

How to perform and write a systematic review and meta-analysis

Gaeun Kim

Professor, College of Nursing · Research Institute of Nursing Science, Keimyung University, Daegu, Korea

We are regularly faced with opportunities to decide on the most effective interventions for our patients. As evidence-based practice becomes increasingly important in nursing and healthcare decision-making and policy formulation, there has been a marked rise in interest in systematic reviews and meta-analyses. Indeed, the quality and volume of systematic reviews and meta-analyses published in nursing and medical journals have seen significant growth and widespread acceptance.

Systematic reviews and meta-analyses are highly regarded in the hierarchical structure of scientific evidence for verifying effectiveness [1,2]. They provide a clear identification of the benefits and harms of interventions, making them a valuable starting point for developing clinical practice guidelines. Consequently, they are becoming some of the most frequently cited publications today [3]. However, despite the quantitative increase in systematic reviews and meta-analyses, there are significant concerns regarding quality and reproducibility. As indicated by several studies [4-6], while the volume of literature on systematic reviews and meta-analyses has grown, the quality of these studies often remains low. The quality of systematic reviews and meta-analyses is directly tied to the rigor of selection and analysis, as flawed meta-analyses can result in inaccurate and misleading conclusions.

Various review methodologies are currently emerging that share characteristics with literature reviews, but have different objectives than a systematic review. According to a study by Grant & Booth, there are approximately 14 types of reviews that resemble a systematic review [7]. Among these, scoping

reviews, also known as mapping reviews, are frequently used today to categorize or group the scale, scope, nature, and characteristics of research within broader subjects of interest. A scoping review is also often conducted as a preliminary step before determining the need for a systematic review [8]. These new types of reviews generally follow a process similar to that of a systematic review. Having a prior understanding of a systematic review and meta-analysis is considered beneficial when selecting a review methodology that is suitable for a specific research purpose.

This paper provides a comprehensive summary of the key points to consider for systematic reviews and meta-analyses to ensure the utility of publications. It covers an overall conceptual understanding, reporting guidelines, and methodological procedures.

1. When and Why Do We Conduct Systematic Reviews and Meta-Analyses?

Systematic reviews and meta-analyses can be utilized to ascertain the certainty of an intervention's effect and provide evidence for its efficacy. Furthermore, by pinpointing the limitations of current research, these methods offer an opportunity to instigate new studies. Existing research, conducted with the same objectives, may yield similar, conflicting, or ambiguous conclusions. In such instances, these methods can be employed to scrutinize the final conclusion from the current perspective of previously conducted research on the same topic. They can also enhance the strength of research findings and refine the accuracy of the study.

2. Understanding the Terminology

A systematic review and meta-analysis is a research methodology that involves the selection, review, critical evaluation, data collection, and analysis of all available literature using a predetermined, systematic, and explicit method to answer specific research questions [9]. Unlike traditional narrative reviews, a systematic review is characterized by its explicitness, reproducibility, and reliance on a transparent

Corresponding author

Gaeun Kim

College of Nursing · Research Institute of Nursing Science, Keimyung University, 1095 Dalgubeol-daero, Dalseo-gu, Daegu 42601, Korea
TEL: +82-53-258-7658
FAX: +82-53-258-3916
E-MAIL: eun0325@kmu.ac.kr

Received Jul 21, 2023 Accepted Jul 26, 2023

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

analysis method. A meta-analysis is a statistical method that quantitatively synthesizes and analyzes the results of various studies conducted on the same topic [9,10]. The potential benefits of synthesizing multiple data sets include improved power and precision compared to individual studies, the ability to identify the root cause of unresolved issues from individual studies, the capacity to reconcile conflicting research results, and the potential to generate new hypotheses.

