REVIEW ARTICLE

WILEY

A brief introduction of meta-analyses in clinical practice and research

Xiao-Meng Wang | Xi-Ru Zhang | Zhi-Hao Li | Wen-Fang Zhong | Pei Yang | Chen Mao

Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou, Guangdong, China

Correspondence
Chen Mao, Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou 510515, Guangdong, China.
Email: maochen9@smu.edu.cn

Abstract

With the explosive growth of medical information, it is almost impossible for healthcare providers to review and evaluate all relevant evidence to make the best clinical decisions. Meta-analyses, which summarize all existing evidence and quantitatively synthesize individual studies, have become the best available evidence for informing clinical practice. This article introduces the common methods, steps, principles, strengths and limitations of meta-analyses and aims to help healthcare providers and researchers obtain a basic understanding of meta-analyses in clinical practice and research.

KEYWORDS
clinical research, meta-analysis

1 | INTRODUCTION

With the explosive growth of medical information, it has become almost impossible for healthcare providers to review and evaluate all related evidence to inform their decision making.[1,2] Furthermore, the inconsistent and often even conflicting conclusions of different studies can confuse these individuals. Systematic reviews were developed to resolve such situations, which comprehensively and systematically summarize all relevant empirical evidence.[3] Many systematic reviews contain meta-analysis, which use statistical methods to combine the results of individual studies.[4] Through meta-analyses, researchers can objectively and quantitatively synthesize results from different studies and increase the statistical strength and precision for estimating effects.[5] In the late 1970s, meta-analysis began to appear regularly in the medical literature.[6] Subsequently, a plethora of meta-analyses have emerged and the growth is exponential over time.[7] When conducted properly, a meta-analysis of medical studies is considered as decisive evidence because it occupies a top level in the hierarchy of evidence.[8]

An understanding of the principles, performance, advantages and weaknesses of meta-analyses is important. Therefore, we aim to provide a basic understanding of meta-analyses for clinicians and researchers in the present article by introducing the common methods, principles, steps, strengths and limitations of meta-analyses.

2 | COMMON META-ANALYSIS METHODS

There are many types of meta-analysis methods (Table 1). In this article, we mainly introduce five meta-analysis methods commonly used in clinical practice.

2.1 | Aggregated data meta-analysis

Although more information can be obtained based on individual participant-level data from original studies, it is usually impossible to obtain these data from all included studies in meta-analysis

---

Xiao-Meng Wang and Xi-Ru Zhang contributed equally to this work.

TABLE 1 Meta-analysis methods

Methods	Definitions
Aggregate data meta-analysis	Extracting summary results of studies available in published accounts
Individual participant data meta-analysis	Collecting individual participant-level data from original studies
Cumulative meta-analysis	Adding studies to a meta-analysis based on a predetermined order
Network meta-analysis	Combining direct and indirect evidence to compare the effectiveness between different interventions
Meta-analysis of diagnostic test accuracy	Identifying and synthesizing evidence on the accuracy of tests
Prospective meta-analysis	Conducting meta-analysis for studies that specify research selection criteria, hypotheses and analysis, but for which the results are not yet known
Sequential meta-analysis	Combining the methodology of cumulative meta-analysis with the technique of formal sequential testing, which can sequentially evaluate the available evidence at consecutive interim steps during the data collection
Meta-analysis of the adverse events	Following the basic meta-analysis principles to analyze the incidences of adverse events of studies

because such data may have been corrupted, or the main investigator may no longer be contacted or refuse to release the data. Therefore, by extracting summary results of studies available in published accounts, an aggregate data meta-analysis (AD-MA) is the most commonly used of all the quantitative approaches.[9] A study has found that > 95% of published meta-analyses were AD-MA.[10] In addition, AD-MA is the mainstay of systematic reviews conducted by the US Preventive Services Task Force, the Cochrane Collaboration and many professional societies.[9] Moreover, AD-MA can be completed relatively quickly at a low cost, and the data are relatively easy to obtain.[11,12] However, AD-MA has very limited control over the data. A challenge with AD-MA is that the association between an individual participant-level covariate and the effect of the interventions at the study level may not reflect the individual-level effect modification of that covariate.[13] It is also difficult to extract sufficient compatible data to undertake meaningful subgroup analyses in AD-MA.[14] Furthermore, AD-MA is prone to ecological bias, as well as to confounding from variables not included in the model, and may have limited power.[15]

