

Development of the systematic review duplication (SRD) tool: a protocol for assessing duplication in intervention-based systematic reviews

Yue-ning Deng^{1,§}, Si-qi Wu^{1,§}, Xu-fei Luo^{2,3}, Bin-bin Tang⁴, Yu-cheng Feng⁴, Hao-yue Feng⁴, Hao-lin Guo⁴, Yi-ming Li⁴, Lin Shi¹, Zhi-rui Xu¹, Zi-feng Zhou¹, Lin Yu¹, Yao-long Chen^{5,6,7,8,✉}, Yu-ting Duan^{1,✉}

¹ The Affiliated Traditional Chinese Medicine Hospital, Guangzhou Medical University, Guangzhou, 510640, China

² Vincent V.C. Woo Chinese Medicine Clinical Research Institute, School of Chinese Medicine, Hong Kong Baptist University, Hong Kong, 999077, China

³ Chinese EQUATOR Centre, Hong Kong, 999077, China

⁴ Clinical School of Integrated Traditional Chinese and Western Medicine, Guangzhou Medical University, Guangzhou, 511436, China

⁵ School of Basic Medical Sciences, Lanzhou University, Lanzhou, 730000, China

⁶ World Health Organization Collaboration Center for Guideline Implementation and Knowledge Translation, Lanzhou University, Lanzhou, 730000, China

⁷ Institute of Health Data Science, Lanzhou University, Lanzhou, 730000, China

⁸ Key Laboratory of Evidence Based Medicine of Gansu Province, Lanzhou University, Lanzhou, 730000, China

Received: December 16, 2025; Revised: February 18, 2026; Accepted: March 20, 2026

ABSTRACT

Introduction The rapid expansion of systematic reviews (SRs) has led to increasing concerns about research duplication and redundancy. While several tools exist to assess the quality and risk of bias of SRs, no standardized instrument specifically evaluates the degree of duplication between SRs on similar topics. This protocol describes the development and validation of the systematic review duplication (SRD) tool, which is designed to systematically assess and quantify duplication between intervention-based SRs in healthcare. **Methods** The development process follows established guidelines for creating reporting and quality assessment tools and comprises three phases: (I) preparatory work, including team formation and tool conceptualization; (II) tool development, involving initial item generation through a systematic literature review and analysis of existing tools, item validation through pilot testing with 40 SRs across 17 disease categories, modified Delphi surveys, and consensus meetings; and (III) dissemination through academic channels and professional networks. The SRD tool will evaluate duplication across four domains: research topic, research methods, research results, and research quality. Expert consensus will be achieved through a two-round modified Delphi process with an agreement threshold of at least 70%. Inter-rater reliability will be assessed using the intraclass correlation coefficient and Kendall's W coefficient. Expected outcomes: The SRD tool is expected to provide a standardized, user-friendly instrument for distinguishing necessary replication from redundant duplication in SRs. It will be available as both web-based and Excel-based applications. **Discussion** The SRD tool is intended to fill a critical gap in evidence synthesis methodology by enabling the systematic assessment of duplication and supporting informed decision-making by researchers, journal editors, guideline developers, and clinicians.

KEYWORDS

systematic review, research duplication, research redundancy, methodological tool, evidence synthesis, protocol

CITE THIS ARTICLE

Deng YN, Wu SQ, Luo XF, et al. Development of the systematic review duplication (SRD) tool: a protocol for assessing duplication in intervention-based systematic reviews. Evidence-Based Chinese Medicine and Technology Assessment, 2026, 2: 9570031. <https://doi.org/10.26599/eCMTA.2026.9570031>

1. Introduction

Systematic review(s) (SR/SRs) represent a gold standard of evidence-based medicine,¹ employing systematic search strategies, pre-established selection criteria, critical appraisal, and evidence synthesis to address specific research questions and inform clinical decision-making.²

