. Validation Study Protocol

Beltran Carrillo[1*] MD; Marta Rubinos-Cuadrado[1*] RNC; Jazmin Parellada-Martin[1] BSCS; Alejandra Palacios-López[1] BME; Beltran Carrillo-Rubinos[1] BME; Fernando Canillas-Del Rey[2, 3] MD, PhD; Juan Jose Baztán-Cortes[3] MD, PhD; Javier Gómez-Pavon[2, 3] MD, PhD

[1]The Umbrella Collaboration Madrid ES
[2]Universidad Alfonso X el Sabio Villanueva de la Cañada, Madrid ES
[3]Hospital Universitario Cruz Roja Madrid ES
[*]these authors contributed equally

Corresponding Author:
Beltran Carrillo MD
The Umbrella Collaboration
C/ Ferraz, 49 - 1º izq
Madrid
ES

Abstract

Background: The synthesis of evidence in healthcare is essential for informed decision-making and policy development. This study aims to validate The Umbrella Collaboration® (TU®), an innovative, semi-automatic tertiary evidence synthesis methodology, by comparing it with Traditional Umbrella Reviews (TUR), which are currently the gold standard.

Objective: The primary objective of the present study is to assess whether a software-driven AI-assisted system of evidence synthesis, TU®, can match the effectiveness of traditional methods of tertiary synthesis, providing a potentially more timely, efficient, and comprehensive approach while remaining open to findings that could demonstrate superior performance. To support the primary objective of evaluating the effectiveness of TU® compared to traditional methodologies, this study also aims to assess the accessibility and comprehensibility of TU®'s outputs as a secondary objective

Methods: This protocol outlines a comparative study divided into two main parts. The first part involves a quantitative comparison of results obtained using TU® and TURs in geriatrics. We will evaluate the identification, size effect, direction, statistical significance, and certainty of outcomes, as well as the time and resources required for each methodology. Data for TURs will be sourced from Medline (via PubMed), while TU® will use AI-assisted informatics to replicate the research questions of the selected URTs. The second part of the study assesses the ease of use and comprehension of TU® through an online survey directed at health professionals, utilizing interactive features and detailed data access.

Results: Expected results include the assessment of concordance in identifying outcomes, the size effect, direction and significance of these outcomes, and the certainty of evidence. Additionally, we will measure the operational efficiency of each methodology by evaluating the time taken to complete projects. User perceptions of the ease of use and comprehension of TU® will be gathered through detailed surveys.

Conclusions: If TU® proves as effective as TURs but more time-efficient, accessible and easily updatable, it could significantly enhance the process of evidence synthesis, facilitating informed decision-making and improving healthcare. This study represents a step towards integrating innovative technologies into routine evidence synthesis practice, potentially transforming health research.

(JMIR Preprints 06/10/2024:67248)
DOI: https://doi.org/10.2196/preprints.67248

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

The Umbrella Collaboration®: An Innovative Tertiary Evidence Synthesis Methodology. Validation Study Protocol

Authors:

Beltrán Carrillo[a,b], Rubinos-Cuadrado, M[a], Parellada-Martin, J[a], Palacios-López, A[a], Carrillo-Rubinos, B[a], Canillas del Rey, F[b, c], Baztán-Cortés, JJ[b], Gómez-Pavón, J[b, c].

a: The Umbrella Collaboration®

b: Hospital Universitario Cruz Roja, (Madrid, Spain)

c: Universidad Alfonso X El Sabio, Villanueva de la Cañada (Madrid, Spain)

email autor: Beltran Carrillo. (bcm@theumbrellacollaboration.org)

ORCID:

Beltran Carrillo: 0000-0001-5748-9423

Rubinos-Cuadrado, M: 0000-0003-0108-2629

Parellada-Martin, J: 0000-0002-8528-4481

Palacios-López, A: 0000-0002-7263-9167

Carrillo-Rubinos, B: 0000-0002-0796-4414

Canillas del Rey, F:

Baztán-Cortes, JJ: 0009-00099792-0404

Gómez-Pavón, J: 0000-0002-4732-6265

ABSTRACT

BACKGROUND:

The synthesis of evidence in healthcare is essential for informed decision-making and policy development. This study aims to validate The Umbrella Collaboration® (TU®), an innovative, semi-automatic tertiary evidence synthesis methodology, by comparing it with Traditional Umbrella Reviews (TUR), which are currently the gold standard.

METHODS:

This protocol outlines a comparative study divided into two main parts. The first part involves a quantitative comparison of results obtained using TU® and TURs in geriatrics. We will evaluate the identification, size effect, direction, statistical significance, and certainty of outcomes, as well as the time and resources required for each methodology. Data for TURs will be sourced from Medline (via PubMed), while TU® will use AI-assisted informatics to replicate the research questions of the

selected URTs. The second part of the study assesses the ease of use and comprehension of TU® through an online survey directed at health professionals, utilizing interactive features and detailed data access.

RESULTS:
Expected results include the assessment of concordance in identifying outcomes, the size effect, direction and significance of these outcomes, and the certainty of evidence. Additionally, we will measure the operational efficiency of each methodology by evaluating the time taken to complete projects. User perceptions of the ease of use and comprehension of TU® will be gathered through detailed surveys.

DISCUSSION:
The implementation of new methodologies in evidence synthesis requires validation. This study will determine whether TU® can match the accuracy and comprehensiveness of TURs while offering benefits in terms of efficiency and user accessibility. The comparative study is designed to address the inherent challenges in validating a new methodology against established standards.

CONCLUSION:
If TU® proves as effective as TURs but more time-efficient, accessible and easily updatable, it could significantly enhance the process of evidence synthesis, facilitating informed decision-making and improving healthcare. This study represents a step towards integrating innovative technologies into routine evidence synthesis practice, potentially transforming health research.

FUNDING:
This project has received a research grant from the Fundación Alfonso X el Sabio.