2.2 | Individual participant data meta-analysis

An individual participant data meta-analysis (IPD-MA) is considered the "gold standard" for meta-analysis; this type of analysis collects individual participant-level data from original studies.[15] Compared with AD-MA, IPD-MA has many advantages, including improved data quality, a greater variety of analytical types that can be performed and the ability to obtain more reliable results.[16,17]

It is crucial to maintain clusters of participants within studies in the statistical implementation of an IPD-MA. Clusters can be retained during the analysis using a one-step or two-step approach.[18] In the one-step approach, the individual participant data from all studies are modeled simultaneously, at the same time as accounting for the clustering of participants within studies.[19] This approach requires a model specific to the type of data being synthesized and an appropriate account of the meta-analysis assumptions (e.g. fixed or random effects across studies). Cheng et al.[20] proposed using a one-step IPD-MA to handle binary rare events and found that this method was superior to traditional methods of inverse variance, the Mantel–Haenszel method and the Yusuf-Peto method. In the two-step approach, the individual participant data from each study are analyzed independently for each separate study to produce aggregate data for each study (e.g. a mean treatment effect estimate and its standard error) using a statistical method appropriate for the type of data being analyzed (e.g. a linear regression model might be fitted for continuous responses, or a Cox regression might be applied for time-to-event data). The aggregate data are then combined to obtain an summary effect in the second step using a suitable model, such as weighting studies by the inverse of the variance.[21] For example, using a two-step IPD-MA, Grams et al.[22] found that apolipoprotein-L1 kidney-risk variants were not associated with incident cardiovascular disease or death independent of kidney measures.

Compared to the two-step approach, the one-step IPD-MA is recommended for small meta-analyses[23] and, conveniently, must only specify one model; however, this requires careful distinction of within-study and between-study variability.[24] The two-step IPD-MA is more laborious, although it allows the use of traditional, well-known meta-analysis techniques in the second step, such as those used by the Cochrane Collaboration (e.g. the Mantel–Haenszel method).

2.3 | Cumulative meta-analysis

Meta-analyses are traditionally used retrospectively to review existing evidence. However, current evidence often undergoes several updates as new studies become available. Thus, updated data must be continuously obtained to simplify and digest the ever-expanding literature. Therefore, cumulative meta-analysis was developed, which adds studies to a meta-analysis based on a predetermined order and then tracks the magnitude of the mean effect and its variance.[25] A cumulative meta-analysis can be performed multiple times; not only can it obtain summary results and provide a comparison of the dynamic results, but also it can assess the impact of newly added studies on the overall conclusions.[26] For example, initial observational studies and systematic reviews and meta-analyses suggested that frozen embryo transfer was better for mothers and babies; however, recent primary studies have begun to challenge these conclusions.[27] Maheshwari et al.[27] therefore conducted a cumulative meta-analysis

to investigate whether these conclusions have remained consistent over time and found that the decreased risks of harmful outcomes associated with pregnancies conceived from frozen embryos have been consistent in terms of direction and magnitude of effect over several years, with an increasing precision around the point estimates. Furthermore, continuously updated cumulative meta-analyses may avoid unnecessary large-scale randomized controlled trials (RCTs) and prevent wasted research efforts.[28]