The number of SRs and meta-analyses has grown exponentially in recent years. A study published in 2016 revealed that MEDLINE included over 8000 SRs annually, representing a threefold increase compared with the previous decade.³ In 2000, PubMed included an estimated 1432 SRs per year, and by 2019, this number had increased to 29 073, representing a 20-fold increase.⁴

However, this growth has also led to frequent occurrences of repetitive research and research redundancy. Overlapping SRs on the same topics are highly prevalent and exhibit substantial variability in methodological quality

and results, often producing inconsistent conclusions that may mislead clinicians and contribute to research waste.⁵ An SR published in 2017 found that 12 SRs and two major guidelines on thrombolytic therapy for pulmonary embolism were published within less than 2 years, and the results of these evidence syntheses were inconsistent.⁶ Similarly, evaluations of thromboembolic SRs demonstrated substantial duplication, and the duplicated SRs showed no improvement in overall methodological quality compared with the original reviews.⁷

Several publications distinguish five related but distinct concepts relevant to this project. Research replication refers to different research teams conducting evidence synthesis using the same, or highly similar, research questions and methods to verify the reliability of original findings.⁸ Research overlap refers to multiple SRs or meta-analyses conducted on the same topic. One study revealed that two-thirds of the meta-analyses published in 2010 had overlapping studies addressing the identical topic, and these studies were not updates undertaken by the same author team.⁹ Research duplication refers to unnecessary, often unintentional or unacknowledged repetition without a clear purpose and rarely with an explanation of the potential additional contributions it might make in the context of existing reviews.¹⁰ Research redundancy occurs when SRs are conducted despite the existence of high-quality

✉ Corresponding authors.

E-mail addresses: chenyaolong@lzu.edu.cn (YL. Cheng); duanyt@gzhmu.edu.cn (YT. Duan)

§ These authors contributed equally to this work.

© 2026 Beijing University of Chinese Medicine. Production and hosting by Tsinghua University Press. This is an open access article under the CC-BY license (<http://creativecommons.org/licenses/by/4.0/>).

evidence, providing no new insights or improvements and creating excessive overlapping content in the literature.¹¹ Research waste encompasses the misallocation of resources (human, financial, and temporal) resulting from duplicative, redundant, and low-quality SRs, potentially leading to evidence confusion and misguided clinical decisions.^{12,13}

Replication in research can contribute to the advancement and development of the field. This is particularly important when the initial research may be misleading because of issues such as the scope of the research question, the methods employed, or calculation errors.¹⁴ Of course, not all overlap is equivalent to redundancy and waste. When overlap serves the purpose of replication, a certain degree of it can be acceptable.¹⁵ However, when replication exceeds a reasonable degree or fails to provide substantive methodological improvements, it may evolve into duplication or even redundancy, thereby constituting research waste.¹³ The extensive duplication of similar or highly similar SRs may pose risks to science, ethics, the economy, and society, resulting in the continued waste of research resources, time, and energy.¹¹

Existing tools such as PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020,¹⁶ AMSTAR (A Measurement Tool to Assess Systematic Reviews) 2,¹⁷ ROBIS (Risk of Bias in Systematic Reviews),¹⁸ and ReMarQ (reporting methodological quality)¹⁹ focus on reporting quality, methodological quality, or risk of bias in SRs. In 2021, Lunny et al. proposed a classification system for SRs of health interventions. Overlap was defined as repetition of the PICO (population, intervention, comparison, outcome) eligibility criteria without being an update or repeated reporting. Overlapping reviews were classified into four categories: identical, nearly identical, partial-

ly included, or fully included.¹³ However, none of these approaches is designed to evaluate whether two or more SRs substantially duplicate one another. A dedicated instrument is therefore needed.

Therefore, we will develop a standardized tool for assessing the degree of duplication between SRs addressing similar topics by comparing structural and methodological elements.

2. Methods

2.1. Ethical approval

This protocol was registered on the Open Science Framework (OSF) in October 2025 (<https://doi.org/10.17605/OSF.IO/P5NGV>). This study is a theoretical and methodological study that does not involve human subjects, patient data, clinical samples, or any identifiable personal information. Therefore, ethical approval and informed consent were not required.