3. Reporting Guidelines

Since research in the fields of nursing, medicine, clinical practice, and healthcare, including systematic reviews and meta-analyses, often assesses effects or efficacy through specific clinical trials or interventions, the transparency and accuracy of such research are of paramount importance. Therefore, it is recommended to adhere to guidelines for research planning, execution, and reporting for each research design. Currently, the EQUATOR (Enhancing the Quality and Transparency of Health Research) Network, an international research network hosted by the University of Oxford in the UK, has registered approximately 256 such guidelines. Notable examples include PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), CONSORT (Consolidated Standards of Reporting Trials), and STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) [11].

PRISMA is a set of guidelines that outlines the recommended items for reporting systematic reviews and meta-analyses. While its primary focus is on reporting reviews that assess the impact of interventions, it can also serve as a reporting guideline for reviews evaluating etiology, prevalence, diagnosis, or prognosis [12]. PRISMA is rooted in the QUOROM (Quality of Reporting of Meta-Analyses) statement, which was published in 1999, and was updated to PRISMA2020 on March 29, 2021 [12-14]. The PRISMA2020 checklist consists of a total of 27 items, the same number as in its previous version (2009), but with many details revised. PRISMA2020 includes an expanded checklist and a PRISMA2020 flow chart template [15,16], as well as requirements for writing components such as the title, abstract, introduction, background, and research objective. Notably, PRISMA2020 mandates the registration of the research protocol and the presentation of the database used for the search, the comprehensive search strategy for registries and websites, the number of reviewers who screened each record, and the method used to define results [12,14].

On the one hand, MOOSE (Meta-analysis of Observational Studies in Epidemiology) serves as a reporting guideline for non-randomized controlled trials [11]. The Cochrane organization also provides a methodological standard known as MECIR (Methodological Expectations of Cochrane Intervention

Reviews), which should be adhered to in the protocol, review, and updates published by Cochrane [9,17]. Furthermore, in June 2019, numerous details were enhanced and expanded in version 6.0 of the Cochrane Handbook, which is regarded as a methodological textbook for systematic reviews and meta-analyses. As of 2022, it has been updated to version 6.4.

4. Methodological Procedure

1) Selection of the review question

The first and most important step is to clarify the review question. The review question requires consideration of several key components which can often be encapsulated by the 'PICO' mnemonic, an acronym for Population, Intervention, Comparison, and Outcome [9,12]. It is necessary to identify which population to study, what intervention to study, which comparison group to involve, and what research results to review.

2) Writing a protocol

Writing a protocol is a crucial step in the systematic review process, and adhering to a pre-established protocol is a primary strategy to prevent selection bias. Although it is not yet standard practice to require a registration (approval) number for a protocol, an increasing number of journals are recommending it, since it can help reduce both publication bias and the risk of duplication in multiple systematic reviews that address the same questions [9,12]. The research protocol can be registered at either PROSPERO (International Prospective Register of Systematic Reviews) or Cochrane [17,18].

3) Literature search

To conduct a comprehensive and objective review of all available publications on a research topic, it is necessary to utilize major search databases such as PubMed (MEDLINE), EMBASE, and the Cochrane Central Register of Trials, along with other specialized databases (for example, CINAHL for nursing). Additionally, it is important to gather all pertinent research from gray literature, references from related studies, manual searches of major studies, conference abstracts, and reports. Two or more databases should be used as the primary search databases. One effective strategy for literature searches is the use of controlled vocabulary. For example, when searching literature in MEDLINE, researchers can employ MeSH (Medical Subject Headings), which are subject headings (a thesaurus) for the medical field, as determined by the National Library of Medicine in the United States. Furthermore, researchers can expand or narrow the results set by using synonyms, alternative words, Boolean operators, truncation, wildcard, and filters such as search limit, search field, and tag [9,12].

4) Inclusion and exclusion of literature

After conducting a literature search, researchers must then proceed to select relevant literature based on the research objectives and the established inclusion and exclusion criteria [9,12]. The outcome of this literature selection process is typically represented through a flow diagrams, with various types of templates available depending on the type of review and sources used to identify studies [16]. Selection bias is one of the most common problems in recent systematic reviews and to mitigate this, Cochrane recommends including literature in all languages [9,12].