2.4 | Network meta-analysis

Although RCTs can directly compare the effectiveness of interventions, most of them compare the effectiveness of an intervention with a placebo, and there is almost no direct comparison between different interventions.[29,30] Network meta-analyses comprise a relatively recent development that combines direct and indirect evidence to compare the effectiveness between different interventions.[31] Evidence obtained from RCTs is considered as direct evidence, whereas evidence obtained through one or more common comparators is considered as indirect evidence. For example, when comparing interventions A and C, direct evidence refers to the estimate of the relative effects between A and C. When no RCTs have directly compared interventions A and C, these interventions can be compared indirectly if both have been compared with B (placebo or some standard treatments) in other studies (forming an A–B–C "loop" of evidence).[32,33]

A valid network meta-analysis can correctly combine the relative effects of more than two studies and obtain a consistent estimate of the relative effectiveness of all interventions in one analysis.[34] This meta-analysis may lead to a greater accuracy of estimating intervention effectiveness and the ability to compare all available interventions to calculate the rank of different interventions.[34,35] For example, phosphodiesterase type 5 inhibitors (PDE5-Is) are the first-line therapy for erectile dysfunction, although there are limited available studies on the comparative effects of different types of PDE5-Is.[36] Using a network meta-analysis, Yuan et al.[36] calculated the absolute effects and the relative rank of different PDE5-Is to provide an overview of the effectiveness and safety of all PDE5-Is.

Notably, a network meta-analysis should satisfy the transitivity assumption, in which there are no systematic differences between the available comparisons other than the interventions being compared[37]; in other words, the participants could be randomized to any of the interventions in a hypothetical RCT consisting of all the interventions included in the network meta-analysis.

2.5 | Meta-analysis of diagnostic test accuracy

Sensitivity and specificity are commonly used to assess diagnostic accuracy. However, diagnostic tests in clinical practice are rarely 100% specific or sensitive.[38] It is difficult to obtain accurate estimates of sensitivity and specificity in small diagnostic accuracy studies.[39,40]

Even in a large sample size study, the number of cases may still be small as a result of the low prevalence. By identifying and synthesizing evidence on the accuracy of tests, the meta-analysis of diagnostic test accuracy (DTA) provides insight into the ability of medical tests to detect the target diseases[41]; it also can provide estimates of test performance, allow comparisons of the accuracy of different tests and facilitate the identification of sources of variability.[42] For example, the FilmArray® (Biomerieux, Marcy-l'Étoile, France) meningitis/encephalitis (ME) panel can detect the most common pathogens in central nervous system infections, although reports of false positives and false negatives are confusing.[43] Based on meta-analysis of DTA, Tansarli et al.[43] calculated that the sensitivity and specificity of the ME panel were both > 90%, indicating that the ME panel has high diagnostic accuracy.

3 | HOW TO PERFORM A META-ANALYSIS

3.1 | Frame a question

Researchers must formulate an appropriate research question at the beginning. A well-formulated question will guide many aspects of the review process, including determining eligibility criteria, searching for studies, collecting data from included studies, structuring the syntheses and presenting results.[44] There are some tools that may facilitate the construction of research questions, including PICO, as used in clinical practice[45]; PEO and SPICE, as used for qualitative research questions[46,47]; and SPIDER, as used for mixed-methods research.[48]

3.2 | Form the search strategy

It is crucial for researchers to formulate a search strategy in advance that includes inclusion and exclusion criteria, as well as a standardized data extraction form. The definition of inclusion and exclusion criteria depends on established question elements, such as publication dates, research design, population and results. A reasonable inclusion and exclusion criteria will reduce the risk of bias, increase transparency and make the review systematic. Broad criteria may increase the heterogeneity between studies, and narrow criteria may make it difficult to find studies; therefore, a compromise should be found.[49]

3.3 | Search of the literature databases

To minimize bias and reduce hampered interpretation of outcomes, the search strategy should be as comprehensive as possible, employing multiple databases, such as PubMed, Embase, Cochrane Central Registry of Controlled Trials, Scopus, Web of Science and Google Scholar.[50,51] Removing language restrictions and actively searching for non-English bibliographic databases may also help researchers to perform a comprehensive meta-analysis.[52]