2.2. Overall design and framework

The development of the systematic review duplication (SRD) tool follows established frameworks for creating reporting guidelines and quality assessment tools.^{20,21} The process comprises three distinct phases, as shown in Figure 1.

- Phase 1: preparatory work (team formation and tool conceptualization);
- Phase 2: tool development (initial item generation, pilot testing and item

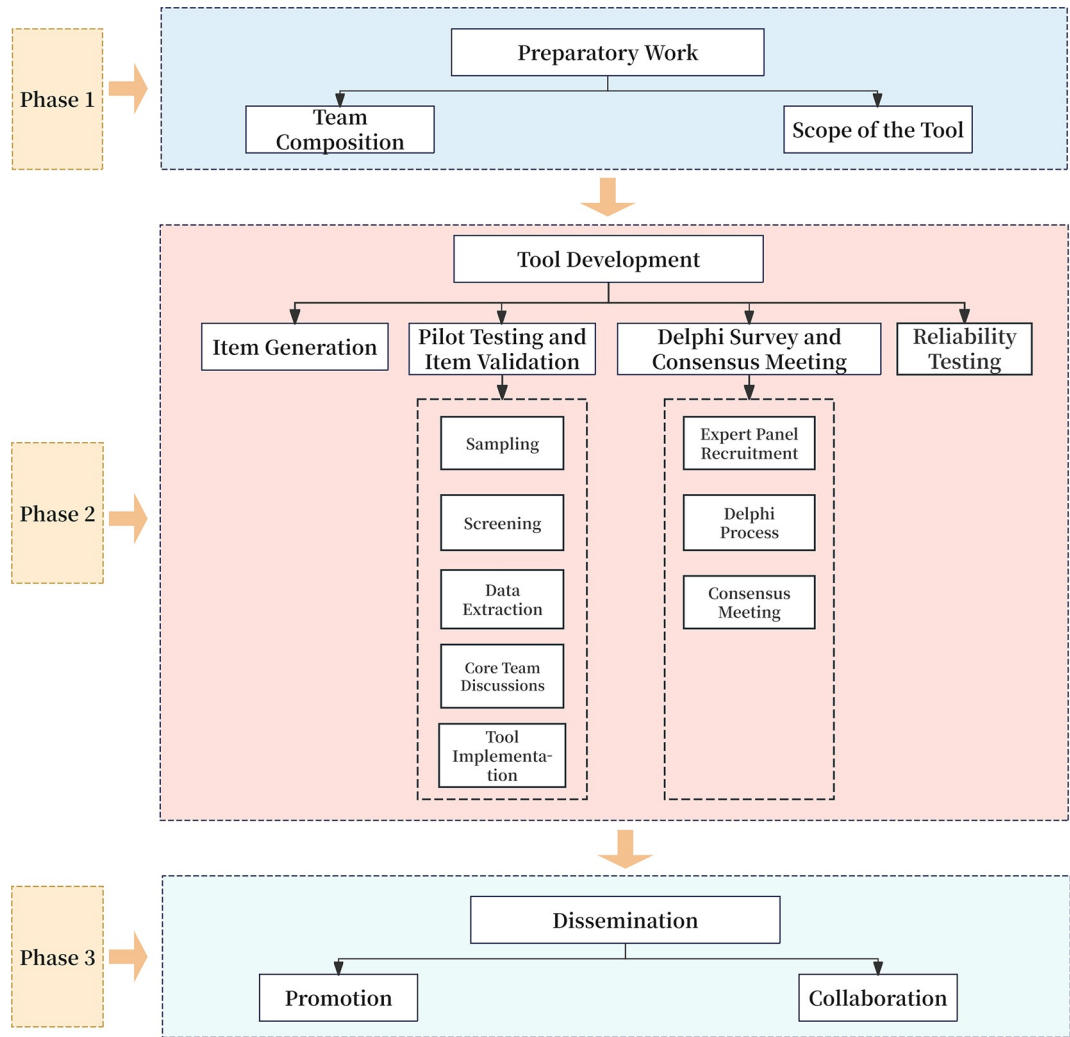


Figure 1 Three distinct phases of tool development

validation, Delphi survey, reliability testing);

- Phase 3: dissemination.