CONFLICTS OF INTEREST:
The authors declare no conflicts of interest.

KEY WORDS: Tertiary Evidence Synthesis, The Umbrella Collaboration, Umbrella Reviews, Health Research Methodology, AI-assisted Synthesis, Evidence-based Decision Making

The Umbrella Collaboration®: An Innovative Tertiary Evidence Synthesis Methodology. Validation Study Protocol

INTRODUCTION

The synthesis of evidence in healthcare is a knowledge acquisition process designed to transform extensive volumes of data into manageable information to support decision-making based on the best possible evidence. It aims to integrate information from multiple sources on complex topics in a comprehensive, precise, transparent, and easily understandable manner. These principles ensure that synthesized evidence is accessible and useful to all stakeholders, including healthcare professionals, policymakers, and patients (1)(The Royal Society of Evidence Synthesis, 2018)[1].

Evidence synthesis plays a pivotal role in Knowledge Translation (KT) (2) (WHO, 2005)[2], serving as a bridge between research and healthcare. Consequently, evidence synthesis is essential for the development of health policies and informed decision-making (3) (Cottrell, 2014)[3].

The role of all stakeholders in health decision-making underscores the importance of their ability to access, understand, and evaluate health information adequately (4) (Sorensen, 2012)[4]. A persistent challenge in KT is the low level of statistical and health literacy among the general population and health professionals, which significantly complicates effective health management (5,6) (Martinez-Garcia, 2022)[5] (Baccolini, 2021)[6]. Therefore, it is crucial to democratize access to high-level health information and to promote active participation of all stakeholders in decisions affecting healthcare (7)(Baumann, 2022)[7].

[1] Accesible: https://royalsociety.org/-/media/policy/projects/evidence-synthesis/evidence-synthesis-statement-principles.pdf

[2] World Health Organization. (2005). Bridging the “know-do” gap meeting on knowledge translation in global health”; 10-12 Oct 205; Geneva, Switzerland. *Geneva: WHO*. (accesible: https://www.measureevaluation.org/resources/training/capacity-building-resources/high-impact-research-training-curricula/bridging-the-know-do-gap.pdf)

[3] Cottrell E, Whitlock E, Kato E, Uhl S, Belinson S, Chang C, Hoomans T, Meltzer D, Noorani H, Robinson K, Schoelles K, Motu'apuaka M, Anderson J, Paynter R, Guise JM. Defining the Benefits of Stakeholder Engagement in Systematic Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014 Mar. Report No.: 14-EHC006-EF. PMID: 24783309. https://www.ncbi.nlm.nih.gov/books/NBK196180/

[4] Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, Brand H; (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health. 2012 Jan 25;12:80. doi: 10.1186/1471-2458-12-80. PMID: 22276600; PMCID: PMC3292515. https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-12-80

[5] Martínez-García, J. J., Canizalez-Román, A., Velázquez-Román, J. A., Flores-Villaseñor, H. M., & León-Sicairos, N. M. (2022). Evaluación del conocimiento de métodos básicos de epidemiología e investigación en médicos residentes. *Revista Médica de la Universidad Autónoma de Sinaloa REVMEDUAS*, *11*(2), 115-123. https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=106198

[6] Baccolini V, Rosso A, Di Paolo C, Isonne C, Salerno C, Migliara G, Prencipe GP, Massimi A, Marzuillo C, De Vito C, Villari P, Romano F. What is the Prevalence of Low Health Literacy in European Union Member States? A Systematic Review and Meta-analysis. J Gen Intern Med. 2021 Mar;36(3):753-761. doi: 10.1007/s11606-020-06407-8. Epub 2021 Jan 5. PMID: 33403622; PMCID: PMC7947142. https://link.springer.com/article/10.1007/s11606-020-06407-8

[7] Baumann LA, Reinhold AK, Brütt AL. Public and patient involvement in health policy decision-making on the health system level - A scoping review. Health Policy. 2022 Oct;126(10):1023-1038. doi: 10.1016/j.healthpol.2022.07.007. Epub 2022 Jul 20. PMID: 35918211. https://www.sciencedirect.com/science/article/pii/S0168851022001919?via%3Dihub

The evolution of evidence synthesis methodologies has led to the development of tertiary synthesis, designed to condense knowledge from multiple systematic reviews with or without meta-analyses (SR/MA). This synthesis, often referred to as Umbrella Reviews now referred to as Traditional Umbrella Reviews (TUR) to distinguish them from the experimental methodology under study, The Umbrella Collaboration® (TU®), builds upon the concept of primary studies (individual studies with participant samples) and secondary studies (systematic reviews and analysis of those primary studies). Tertiary synthesis represents a third level, named with other terms such as overviews, meta-epidemiological studies, meta-meta-analyses, meta-synthesis, and meta-reviews also describing this approach (8,9) (Giuseppe Biondi-Zoccai, 2016)[8] (Choi, 2023)[9]. Tertiary synthesis has gained prominence in contexts where broad research questions are posed, rapid results are needed, and resources for extensive systematic reviews are limited. TURs follow a structured methodological process that involves several clearly defined stages (9–14) (Fusar-Poli, 2018)[10] (Aromataris, 2020)[11] (Cant, 2022)[12] (Belbasis, 2022)[13] (Pollock, 2023)[14] (Choi, 2023)[15] (Aromataris, 2024)[16]. Although organizations like Cochrane Collaboration (14) (Pollock, 2023)[17] and the Joanna Briggs Institute (JBI) (11) (Aromataris, 2024)[18] have developed and continually updated detailed methodologies for these reviews, there remains considerable divergence in how TUR authors implement these steps in practice. This methodological freedom leads to significant variations among different TURs in terms of rigor and approach.

The urgency for high-quality, timely information during crises like the SARS-CoV-2 pandemic has

[8] Biondi-Zoccai, G. (2016). Umbrella reviews. *Evidence synthesis with overviews of reviews and meta-epidemiologic studies. Cham, Switzerland: Springer International*.