3.4 | Select the articles

The selection or rejection of the included articles should be guided by the criteria.[53] Two independent reviewers may screen the included articles, and any disagreements should be resolved by consensus through discussion. First, the titles and abstracts of all relevant searched papers should be read, and inclusion or exclusion criteria applied to determine whether these papers meet. Then, the full texts of the included articles should be reviewed once more to perform the rejection again. Finally, the reference lists of these articles should be searched to widen the research as much as possible.[54]

3.5 | Data extraction

A pre-formed standardized data extraction form should be used to extract data of included studies. All data should be carefully converted using uniform standards. Simultaneous extraction by multiple researchers might also make the extracted data more accurate.

3.6 | Assess quality of articles

Checklists and scales are often used to assess the quality of articles. For example, the Cochrane Collaboration's tool[55] is usually used to assess the quality of RCTs, whereas the Newcastle Ottawa Scale[56] is one of the most common method to assess the quality of non-randomized trials. In addition, Quality Assessment of Diagnostic Accuracy Studies 2[57] is often used to evaluate the quality of diagnostic accuracy studies.

3.7 | Test for heterogeneity

Several methods have been proposed to detect and quantify heterogeneity, such as Cochran's Q and I^2 values. Cochran's Q test is used to determine whether there is heterogeneity in primary studies or whether the variation observed is due to chance,[58] but it may be underpowered because of the inclusion of a small number of studies or low event rates.[59] Therefore, $p < 0.10$ (not 0.05) indicates the presence of heterogeneity given the low statistical strength and insensitivity of Cochran's Q test.[60] Another common method for testing heterogeneity is the I^2 value, which describes the percentage of variation across studies that is attributable to heterogeneity rather than chance; this value does not depend on the number of studies.[61] I^2 values of 25%, 50% and 75% are considered to indicate low, moderate and high heterogeneity, respectively.[60]

3.8 | Estimate the summary effect

Fixed effects and random effects models are commonly used to estimate the summary effect in a meta-analysis.[62] Fixed effects models, which consider the variability of the results as "random variation", simply weight individual studies by their precision (inverse of the variance). Conversely, random effects models assume a different underlying effect for each study and consider this an additional source of variation that is randomly distributed. A substantial difference in the summary effect calculated by fixed effects models and random effects models will be observed only if the studies are markedly heterogeneous (heterogeneity $p < 0.10$) and the random effects model typically provides wider confidence intervals than the fixed effect model.[63,64]

3.9 | Evaluate sources of heterogeneity

Several methods have been proposed to explore the possible reasons for heterogeneity. According to factors such as ethnicity, the number of studies or clinical features, subgroup analyses can be performed that divide the total data into several groups to assess the impact of a potential source of heterogeneity. Sensitivity analysis is a common approach for examining the sources of heterogeneity on a case-by-case basis.[65] In sensitivity analysis, one or more studies are excluded at a time and the impact of removing each or several studies is evaluated on the summary results and the between-study heterogeneity. Sequential and combinatorial algorithms are usually implemented to evaluate the change in between-study heterogeneity as one or more studies are excluded from the calculations.[66] Moreover, a meta-regression model can explain heterogeneity based on study-level covariates.[67]

3.10 | Assess publication bias

A funnel plot is a scatterplot that is commonly used to assess publication bias. In a funnel plot, the x-axis indicates the study effect and the y-axis indicates the study precision, such as the standard error or sample size.[68,69] If there is no publication bias, the plot will have a symmetrical inverted funnel; conversely, asymmetry indicates the possibility of publication bias.