2.3. Phase 1: preparatory work

2.3.1. Team composition

A two-tier organizational structure will be established. The core working group will be responsible for drafting the initial versions and coordinating daily activities. This group comprises researchers with expertise in evidence-based medicine, clinical practice (three members with multiple years of experience), medical statistics (one member), and SR methodology (one graduate student with two years of experience). Patient or public representatives are not included in the working group. The extended expert panel will be responsible for critical decisions through a modified Delphi process, including item selection and rating criteria. The core group will compile, analyze, and revise drafts after each Delphi round, providing structured feedback until stable consensus is achieved. Unresolved disagreements will be addressed through online consensus meetings.

2.3.2. Scope of the tool

The SRD tool will target intervention-based SRs that evaluate the effectiveness and safety of specific interventions (e.g., medications, surgery, rehabilitation, acupuncture) for predetermined outcomes. Network meta-analyses and umbrella reviews also fall within its scope. While the tool will primarily focus on randomized controlled trials (RCTs), non-RCTs and observational studies may be included when evidence is limited or when required by specific research questions. Duplication will be assessed between pairs of SRs addressing the same disease or condition.

2.3.3. Target users and application scenarios

The tool is intended for (I) SR authors (both experienced and prospective), to assess whether a proposed topic has been adequately addressed, thereby avoiding redundancy and justifying new syntheses; (II) journal editors and peer reviewers, to evaluate the novelty and duplication risk of submitted manuscripts and support editorial decisions; (III) guideline developers, to select the most appropriate and current SRs as evidence bases for clinical recommendations; and (IV) health technology assessment personnel, to streamline the identification of relevant, non-duplicative evidence syntheses for reimbursement and policy decisions.

The tool supports evidence-based decision-making in four key contexts: (I) pre-project planning, in which researchers compare their planned SR (PICO, methods, etc.) with published reviews to determine whether the proposed work is genuinely needed; (II) comparative assessment of existing reviews, in which users evaluate duplication among multiple reviews addressing similar questions and, together with quality and publication date, select the most reliable and timely review; (III) manuscript evaluation, in which editors and reviewers objectively assess a submission's novelty relative to prior work, thereby informing acceptance or revision decisions; and (IV) evidence mapping and gap analysis, in which application of the tool to a body of literature (e.g., SRs on a specific intervention or condition) may help identify clusters of duplicative reviews as well as genuine evidence gaps, thereby supporting research prioritization and funding allocation.

2.4. Phase 2: tool development

2.4.1. Item generation

Existing tools for assessing the quality of SRs or meta-analyses will be reviewed. Items will be categorized according to factors that may influence duplication (e.g., research topic, inclusion criteria, search strategy, assessment process, conclusions). Databases including PubMed, Web of Science, and the Cochrane Library will be searched for studies addressing SR duplication, overlap, research waste, and updates of SRs. Building on the foundation established above and drawing on PRISMA 2020 and AMSTAR 2, the

Cochrane Handbook for Systematic Reviews of Interventions will serve as the primary reference framework. Items will be prioritized if they are theoretically justified, commonly reported, and feasible to assess.

The SRD tool assesses duplication across SRs using a multidimensional framework that generates a qualitative duplication profile rather than a single numerical score. The tool evaluates four key domains: (I) Research topic: compares the scope and core PICO to determine whether reviews address the same clinical question; (II) Research methods: assesses similarity in search strategies, selection criteria, data extraction, synthesis approaches, and effect measures; (III) Research results: examines overlap in included primary studies and consistency of findings, including effect estimates and subgroup analyses; and (IV) Methodological quality: evaluates risk of bias and evidence certainty using established tools (e.g., AMSTAR 2, ROBIS, Grading of Recommendations Assessment, Development and Evaluation [GRADE]), enabling users to distinguish between redundant low-quality reviews and methodologically rigorous updates. The final output consists of two components: domain-level judgments, which indicate the degree of overlap or divergence for each of the four domains, and an overall interpretive category that synthesizes these findings into a summary assessment.