5) Risk of bias assessment

The risk of bias must be evaluated in two contexts—first, in the outcomes of individual studies, and second, in the meta-analysis of the incorporated research findings [9,12]. Initially, the evaluation should scrutinize the potential bias risk in the research outcomes via an internal validity assessment of the included individual studies. Essentially, this means examining the risk of either overestimating or underestimating the effect.

Bias refers to systemic error, and several tools are commonly used in clinical research to assess the risk of this bias. Examples of such tools include Cochrane's Risk of Bias (RoB) tool, NOS (the Newcastle-Ottawa Scale), MINORS (Methodological Index for Non-Randomized Studies), QUADAS (Quality Assessment of Diagnostic Accuracy Studies), and QUADAS-2. Recently, Cochrane introduced 'revision of the RoB tool for randomized clinical trials (RoB 2)' and 'ROBINS-I (the Risk of Bias in Non-Randomized Studies of Interventions)', a risk of bias assessment tool for non-randomized clinical trials [19-22]. The RoB 2 tool divides the risk of bias assessment into several areas: bias arising from the randomization process, bias due to deviation from intended interventions, missing outcome data, risk of bias in the measurement of the outcome, and bias in the selection of the reported result. The result of the risk of bias assessment is determined by a mapping algorithm that verifies its outcome based on "signalling questions." The judgment can be categorized as "low" or "high" risk of bias, or it may be expressed as "some concerns" [20,22]. In addition, the meta-analysis result should also undergo a review for overall bias. Visual assessments often use funnel plots or contour-enhanced funnel plots. Statistical tests for bias, such as the Begg and Mazumdar test and the Egger test, can also be employed in the assessment [9,12].

6) Data extraction

Data extraction refers to the process of pulling data from individual studies for inclusion in the final analysis. Elements that may be extracted include information about the research

(such as the author and year of publication), the rationale for inclusion or exclusion, the research methodology, the subject of the research, the comparative intervention, the results of the intervention, and more. This is typically presented in the form of a "Characteristics of Included Studies" table in the paper [9,12].

7) Data analysis and presentation of results

Data analysis is a phase in which data is analyzed, synthesized, and research findings are summarized. This process can be divided into qualitative and quantitative synthesis. If quantitative synthesis is deemed feasible, a meta-analysis is performed. The outcome of the meta-analysis is displayed as a forest plot, featuring point estimates and confidence intervals. The pooled effect estimate is adjusted according to the weight value that contributes to the final analysis. This value is a combination of the point estimate and confidence interval from individual research [9,12]. Most statistical packages, whether free or subscription-based, offer meta-analysis capabilities. Widely used statistical software includes Cochrane Collaboration's Revman 5.3.5 (Nordic Cochrane Centre, The Cochrane Collaboration, 2014) [23], Comprehensive Meta-Analysis (CMA) [24], R (R Core Team), and Stata. The principle of a meta-analysis involves calculating the summary estimate of individual research, assigning value to each research study, estimating the weighted average, and calculating the integrated pooled effect estimate. A fixed-effect model and random-effect model are typically used for the statistical model of a meta-analysis when combining pooled effect estimates [9,12]. Pooled effect estimates in a meta-analysis are categorized into dichotomous data and continuous data. For dichotomous data, the relative risk, odds ratio, risk difference, hazard ratio, and number needed to treat are utilized. For continuous data, the mean difference and standardized mean difference are employed [9,12].

If a meta-analysis is not conducted properly, it could lead to distorted conclusions. Therefore, if the research included in the systematic review shows variation or inconsistency, if there is a potential for bias in the individual studies included, or if there is a serious risk of reporting bias, a meta-analysis should not be carried out.