[9] Choi GJ, Kang H. Introduction to Umbrella Reviews as a Useful Evidence-Based Practice. J Lipid Atheroscler. 2023 Jan;12(1):3-11. doi: 10.12997/jla.2023.12.1.3. Epub 2022 Oct 21. PMID: 36761061; PMCID: PMC9884555. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9884555/

[10] Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evid Based Ment Health. 2018 Aug;21(3):95-100. doi: 10.1136/ebmental-2018-300014. Epub 2018 Jul 13. PMID: 30006442; PMCID: PMC10270421. https://mentalhealth.bmj.com/content/21/3/95.long

[11] Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Chapter 10: Umbrella Reviews. Aromataris E, Munn Z, editors. *JBI Manual for Evidence Synthesis.* JBI; 2020. Available from https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-20-11

[12] Cant, R., Ryan, C., & Kelly, M. A. (2022). A nine-step pathway to conduct an umbrella review of literature. https://onlinelibrary.wiley.com/doi/full/10.1111/nae2.12039

[13] Belbasis L, Bellou V, Ioannidis JPA. Conducting umbrella reviews. BMJ Med. 2022 Nov 22;1(1):e000071. doi: 10.1136/bmjmed-2021-000071. PMID: 36936579; PMCID: PMC9951359.

[14] 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, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook.

[15] Choi GJ, Kang H. Introduction to Umbrella Reviews as a Useful Evidence-Based Practice. J Lipid Atheroscler. 2023 Jan;12(1):3-11. doi: 10.12997/jla.2023.12.1.3. Epub 2022 Oct 21. PMID: 36761061; PMCID: PMC9884555.

[16] Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JBI Manual for Evidence Synthesis*. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-01

[17] 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, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook.

[18] Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JBI Manual for Evidence Synthesis*. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-01

highlighted the critical need for faster evidence synthesis methods, even if it means accepting certain limitations in comprehensiveness, detail, and precision (15) (Tricco, 2022)[19]. This demand has driven the development of innovative approaches such as The Umbrella Collaboration® (TU®), which leverages AI-assisted software to facilitate tertiary evidence synthesis under human oversight. While AI tools such as PICO Portal, DistillerSR, Covidence, and Rayyan have already improved secondary evidence synthesis by streamlining data management and analysis (16) (Qureshi, 2023)[20], technologies like large language models (LLMs), including ChatGPT, are beginning to demonstrate potential for conducting systematic reviews autonomously, though human supervision remains essential to mitigate risks such as errors and hallucinations (17) (Teperikidis, 2024)[21]. Despite these developments in secondary synthesis, the application of AI and software engineering in tertiary synthesis is still in its early stages, with no dedicated software currently available. TU® is at the forefront of this field, pioneering the integration of AI and software engineering with human oversight to ensure accuracy and minimize technology-induced errors. As AI continues to evolve, it is likely that fully automated processes for both secondary and tertiary synthesis will emerge, potentially revolutionizing clinical research and practice. However, building confidence in these technologies will require ongoing development and rigorous validation.

This article was previously presented as a meeting abstract at the 64º Spanish Society of Geriatrics and Gerontology (Sociedad Española de Geriatria y Gerontologia) Congress on June 26 - 28, 2024.

The Umbrella Collaboration® (patent pending)

TU® is primarily a software-driven system engineered to streamline tertiary evidence synthesis, relying on programmed algorithms to automate the majority of its functions. The core of the system is built on a software infrastructure that processes and synthesizes data from SR/MA abstracts stored in Medline. While AI plays a crucial role, particularly through the use of LLMs and machine learning (ML), it is utilized selectively within the broader software framework to enhance specific tasks.

LLMs are used in generating related search terms, expanding upon human-generated queries to enhance the comprehensiveness of literature searches. Any LLM can be adapted to TU® software, up

[19] Tricco AC, Straus SE, Ghaffar A, Langlois EV. Rapid reviews for health policy and systems decision-making: more important than ever before. Syst Rev. 2022 Jul 30;11(1):153. doi: 10.1186/s13643-022-01887-7. PMID: 35906637; PMCID: PMC9338614.

[20] Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models "the answer" to bringing us closer to systematic review automation? Syst Rev. 2023 Apr 29;12(1):72. doi: 10.1186/s13643-023-02243-z. PMID: 37120563; PMCID: PMC10148473.

[21] Teperikidis L, Boulmpou A, Papadopoulos C, Biondi-Zoccai G. Using ChatGPT to perform a systematic review: a tutorial. Minerva Cardiol Angiol. 2024 Jul 26. doi: 10.23736/S2724-5683.24.06568-2. Epub ahead of print. PMID: 39056432.

to date we have used ChatGPT 4 (18) (OpenAI)[22]. This function is crucial in broadening the scope of the searches while ensuring that the results remain relevant and precise. The AI component is designed to support, not replace, human oversight, ensuring that the final selection of literature is both accurate and comprehensive (19) (Gwon, 2024)[23].

As the TU® database grows, ML will be integrated to further refine and optimize the software's performance. Training the system on an expanding dataset is expected to enhance its ability to select, categorize, and analyze relevant research, thereby improving both efficiency and accuracy over time. This approach allows TU® to evolve, continuously improving its utility in evidence synthesis through the iterative learning process (20) (Kothamali, 2024)[24].

Overall, TU® represents a hybrid model where traditional software engineering and targeted AI applications work in tandem. This balance ensures that while the software performs most functions automatically, AI enhances specific tasks, such as search term generation and future predictive analysis, under strict human supervision. This strategic integration of AI elements within a primarily software-driven system ensures the reliability and precision of the evidence synthesis process.

The outcomes of TU® are made available through an interactive public web application specifically designed to be accessible and comprehensible to all stakeholders, regardless of their statistical expertise. The combination of intuitive graphical formats and plain language ensures that the findings are easily interpreted by a wide audience.