2.4.2. Pilot testing and item validation

Sampling strategy

SRs will be sampled across disease categories based on MeSH (Medical Subject Headings) classification. To ensure balanced representation, a stratified sampling approach will be used. Within each disease category, a maximum of 10–15 reviews will be randomly selected, with oversampled categories capped to prevent dominance. For categories containing fewer eligible reviews, all will be included. Disease categories that do not align well with standard classifications will be adjusted accordingly.

Simple random sampling will be conducted using the R programming language. For each database search result, SRs will be selected based on automatically generated sequential numbers. If titles clearly indicate ineligibility, the next sequentially numbered record will be included.

Screening process

Citations will be imported into Rayyan (<https://www.rayyan.ai/>) for screening and removal of duplicate publications by two reviewers independently. Screening of titles, abstracts, and publication types, as well as topic duplication analysis, will be conducted in Excel. Finally, full-text screening will be conducted for all included publications.

Eligibility criteria

The inclusion criteria include (I) SRs/meta-analyses evaluating intervention effectiveness, including empirical studies, with mixed designs potentially including both randomized and non-randomized studies; (II) explicit identification as an SR, with or without meta-analysis; (III) SRs in the medical field; (IV) formal publication (journal articles or book chapters) with a clear publication date, authorship, and journal or publisher information, excluding preprints, conference abstracts, and dissertations; and (V) accessible full text, or at least a complete methodological section, enabling extraction of the required information.

The exclusion criteria include (I) diagnostic test accuracy reviews, prognostic reviews, methodological reviews, narrative reviews, scoping reviews, living systematic reviews, commentaries, perspectives, and case reports; (II) protocols and other non-SR literature; (III) veterinary medicine, animal medicine, social science (e.g., pure medical ethics reviews), or other non-medical fields; (IV) duplicate publications of the same review appearing in different journals, as well as versions differing only by language or format, will be handled by retaining a single version. Priority should be given to the earliest publication; if the earliest publication is unavailable or of questionable quality, the version published in the highest-impact journal should be retained instead; (V) inaccessible full text or missing critical information (e.g., unreported search strategy or unclear research question); and (VI) retracted articles.

Data extraction

Data extraction will involve both textual descriptions and rating-level

selections for each item. Different items will employ different extraction approaches.

Type A: Standardized coding (e.g., disease classification using International Classification of Diseases codes to prevent mislabeling).

Type B: Predefined numerical coding (e.g., 1 = Western medicine, 2 = traditional medicine, 3 = physical therapy for interventions).

Type C: Free-text description (when standardized selection is infeasible, researchers will make independent judgments).

Two reviewers will independently complete the data extraction without a prior calibration exercise. Any disagreements will be resolved by a senior researcher in consultation with two additional senior researchers.

Core team discussions

The core research team will conduct iterative discussions to ensure accuracy. Team members will independently propose item modifications based on pilot testing results and the relevant literature, followed by face-to-face consensus-building. Through multiple rounds, the final assessment dimensions, item list, and guidance documents will be established.

Tool implementation

Following iterative refinement, the project team will develop version 1.0 with accompanying user guidance. The tool will be implemented as web-based and/or Excel-based applications, requiring close integration of medical requirements with technical implementation. The medical team will define the core duplication assessment indicators and evaluation logic, including input formats and conditional branching, while technical personnel will design the interactive models. Joint reviews will ensure that medical evaluation practices align with technical feasibility.

In all scenarios, the tool will be designed for use either in its full version, for in-depth duplication analysis, or in a simplified version, for rapid screening, thereby ensuring broad applicability across different user needs and time constraints. Expert consensus and reliability testing will be conducted following the Delphi survey to finalize the tool, as detailed below.