In a meta-analysis, heterogeneity refers to any variation observed among the included studies that exceeds the sampling error and cannot be attributed to chance. Clinical heterogeneity may arise if there is variability in the results among research subjects or interventions. Methodological heterogeneity may occur if there is variability in the research design or if there is a risk of bias in the overall included research [9,12]. Statistical heterogeneity can result from clinical or methodological heterogeneity, or a combination of both. This leads to

variability in the intervention effect assessed in the included research. To verify and assess heterogeneity, visual investigation using graphs (e.g., forest plots, L'Abbé plots, Galbraith plots, etc.) or statistical tests (e.g., the Q statistic, Higgins's I-squared statistic, etc.) can be employed [9,12]. If heterogeneity is present, it is necessary to conduct a subgroup analysis or meta-regression, or consider statistical measures to identify the cause.

Sensitivity analysis is a method used to examine the robustness of the meta-analysis results. This evaluation is carried out by comparing the outcome of an alternative analysis, which differs from the chosen option, to the initial analysis result. For example, it could involve comparing the results of a fixed-effect model with those of a random-effect model, or conducting a comparative analysis based on the characteristics of the research.

8) Level of evidence assessment and drawing a conclusion

The quality of conclusions drawn from all systematic reviews and meta-analyses reflects the level of evidence supplied by each individual study included in the review. Consequently, it is crucial to clinically apply these conclusions only after assessing the advantages and disadvantages of each individual study and carefully interpreting the results.

The quality of results or the level of evidence is typically evaluated using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) tool [25-27] or AHRQ's Strength of Evidence. Factors considered in these evaluations include research design, risk of bias, volume of evidence, inconsistency in evidence, indirectness of evidence, imprecision of effect estimates, and the risk of publication bias. The GRADE method is the most commonly used for assessing the level of evidence, and its results can be displayed in terms of level (high, moderate, low, very low) and through a "summary of findings" table [9,26,27].

For the critical evaluation of systematic reviews, the AMSTAR (A Measurement Tool to Assess Systematic Reviews) was developed in 2007 and later updated to AMSTAR-2. This tool, consisting of 11 items, was designed to assess the quality of systematic reviews and determine their suitability [28].

CONCLUSION

If research adheres strictly to the standard rules for conducting a systematic review and meta-analysis, it minimizes potential bias and enhances the transparency and reliability of the results. A well-executed systematic review and meta-analysis of data can serve as a vital tool not only in clinical settings but also in educational domains, supporting evidence-based nursing care. This paper anticipates that by rigorously apply-

ing standardized methodologies in systematic reviews and meta-analyses, the scope for both quantitative and qualitative analysis of the study can be broadened. Moreover, it is expected to yield useful and accurate analysis results that can be applied in clinical practice, education, and healthcare policy-making.

ORCID and ResearcherID

Gaeun Kim <https://orcid.org/0000-0003-1914-8331>
<https://researcherid.com/rid/AFS-1703-2022>

Authors' contribution

All the work was done by Gaeun Kim.

Conflict of interest

No existing or potential conflict of interest relevant to this article was reported.

Funding

This study was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (No. 2022R1A2C1012697).

Data availability

Please contact the corresponding author for data availability.

Acknowledgements

None.

REFERENCES

1. Rubin A, Bellamy J. Practitioner's guide to using research for evidence-based practice. 3rd ed. Wiley; 2022. p. 48-49.
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *British Medical Journal*. 1996;312(7023):71-72.
<https://doi.org/10.1136/bmj.312.7023.71>
3. Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. *Journal of the American Medical Association*. 2005;293(19):2362-2366.
<https://doi.org/10.1001/jama.293.19.2362>
4. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *The Milbank Quarterly*. 2016;94(3):485-514.