Additionally, the platform supports continuous updates, automatically integrating new SR/MA data every 24 hours, thereby ensuring the most current and reliable evidence synthesis. This approach aligns with the concept of Living Systematic Reviews (LSRs) (21) (Elliot, 2017)[25], which advocate for frequent updates to maintain relevance in rapidly evolving fields. While Cochrane Collaboration recommends updating LSRs monthly (22) (Cochrane, 2019)[26], TU® surpasses this standard by ensuring updates are incorporated daily, providing near real-time evidence synthesis.

[22] OpenAI. 2024. *ChatGPT* (versión GPT-4). https://chat.openai.com/chat.

[23] Gwon YN, Kim JH, Chung HS, Jung EJ, Chun J, Lee S, Shim SR. The Use of Generative AI for Scientific Literature Searches for Systematic Reviews: ChatGPT and Microsoft Bing AI Performance Evaluation. JMIR Med Inform. 2024 May 14;12:e51187. doi: 10.2196/51187. PMID: 38771247; PMCID: PMC11107769.

[24] Kothamali, P. R., Karne, V. K., & Dandyala, S. S. M. (2024, July). Integrating AI and Machine Learning in Quality Assurance for Automation Engineering. In *International Journal for Research Publication and Seminar* (Vol. 15, No. 3, pp. 93-102).

[25] Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., ... & Pearson, L. (2017). Living systematic review: 1. Introduction—the why, what, when, and how. *Journal of clinical epidemiology*, *91*, 23-30.

[26] Cochrane: Guidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode. 2019. https://community.cochrane.org/sites/default/files/uploads/inline-files/Transform/201912_LSR_Revised_Guidance.pdf

The daily updates provided by TU® are further complemented by its remarkable efficiency. Pilot tests have demonstrated TU®'s capability to complete tertiary evidence synthesis projects within hours, a significant reduction in time compared to traditional methods. If validated by the upcoming research, this advancement could demonstrate TU®'s potential to streamline the synthesis process, delivering rapid yet reliable results while upholding the highest standards of accuracy. Should these findings be confirmed, TU® may emerge as an invaluable tool for accelerating the pace of evidence-based research.

A schematic diagram is provided to clearly illustrate TU®'s workflow, showing the process from data acquisition to the generation of synthesized results. Abstracts of SR/MA, retrieved from the Medline database via PubMed, form the foundation of the analysis. The data is processed through a range of techniques, including Natural Language Processing (NLP), sentiment analysis (SA), web scraping (WS) and ML. The expected results comprise synthesized evidence on intervention effectiveness and risk exposures, presented in a graphical and visual format. These results are conveyed in plain language, making them easily understandable by all stakeholders, regardless of their statistical literacy.

The project is continuously updated through automated and on-demand searches, with data from new studies seamlessly integrated into the existing body of evidence. Each inclusion restarts the synthesis process, creating a dynamic, cyclical workflow that ensures the results of the project remain up-to-date (Figure 1).

Figure 1. The Umbrella Collaboration® Workflow.

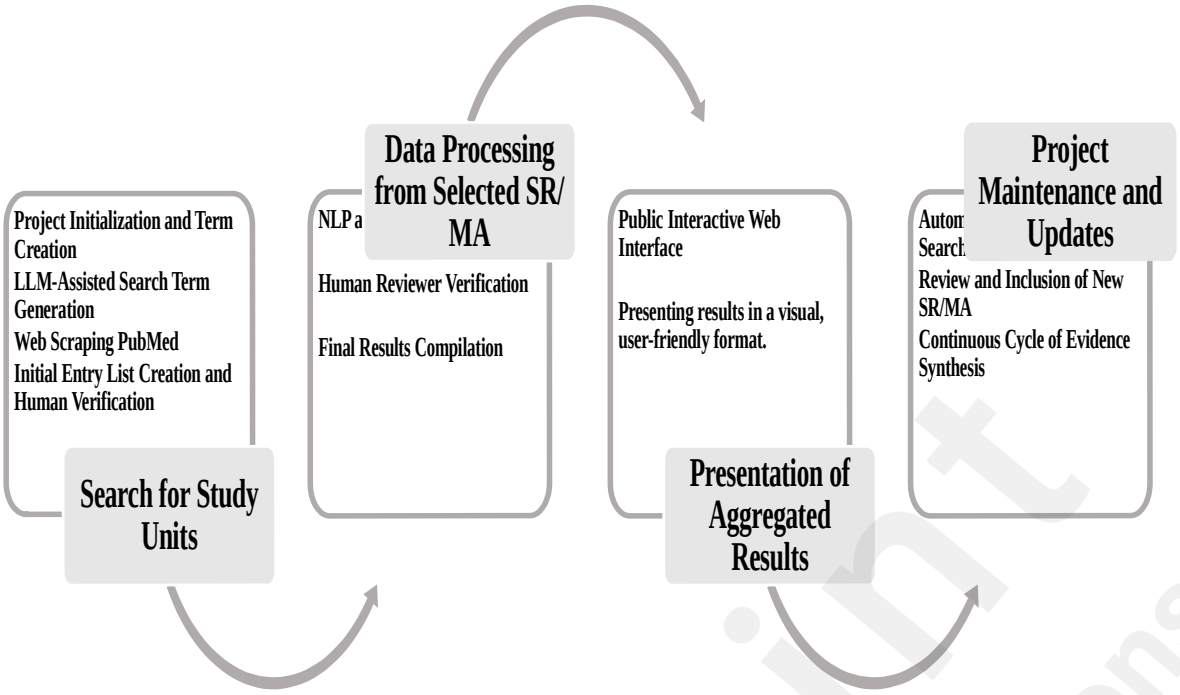

Figure 1 The Umbrella Collaboration® Workflow

The implementation of new methodologies in the scientific field requires a comparative validation process with established methods to ensure their reliability and effectiveness. TU®, being an innovative methodology still in its theoretical-conceptual stage, must be evaluated against established methodologies. Therefore, the aim of this study is to validate TU® by comparing its performance and outcomes with the gold standard, TURs, to establish its credibility and potential superiority.