Full Version: This will be available as both an online application and an Excel-based tool and will contain all domains and items for in-depth duplication assessment.

Simplified Version: This version will focus on the core predictive items or key domains that are most critical for rapidly identifying duplication. It will be designed for users who need a quick screening tool, such as clinicians or researchers conducting rapid reviews. Clear guidance will also be provided on how to interpret the results of the simplified version.

2.4.3. Modified Delphi survey and consensus meeting

Expert panel recruitment

Experts will be identified through screening the authors of relevant publications, social media searches, previous collaborations, and peer recommendations. The panel is expected to include approximately 15–20 participants, with balanced representation across key stakeholder groups from diverse settings, including universities, primary care institutions, hospitals, medical journals, patient organizations, and both non-profit and for-profit organizations. Attention will also be given to geographic distribution and other diversity factors.

Delphi process

Experts will receive the initial tool draft, guidance documents, and participant testing results. The Delphi survey will proceed in two rounds. In Round 1, a questionnaire using a 5-point Likert scale for item importance (*not at all important*, *not important*, *neutral*, *important*, and *very important*) will be used. Items will be retained if $\geq 70\%$ of experts rate them as combined categories. Open-ended questions will be used to solicit feedback on item content, definitions, and scope. Responses will be organized, categorized, and preliminarily revised to form a structured item list.

In Round 2, a questionnaire using a 3-point scale (*agree*, *neutral*, and *disagree*) will be used together with summary statistics from Round 1 (median, interquartile range, and group distribution). Experts will re-rate items with knowledge of group consensus, while maintaining the $\geq 70\%$ agreement threshold.

Consensus meeting

If unresolved issues persist, an online expert consultation meeting will be convened for open discussion of assessment dimensions, item content, terminology definitions, wording, and rating criteria. Expert feedback will inform revisions to the dimensions and items. Consensus items will be retained, whereas contentious items will be further discussed or removed, yielding a final expert-endorsed item set.

2.4.4. Reliability testing

Purposive sampling will be used to recruit participants with experience in SRs and methodological data extraction. To ensure practical experience, first authors of SRs (≥ 2 RCT-based SRs published from 2020 to 2024 in the Cochrane Database) will be invited. Participants will include graduate students, junior researchers, and senior methodologists. All participants will receive the tool and guidance documents.

Four independent reviewers will assess each sample using the tool to measure inter-rater reliability. The intraclass correlation coefficient (ICC) will be used for item-level ratings, while Kendall's W will be used to assess concordance among reviewers for the overall duplication classification (i.e., the qualitative profile generated by the tool). An ICC ≥ 0.75 or Kendall's W > 0.7 will indicate satisfactory inter-rater agreement. Assessment time and operational difficulty will be recorded to evaluate usability.

Semi-structured interviews will be conducted to collect participant feedback on tool content, workflow, item comprehensibility, and rating criteria. Interview data will undergo thematic analysis to identify key issues, thereby informing revisions to the tool and guidance documents through team and methodologist discussions.

The final version of the tool will incorporate feedback from team discussions, Delphi surveys, expert consultations, and user testing.

2.5. Phase 3: dissemination

The SRD tool will be disseminated through presentations at academic conferences, training courses, and methodology workshops; collaboration with professional associations (e.g., GRADE Working Group, Cochrane Methods Group); and an official website or hosting on research networks (e.g., Enhancing the QUALity and Transparency Of health Research Network [EQUATOR] Network). To support uptake and proper use, dissemination will also include the development of user manuals, training modules, and example case studies.

3. Discussion

The SRD tool is designed to define the concept of duplication in evidence synthesis by translating abstract concerns regarding overlap and redundancy into structured, assessable domains. By decomposing duplication into four domains (research topic, research methods, research results, and research quality), the tool acknowledges that duplication is multidimensional rather than a binary phenomenon. Two SRs may share an identical research question yet differ meaningfully in eligibility criteria, analytical approaches, or certainty-of-evidence assessments, leading to distinct contributions. Conversely, SRs that differ superficially in phrasing may still rely on essentially identical datasets and methods. The SRD tool is therefore expected to facilitate nuanced differentiation between justified replication and problematic redundancy.