- <https://doi.org/10.1111/1468-0009.12210>
5. Fontelo P, Liu F. A review of recent publication trends from top publishing countries. *Systematic Reviews*. 2018;7(1):147. <https://doi.org/10.1186/s13643-018-0819-1>
 6. Kim G, Baik SK. Overview and recent trends of systematic reviews and meta-analyses in hepatology. *Clinical and Molecular Hepatology*. 2014;20(2):137-150. <https://doi.org/10.3350/cmh.2014.20.2.137>
 7. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*. 2009;26(2):91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
 8. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*. 2005;8(1):19-32. <https://doi.org/10.1080/1364557032000119616>
 9. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022)* [Internet]. Cochrane; 2022 [cited 2023 July 3]. Available from: <https://www.training.cochrane.org/handbook>
 10. Granados-Duque V, García-Perdomo HA. Systematic review and meta-analysis: which pitfalls to avoid during this process. *International Brazilian Journal of Urology*. 2021;47(5):1037-1041. <https://doi.org/10.1590/s1677-5538.ijbu.2020.0746>
 11. Enhancing the QUALity and Transparency Of health Research (EQUATOR) Network and UK EQUATOR Centre. EQUATOR Network [Internet]. [cited 2023 July 1]. Available from: <https://www.equator-network.org>
 12. 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. *PLoS Medicine*. 2021;18(3):e1003583. <https://doi.org/10.1371/journal.pmed.1003583>
 13. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine*. 2009;151(4):W65-W94. <https://doi.org/10.7326/0003-4819-151-4-200908180-00136>
 14. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). PRISMA 2020 Statement [Internet]. 2021 [cited 2023 July 2]. Available from: <http://www.prisma-statement.org/PRISMAStatement/PRISMAStatement>
 15. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). PRISMA 2020 Checklist [Internet]. 2021 [cited 2023 July 2]. Available from: <http://www.prisma-statement.org/PRISMAStatement/Checklist>
 16. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). PRISMA 2020 Flow diagram template [Internet]. 2021 [cited 2023 July 2]. Available from: <http://www.prisma-statement.org/PRISMAStatement/FlowDiagram>
 17. Cochrane. Cochrane database of systematic reviews [Internet]. [cited 2023 July 2]. Available from: <https://www.cochranelibrary.com/cdsr/about-cdsr>
 18. National Institute for Health and Care Research (NIHR). PROSPERO: International Prospective Register of Systematic Reviews [Internet]. [cited 2023 July 1]. Available from: <https://www.crd.york.ac.uk/PROSPERO>
 19. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*. 2011;343:d5928. <https://doi.org/10.1136/bmj.d5928>
 20. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *British Medical Journal*. 2019;366:l4898. <https://doi.org/10.1136/bmj.l4898>
 21. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *British Medical Journal*. 2016;355:i4919. <https://doi.org/10.1136/bmj.i4919>
 22. Cochrane. Cochrane methodology. Risk of bias tool for randomized trials (RoB 2) [Internet]. [cited 2023 July 1]. Available from: <https://training.cochrane.org/online-learning/cochrane-methodology/risk-bias>
 23. Cochrane. Review Manager 5 (RevMan 5) [Internet]. 2014 [cited 2023 July 1]. Available from: <https://community.cochrane.org/help/tools-and-software/revman-5>
 24. Biostat. Comprehensive meta-analysis (CMA) [Internet]. [cited 2023 July 2]. Available from: <https://ko.meta-analysis.com/>
 25. 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. *Journal of Clinical Epidemiology*. 2011;64(4):383-394. <https://doi.org/10.1016/j.jclinepi.2010.04.026>
 26. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al.; GRADE Working Group. Grading quality of evidence and strength of recommendations. *British Medical Journal*. 2004;328(7454):1490. <https://doi.org/10.1136/bmj.328.7454.1490>
 27. GRADEpro GDT: GRADEpro guideline development tool [software]. 2015 [cited 2023 July 1]. Available from: <https://grade.pro>
 28. 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. *British Medical Journal*. 2017;358:j4008. <https://doi.org/10.1136/bmj.j4008>