Objectives

The primary objective of the present study is to assess whether a software-driven AI-assisted system of evidence synthesis, TU®, can match the effectiveness of traditional methods of tertiary synthesis, providing a potentially more timely, efficient, and comprehensive approach while remaining open to findings that could demonstrate superior performance. To support the primary objective of evaluating the effectiveness of TU® compared to traditional methodologies, this study also aims to assess the accessibility and comprehensibility of TU®'s outputs as a secondary objective.

METHODS AND MATERIALS

Study Design

Part 1: Quantitative Comparison of Methodologies

The first part of this study focuses on a quantitative comparison between the two tertiary synthesis methodologies. To facilitate this comparison, a targeted search in PubMed will identify relevant TURs in geriatrics, focusing on representative reviews rather than an exhaustive search. Our approach will involve a focused search in PubMed, using specific terms relevant to geriatrics, to find suitable TURs that serve as a benchmark for this comparative analysis. This targeted search is sufficient for our methodological comparison and does not require the comprehensive search strategy typical of systematic reviews, as our goal is not to cover the entire scope of available literature but to enable a parallel evaluation of synthesis methodologies. Therefore, while the search strategy may appear basic, it is intentionally designed to fulfill the specific needs of our project without aiming for exhaustive literature retrieval, which is beyond the scope of this project.

Study Variables

In the quantitative comparison, several critical variables will be analyzed. Key among these is the identification and evaluation of Outcomes of Interest (outcomes). This includes assessing the degree of concordance between the methodologies in identifying outcomes, using a concordance matrix to document and compare the outcomes identified by TU® and TURs. Additionally, we will analyze the total number of outcomes identified by each methodology, providing a descriptive and statistical comparison. It is essential to define the concept of an "outcome of interest" within the context of tertiary evidence synthesis. An outcome of interest refers to specific aspects identified and evaluated by systematic reviews with meta-analyses (SR/MA) that examine the same research question. These outcomes are critical for understanding the overall impact of various interventions on health conditions or the effects of exposures to risks.

A crucial aspect of this analysis involves the comparison of effect sizes for the identified outcomes. Given the diverse metrics used in TURs—such as Standardized Mean Difference (SMD), Mean Difference (MD), Relative Risk (RR), Odds Ratio (OR), and Hazard Ratio (HR)—and the unique metric employed by TU®, we will standardize all effect sizes to Cohen’s d. This standardization facilitates a direct comparison, ensuring consistency in the interpretation of results.

We will also examine the direction of the effects for each outcome, categorizing them as favorable, unfavorable, or unknown for interventions, and as increasing, decreasing, or unknown for exposures

to risks. The statistical significance of the outcomes will be compared by analyzing p-values and confidence intervals, assessing whether the results are statistically significant across both methodologies.

Furthermore, the certainty of the evidence associated with each outcome will be evaluated. For TURs, this will be done using the GRADE system (Grading of Recommendations Assessment, Development and Evaluation) (23) (Guyatt, 2008)[27], which categorizes evidence into very low, low, moderate, and high levels. TU® applies a SA-based scoring system on a scale from -1 to +1. Both scales will be normalized to a similar quantitative range (0-1) to facilitate comparison.

Finally, the execution time of each methodology will be assessed, with TU® providing exact time measurements and TURs relying on an estimated timeframe of 6 to 12 months based on existing literature.

Data Collection and Research Question Replication

Data collection will begin with a targeted search in PubMed to identify TURs in geriatrics, using the search terms "umbrella" AND "geriatric". The identified TURs will serve as benchmarks for our comparative analysis. The research questions from these selected umbrella reviews will be directly replicated in TU® without modification, ensuring a precise comparison of outcomes generated by each methodology. TU® will be configured to replicate these questions, employing automated searches and synthesis through NLP, WS, SA and ML, with human reviewer verifying and extracting data as necessary. This approach allows us to assess the comparative effectiveness and efficiency of TU® relative to traditional methods, particularly in identifying and analyzing outcomes critical to evaluating health interventions and exposure risks. Data from both TURs and TU® will be systematically collected and recorded in a database to facilitate precise comparisons of outcomes, effect sizes, and other critical variables, ensuring a thorough evaluation of both methodologies.

Part 2: Evaluation of Ease of Use and Comprehension

The second part of the study focuses on evaluating and comparing the ease of use and comprehension of the results generated by TU® with those from TURs. This evaluation will be conducted through an anonymous and voluntary online survey directed at health professionals, designed to assess their experience with both methodologies.

[27] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6. doi: 10.1136/bmj.39489.470347.AD. PMID: 18436948; PMCID: PMC2335261. https://www.bmj.com/content/336/7650/924.long

The survey, developed using Google Survey for ease of access and data analysis, comprises 16 items. The initial 6 items gather demographic information about the survey respondents, while the subsequent 10 questions directly compare the usefulness and clarity of the results produced by TU® and TURs, using a Likert scale ranging from 1 to 5. This scale will measure respondent's perceptions of the clarity, comprehensibility, and ease of use of the results provided by both methodologies. (Table 1).

To ensure thorough evaluation, informational sessions will be held in geriatric departments of university hospitals in Madrid (Spain). During these sessions, the concept of tertiary evidence synthesis and the functionality of TU® will be introduced. Health professionals, including those from geriatrics and other rotating specialties, will be given access to the TU® platform and guided through its use by an expert. This hands-on experience will be complemented by providing the participants with TUR result tables to facilitate a direct comparison.