Existing instruments such as PRISMA 2020, AMSTAR 2, ROBIS, and ReMarQ primarily focus on transparency, methodological rigor, or risk of bias within individual SRs. The SRD tool is conceptually complementary to these instruments. Rather than evaluating whether a single SR is well conducted, the SRD tool evaluates the relationship between two SRs. This relational perspective represents a methodological extension of current appraisal paradigms and aligns with growing attention to research waste and evidence ecosystem efficiency.

In this sense, the SRD tool may serve as an upstream decision-support instrument. For example, prior to initiating a new SR, investigators could

apply the tool to existing reviews to determine whether meaningful knowledge gaps remain. Similarly, journal editors and peer reviewers could use the tool alongside conventional appraisal instruments to inform decisions about novelty and contribution.

For researchers, the SRD tool may support more strategic topic selection, clinical trial planning, and protocol development. For editors and peer reviewers, it offers a transparent framework for evaluating novelty. For guideline developers and health technology assessment bodies, it may assist in selecting among multiple SRs or determining whether new syntheses are warranted. Within the field of Chinese medicine, where heterogeneous interventions and diverse outcome measures are common,²² a structured approach to evaluating duplication may be particularly valuable for promoting coherent evidence accumulation and avoiding fragmented syntheses.

Nevertheless, several limitations warrant consideration. First, judgments regarding duplication inevitably involve a degree of subjectivity, particularly when assessing conceptual similarity across research topics or qualitative distinctions in methodological approaches. Second, applying the tool across highly heterogeneous review designs, such as those differing in intervention complexity, outcome diversity, or analytic frameworks, may pose challenges to consistent interpretation. Third, the tool's performance may vary depending on the completeness and clarity of reporting in the original reviews. Further improvement is needed to enable reliable application across diverse studies.

4. Conclusion

The SRD tool will provide a rigorous, standardized approach to assessing duplication between SRs. By following established frameworks for tool development and engaging diverse stakeholders throughout the process, we aim to create an instrument that will meaningfully contribute to improving the efficiency and quality of evidence synthesis. This protocol ensures transparency and methodological rigor in the tool's development, providing a foundation for future efforts to optimize research resource allocation and enhance evidence-based decision-making.

Funding

The study was sponsored by the Project of Administration of Traditional Chinese Medicine of Guangdong Province (grant no.: 20251280), Guangzhou Health Science and Technology Project (grant no.: 2024A011005), Guangzhou Science and Technology Fund (grant nos.: 2024A03J0791, 2025A03J3510, 2025A03J3512, and 2025A03J3428), Guangzhou Key Science and Technology Project of TCM (grant no.: 2025ZD010), Plan on Enhancing Scientific Research in GMU (grant nos.: GMUCR 2024-01018 and GMUCR2024-02030), and the Young Scientific and Technological Talents Research Project of the Affiliated Traditional Chinese Medicine Hospital (grant nos.: 2022RC07, 2024SZYRC13 and 2024SZYRC15).

Data Availability

Not applicable.

Author Contribution Statement

Yue-ning Deng: Conceptualization, methodology, writing – original draft, and visualization. **Si-qi Wu:** Conceptualization, methodology, and writing – review & editing. **Xu-fei Luo:** Investigation and methodology. **Bin-bin Tang:** Data curation. **Yu-cheng Feng:** Data curation. **Hao-yue Feng:** Data curation. **Hao-lin Guo:** Data curation. **Yi-ming Li:** Data curation. **Lin Shi:** Supervi-

sion. **Zhi-rui Xu:** Supervision. **Zi-feng Zhou:** Supervision. **Lin Yu:** Project administration and funding acquisition. **Yao-long Chen:** Conceptualization, methodology, and project administration. **Yu-ting Duan:** Conceptualization, methodology, project administration, funding acquisition, and writing – review & editing.