Participants will be able to access the survey via a QR code (Figure 2), provided during the sessions, allowing them to complete it either immediately or at their convenience. The survey's responses will be analyzed descriptively, focusing on the overall user experience with TU® and its potential advantages in terms of ease of interpretation and presentation compared to traditional methods.

Table 1. Survey Questions on the Utility of The Umbrella Collaboration®

	Question	Answer Likert Scale
1 - 6	Respondent Affiliation Data	
7	Do you consider that the interactive interface of the "The Umbrella" (TU®) methodology facilitates the understanding of results compared to the static tables of Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
8	¿Do you believe that the visualisation of results using bubble plots in "The Umbrella" (TU®) helps to quickly identify the most relevant outcomes?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
9	Does the graphical representation of the outcomes (illustrated by figures, colours, and sizes) in "The Umbrella" (TU®) enhance your ability to assess the clinical relevance of the results?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
10	Do you find that access to detailed data by clicking on the figures of the outcomes in "The Umbrella" (TU®) interface enhances your evidence analysis experience?	1: Strongly disagree 2: 3:

		4: 5: Strongly agree
11	Do you think that "The Umbrella" (TU®) methodology allows for a quicker interpretation of data compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
12	Is the information provided by "The Umbrella" (TU®) useful in your field of work (clinical, research, or educational) for evidence-based decision making?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
13	Does the ease of use of "The Umbrella" (TU®) interface facilitate greater data exploration compared to traditional methods?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
14	Do you consider the historical evolution of evidence provided by "The Umbrella" (TU®) methodology to be useful?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
15	Does "The Umbrella" (TU®) methodology require less statistical knowledge to interpret the results compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
16	Overall, are you satisfied with "The Umbrella" (TU®) methodology as a tool for tertiary evidence synthesis?	1: Strongly disagree 2: 3: 4: 5: Strongly agree

.

Figure 2. QR Survey Accessing Code on the Ease of Use and Comprehension of TU®

Statistical Analysis: Data Analysis and Statistical Methods

The quantitative analysis will compare the results obtained by both methodologies. Contingency tables will be constructed to contrast the identification of outcomes, the size effect, the direction of the effect, and the statistical significance, ensuring statistical congruence between the methods. For the evaluation of certainty levels, TU® scores will be normalized to a scale comparable to the GRADE levels, which will also be transformed into a numerical scale between 0-1.

The chi-squared test will be used to determine significant differences between the methodologies. Additionally, Pearson and Spearman correlation analyses will be conducted to quantify the relationship between TU® certainty levels and GRADE certainty levels of the TURs. Pearson correlation is useful for quantifying the strength and direction of a linear relationship between two continuous variables. Moreover, the analysis will be complemented with Spearman correlation due to potential violations of normality assumptions in the data.

Statistical analysis will be performed using IBM SPSS Statistics 26, with an alpha level of 0.05 to determine statistical significance.

RESULTS

In the results section of this study, we will focus on evaluating the effectiveness of the experimental methodology in accurately identifying and analyzing relevant outcomes. Among the results we will include detailed assessments of effect sizes, the direction of effects, statistical significance, and the certainty of evidence for each outcome. We will compare these findings with those derived from TUR to determine if the experimental approach yields results that are at least equivalent in quality and comprehensiveness.

Finally, the efficiency of the experimental methodology will be evaluated by recording the time taken to complete the synthesis process. While the time required for TUR is estimated to range between 6 to 12 months based on existing literature, we will document the actual time taken by the experimental approach. This will provide a practical measure of the potential time savings offered by the software driven AI-assisted method, highlighting its feasibility and effectiveness in a real-world context.

To evaluate the ease of use and comprehension of TU® within its environment, a detailed survey will be specifically designed. Survey respondents will have access to the TU® platform, where they can interact with various interactive screens displaying the results of the synthesis process across different projects completed to date.

ETHICAL CONSIDERATIONS

This study does not involve human participants, personal data, or animals. All data will be sourced from published materials, and the analyses will be conducted in accordance with established ethical standards for secondary data analysis. Additionally, the survey component will be conducted anonymously and on a voluntary basis. Given these considerations, we do not deem it necessary to seek approval from an ethics committee, as the study adheres to standard ethical practices for research of this nature.

DISCUSSION

The primary objective of this study is to evaluate whether The Umbrella Collaboration® (TU®), a software-driven system designed to facilitate tertiary evidence synthesis with AI-assisted methodologies, can match the effectiveness of traditional umbrella reviews (TURs). TU® integrates within its software development advanced technologies such as NLP, SA, WS and ML to enhance the efficiency of evidence synthesis.

The semi-automated processes implemented by TU® could signify a significant advancement in making evidence synthesis more accessible and timely, with the potential for continuous updates as new data becomes available. The study's findings could pave the way for broader adoption of AI-driven methodologies in evidence synthesis, potentially reducing the time and resources needed for comprehensive reviews. TU®'s integration of software engineering project and AI with traditional methods could streamline the review process, enabling faster aggregation and interpretation of data across various research domains.

By offering a platform that supports real-time updates and provides accessible synthesis outputs, TU® has the potential to enhance the utility of tertiary synthesis for a wide range of stakeholders, including those with limited statistical expertise. However, a key limitation of TU® is its reliance on a single database (Medline via PubMed), which might not capture all relevant studies. This study will

examine whether this limitation can be mitigated by the system's other capabilities. Ultimately, the timely and comprehensible evidence synthesis provided by TU® could facilitate more informed decision-making in clinical settings, particularly in rapidly evolving areas of medical research.

LIMITATIONS

The decision to utilize only one database, Medline via PubMed, in TU® is both a recognized limitation and a deliberate choice shaped by resource constraints and technical considerations. While systematic reviews typically require searching multiple databases to capture all relevant literature (24) (Gusenbauer, 2020)[28], our approach focuses on testing whether TU® can achieve outcomes comparable to TURs despite these limitations. PubMed's strong coverage and the use of LLMs to expand search terms have proven effective, with a pilot study showing that 86% of references from TURs were identified using only PubMed, and of the 511 references included, only 11 were not indexed in PubMed. This highlights the potential of our approach to match traditional methods, although the limitation of using a single database is acknowledged.

Furthermore, while searching multiple databases is often recommended to avoid language and indexing biases, especially those related to non-English literature (25,26) (Jia, 2020)[29] (Mao, 2020)[30], TU® mitigates some of these biases by focusing on abstracts in English, as all PubMed abstracts are provided in this language regardless of the original publication's language. However, the absence of Chinese databases in our approach is a notable limitation, given that only a small proportion of Chinese journals are indexed in Medline.

CONCLUSION

This study aims to validate TU® as a tool for tertiary evidence synthesis in health. If this methodology proves to be as effective as TURs, but more efficient in terms of project execution time and more accessible in terms of ease of use and comprehension, it could significantly enhance the way evidence synthesis is conducted, facilitating informed decision-making, and improving health outcomes. The results of this study may represent a step towards the integration of innovative technologies into the routine practice of evidence synthesis, with the potential to transform the field

[28] Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res Synth Methods. 2020 Mar;11(2):181-217. doi: 10.1002/jrsm.1378. Epub 2020 Jan 28. PMID: 31614060; PMCID: PMC7079055.

[29] Jia Y, Huang D, Wen J, Wang Y, Rosman L, Chen Q, Robinson KA, Gagnier JJ, Ehrhardt S, Celentano DD. Assessment of Language and Indexing Biases Among Chinese-Sponsored Randomized Clinical Trials. JAMA Netw Open. 2020 May 1;3(5):e205894. doi: 10.1001/jamanetworkopen.2020.5894. PMID: 32463469; PMCID: PMC7256669.

[30] Mao C, Li M. Language Bias Among Chinese-Sponsored Randomized Clinical Trials in Systematic Reviews and Meta-analyses—Can Anything Be Done? *JAMA Netw Open.* 2020;3(5):e206370. doi:10.1001/jamanetworkopen.2020.6370

of health research.

FUNDING

This project has received a research grant from the *Fundación Alfonso X el Sabio*.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES:

1. Evidence synthesis for policy A STATEMENT OF PRINCIPLES [Internet]. The Royal Society of Evidence Synthesis; 2018. Available from: https://royalsociety.org/-/media/policy/projects/evidence-synthesis/evidence-synthesis-statement-principles.pdf
2. World Health Organization, 2005. bridging-the-know-do-gap.pdf [Internet]. [cited 2024 Jun 2]. Available from: https://www.measureevaluation.org/resources/training/capacity-building-resources/high-impact-research-training-curricula/bridging-the-know-do-gap.pdf
3. Cottrell E, Whitlock E, Kato E, Uhl S, Belinson S, Chang C, et al. Defining the Benefits of Stakeholder Engagement in Systematic Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014 [cited 2024 May 26]. (AHRQ Methods for Effective Health Care). Available from: http://www.ncbi.nlm.nih.gov/books/NBK196180/
4. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health [Internet]. 2012 Jan 25 [cited 2024 May 26];12(1):80. Available from: https://doi.org/10.1186/1471-2458-12-80
5. Martínez-García JJ, Canizalez-Román A, Velázquez-Román JA, Flores-Villaseñor HM, León-Sicairos NM. Evaluación del conocimiento de métodos básicos de epidemiología e investigación en médicos residentes. Rev Médica Univ Autónoma Sinaloa REVMEDUAS [Internet]. 2022 Jul 13 [cited 2024 May 26];11(2):115–23. Available from: https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=106198
6. Baccolini V, Rosso A, Di Paolo C, Isonne C, Salerno C, Migliara G, et al. What is the Prevalence of Low Health Literacy in European Union Member States? A Systematic Review and Meta-analysis. J Gen Intern Med [Internet]. 2021 Mar 1 [cited 2024 May 26];36(3):753–61. Available from: https://doi.org/10.1007/s11606-020-06407-8
7. Baumann LA, Reinhold AK, Brütt AL. Public and patient involvement in health policy decision-making on the health system level – A scoping review. Health Policy [Internet]. 2022 Oct 1 [cited 2024 May 26];126(10):1023–38. Available from: https://www.sciencedirect.com/science/article/pii/S0168851022001919
8. Biondi-Zoccai G, editor. Umbrella Reviews: Evidence Synthesis with Overviews of Reviews and Meta-Epidemiologic Studies [Internet]. Cham: Springer International Publishing; 2016 [cited 2024 May 26]. Available from: http://link.springer.com/10.1007/978-3-319-25655-9
9. Choi GJ, Kang H. Introduction to Umbrella Reviews as a Useful Evidence-Based Practice. J Lipid Atheroscler [Internet]. 2023 Jan [cited 2024 May 26];12(1):3–11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9884555/
10. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. BMJ Ment Health [Internet]. 2018 Aug 1 [cited 2024 May 26];21(3):95–100. Available from: https://mentalhealth.bmj.com/content/21/3/95
11. Aromataris E, Lockwood C, Porritt K, Pilla B. JBI Manual for Evidence Synthesis - JBI Global Wiki [Internet]. [cited 2024 Jun 5]. Available from: https://jbi-global-wiki.refined.site/space/MANUAL
12. Cant R, Ryan C, Kelly MA. A nine-step pathway to conduct an umbrella review of literature. Nurse Author Ed [Internet]. 2022 [cited 2024 May 26];32(2):31–4. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/nae2.12039
13. Belbasis L, Brooker RD, Zavalis E, Pezzullo AM, Axfors C, Ioannidis JP. Mapping and systematic appraisal of umbrella reviews in epidemiological research: a protocol for a meta-epidemiological study. Syst Rev [Internet]. 2023 Jul 14 [cited 2024 May 26];12(1):123. Available from: https://doi.org/10.1186/s13643-023-02265-7
14. Pollock M. Chapter V: Overviews of Reviews [Internet]. 2023 [cited 2024 May 26]. Available