Use of AI Statement

None.

Declaration of Competing Interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

- [1] Sur R, Dahm P. History of evidence-based medicine. *Indian J Urol.* 2011;27(4):487.
- [2] Kranke P. Evidence-based practice: how to perform and use systematic reviews for clinical decision-making. *Eur J Anaesthesiol.* 2010;27(9):763–772.
- [3] Page MJ, Shamseer L, Altman DG, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. *PLoS Med.* 2016;13(5):e1002028.
- [4] Hoffmann F, Allers K, Rombey T, et al. Nearly 80 systematic reviews were published each day: observational study on trends in epidemiology and reporting over the years 2000–2019. *J Clin Epidemiol.* 2021;138:1–11.
- [5] Katsura M, Kuriyama A, Tada M, et al. High variability in results and methodological quality among overlapping systematic reviews on the same topics in surgery: a meta-epidemiological study. *Br J Surg.* 2021;108(12):1521–1529.
- [6] Riva N, Puljak L, Moja L, et al. Multiple overlapping systematic reviews facilitate the origin of disputes: the case of thrombolytic therapy for pulmonary embolism. *J Clin Epidemiol.* 2018;97:1–13.
- [7] Chapelle C, Ollier E, Bonjean P, et al. Replication of systematic reviews: is it to the benefit or detriment of methodological quality? *J Clin Epidemiol.* 2023;162: 98–106.
- [8] Nosek BA, Hardwicke TE, Moshontz H, et al. Replicability, robustness, and reproducibility in psychological science. *Annu Rev Psychol.* 2022;73:719–748.
- [9] Siontis KC, Hernandez-Boussard T, Ioannidis JPA. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ.* 2013;347:f4501.
- [10] Page MJ, Welch VA, Haddaway NR, Karunanathan S, Maxwell LJ, Tugwell P. “One more time”: why replicating some syntheses of evidence relevant to COVID-19 makes sense. *J Clin Epidemiol.* 2020;125: 179–182.
- [11] Puljak L, Lund H. Definition, harms, and prevention of redundant systematic reviews. *Syst Rev.* 2023;12(1):63.
- [12] Siontis KC, Ioannidis JPA. Replication, duplication, and waste in a quarter million systematic reviews and meta-analyses. *Circ Cardiovasc Qual Outcomes.* 2018;11(12):e005212.
- [13] Lunney C, Reid EK, Neelakant T, et al. A new taxonomy was developed for overlap across ‘overviews of systematic reviews’: a meta-research study of research waste. *Res Synth Meth.* 2022;13(3):315–329.
- [14] Karunanathan S, Grimshaw JM, Maxwell L, et al. Can a replication revolution resolve the duplication crisis in systematic reviews? *BMJ Evid Based Med.* 2024;29(5):285–288.
- [15] Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* 2016;94(3):485–514.
- [16] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021: n71.
- [17] Shea BJ, Reeves BC, Wells G, 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: j4008.
- [18] Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* 2016;69:225–234.
- [19] Marques-Cruz M, Vieira RJ, Martinho-Dias D, et al. Reported methodological quality of medical systematic reviews: development of an assessment tool (ReMarQ) and meta-research study. *Res Synth Methods.* 2025;16(1):175–193.
- [20] Simera I, Moher D, Hoey J, Schulz KF, Altman DG. The EQUATOR network and reporting guidelines: helping to achieve high standards in reporting health research studies. *Maturitas.* 2009;63(1):4–6.
- [21] Whiting P, Wolff R, Mallett S, Simera I, Savović J. A proposed framework for developing quality assessment tools. *Syst Rev.* 2017;6(1):204.
- [22] Zhang Z, Li R, Chen Y, et al. Integration of traditional, complementary, and alternative medicine with modern biomedicine: the scientization, evidence, and challenges for integration of traditional Chinese medicine. *Acupunct Herb Med.* 2024;4(1):68–78.