from: https://training.cochrane.org/handbook/current/chapter-v
15. Tricco AC, Straus SE, Ghaffar A, Langlois EV. Rapid reviews for health policy and systems decision-making: more important than ever before. Syst Rev [Internet]. 2022 Jul 30 [cited 2024 Jun 5];11(1):153. Available from: https://doi.org/10.1186/s13643-022-01887-7
16. Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? Syst Rev [Internet]. 2023 Apr 29 [cited 2024 Aug 17];12(1):72. Available from: https://doi.org/10.1186/s13643-023-02243-z
17. Teperikidis L, Boulmpou A, Papadopoulos C, Biondi-Zoccai G. Using ChatGPT to perform a systematic review: a tutorial. Minerva Cardiol Angiol. 2024 Jul 26;
18. ChatGPT [Internet]. [cited 2024 Sep 7]. Available from: https://openai.com/chatgpt/
19. Gwon YN, Kim JH, Chung HS, Jung EJ, Chun J, Lee S, et al. The Use of Generative AI for Scientific Literature Searches for Systematic Reviews: ChatGPT and Microsoft Bing AI Performance Evaluation. JMIR Med Inform [Internet]. 2024 May 14 [cited 2024 Aug 15];12:51187. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11107769/
20. Kothamali PR, Karne VK, Dandyala SSM. Integrating AI and Machine Learning in Quality Assurance for Automation Engineering. Int J Res Publ Semin [Internet]. 2024 Jul 25 [cited 2024 Aug 15];15(3):93–102. Available from: https://jrps.shodhsagar.com/index.php/j/article/view/1445
21. 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 Nov;91:23–30.
22. 201912_LSR_Revised_Guidance.pdf [Internet]. [cited 2024 Aug 15]. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/Transform/201912_LSR_Revised_Guidance.pdf
23. 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 [Internet]. 2008 Apr 24 [cited 2024 Jun 6];336(7650):924–6. Available from: https://www.bmj.com/content/336/7650/924
24. Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res Synth Methods [Internet]. 2020 Mar [cited 2024 Aug 16];11(2):181–217. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7079055/
25. Jia Y, Huang D, Wen J, Wang Y, Rosman L, Chen Q, et al. Assessment of Language and Indexing Biases Among Chinese-Sponsored Randomized Clinical Trials. JAMA Netw Open [Internet]. 2020 May 28 [cited 2024 Aug 16];3(5):e205894. Available from: https://doi.org/10.1001/jamanetworkopen.2020.5894
26. Mao C, Li M. Language Bias Among Chinese-Sponsored Randomized Clinical Trials in Systematic Reviews and Meta-analyses—Can Anything Be Done? JAMA Netw Open [Internet]. 2020 May 28 [cited 2024 Aug 16];3(5):e206370. Available from: https://doi.org/10.1001/jamanetworkopen.2020.6370

Supplementary Files

Untitled.

Table 1. Survey Questions on the Utility of The Umbrella Collaboration*

	Question	Answer Likert Scale
1 - 6	Respondent Affiliation Data	
7	Do you consider that the interactive interface of the "The Umbrella" (TU') methodology facilitates the understanding of results compared to the static tables of Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
8	¿Do you believe that the visualisation of results using bubble plots in "The Umbrella" (TU') helps to quickly identify the most relevant outcomes?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
9	Does the graphical representation of the outcomes (illustrated by figures, colours, and sizes) in "The Umbrella" (TU') enhance your ability to assess the clinical relevance of the results?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
10	Do you find that access to detailed data by clicking on the figures of the outcomes in "The Umbrella" (TU') interface enhances your evidence analysis experience?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
11	Do you think that "The Umbrella" (TU') methodology allows for a quicker interpretation of data compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
12	Is the information provided by "The Umbrella" (TU') useful in your field of work (clinical, research, or educational) for evidence-based decision making?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
13	Does the ease of use of "The Umbrella" (TU') interface facilitate greater data exploration compared to traditional methods?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
14	Do you consider the historical evolution of evidence provided by "The Umbrella" (TU') methodology to be useful?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
15	Does "The Umbrella" (TU') methodology require less statistical knowledge to interpret the results compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3:

Untitled.

Table 1. Survey Questions on the Utility of The Umbrella Collaboration*

	Question	Answer Likert Scale
1 - 6	Respondent Affiliation Data	
7	Do you consider that the interactive interface of the "The Umbrella" (TU') methodology facilitates the understanding of results compared to the static tables of Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
8	¿Do you believe that the visualisation of results using bubble plots in "The Umbrella" (TU') helps to quickly identify the most relevant outcomes?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
9	Does the graphical representation of the outcomes (illustrated by figures, colours, and sizes) in "The Umbrella" (TU') enhance your ability to assess the clinical relevance of the results?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
10	Do you find that access to detailed data by clicking on the figures of the outcomes in "The Umbrella" (TU') interface enhances your evidence analysis experience?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
11	Do you think that "The Umbrella" (TU') methodology allows for a quicker interpretation of data compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
12	Is the information provided by "The Umbrella" (TU') useful in your field of work (clinical, research, or educational) for evidence-based decision making?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
13	Does the ease of use of "The Umbrella" (TU') interface facilitate greater data exploration compared to traditional methods?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
14	Do you consider the historical evolution of evidence provided by "The Umbrella" (TU') methodology to be useful?	1: Strongly disagree 2: 3: 4: 5: Strongly agree
15	Does "The Umbrella" (TU') methodology require less statistical knowledge to interpret the results compared to Traditional Umbrella Reviews (URT)?	1: Strongly disagree 2: 3:

Figures

The Umbrella Collaboration Workflow.

Figure 1. The Umbrella Collaboration® Workflow.