3.11 | Present results

A forest plot is a valid and useful tool for summarizing the results of a meta-analysis. In a forest plot, the results from each individual study are shown as a blob or square; the confidence interval, usually representing 95% confidence, is shown as a horizontal line that passes through the square; and the summary effect is shown as a diamond.[70]

4 | PRINCIPLES OF META-ANALYSIS PERFORMANCE

There are four most important principles of meta-analysis performance that should be emphasized. First, the search scope of meta-

analysis should be expanded as much as possible to contain all relevant research, and it is important to remove language restrictions and actively search for non-English bibliographic databases. Second, any meta-analysis should include studies selected based on strict criteria established in advance. Third, appropriate tools must be selected to evaluate the quality of evidence according to different types of primary studies. Fourth, the most suitable statistical model should be chosen for the meta-analysis and a weighted mean estimate of the effect size should be calculated. Finally, the possible causes of heterogeneity should be identified and publication bias in the meta-analysis must be assessed.

5 | STRENGTHS OF META-ANALYSIS

Meta-analyses have several strengths. First, a major advantage is their ability to improve the precision of effect estimates with considerably increased statistical power, which is particularly important when the power of the primary study is limited as a result of the small sample size. Second, a meta-analysis has more power to detect small but clinically significant effects and to examine the effectiveness of interventions in demographic or clinical subgroups of participants, which can help researchers identify beneficial (or harmful) effects in specific groups of patients.[71,72] Third, meta-analyses can be used to analyze rare outcomes and outcomes that individual studies were not designed to test (e.g. adverse events). Fourth, meta-analyses can be used to examine heterogeneity in study results and explore possible sources in case this heterogeneity would lead to bias from "mixing apples and oranges".[73] Furthermore, meta-analyses can compare the effectiveness of various interventions, supplement the existing evidence, and then offer a rational and helpful way of addressing a series of practical difficulties that plague healthcare providers and researchers. Lastly, meta-analyses may resolve disputes caused by apparently conflicting studies, determine whether new studies are necessary for further investigation and generate new hypotheses for future studies.[7,74]

6 | LIMITATIONS OF META-ANALYSIS

6.1 | Missing related research

The primary limitation of a meta-analysis is missing related research. Even in the ideal case in which all relevant studies are available, a faulty search strategy can miss some of these studies. Small differences in search strategies can produce large differences in the set of studies found.[75] When searching databases, relevant research can be missed as a result of the omission of keywords. The search engine (e.g. PubMed, Google) may also affect the type and number of studies that are found.[76] Moreover, it may be impossible to identify all relevant evidence if the search scope is limited to one or two databases.[51,77] Finally, language restrictions and the failure to search non-English bibliographic databases may also lead to an incomplete

meta-analysis.[52] Comprehensive search strategies for different databases and languages might help solve this issue.

6.2 | Publication bias

Publication bias means that positive findings are more likely to be published and then identified through literature searches rather than ambiguous or negative findings.[78] This is an important and key source of bias that is recognized as a potential threat to the validity of results.[79] The real research effect may be exaggerated or even falsely positive if only published articles are included.[80] For example, based on studies registered with the US Food and Drug Administration, Turner et al.[81] reviewed 74 trials of 12 antidepressants to assess publication bias and its influence on apparent efficacy. It was found that antidepressant studies with favorable outcomes were 16 times more likely to be published than those with unfavorable outcomes, and the apparent efficacy of antidepressants increased between 11% and 69% when the non-published studies were not included in the analysis.[81] Moreover, failing to identify and include non-English language studies may also increase publication bias.[82] Therefore, all relevant studies should be identified to reduce the impact of publication bias on meta-analysis.

6.3 | Selection bias

Because many of the studies identified are not directly related to the subject of the meta-analysis, it is crucial for researchers to select which studies to include based on defined criteria. Failing to evaluate, select or reject relevant studies based on stricter criteria regarding the study quality may also increase the possibility of selection bias. Missing or inappropriate quality assessment tools may lead to the inclusion of low-quality studies. If a meta-analysis includes low-quality studies, its results will be biased and incorrect, which is also called "garbage in, garbage out".[83] Strictly defined criteria for included studies and scoring by at least two researchers might help reduce the possibility of selection bias.[84,85]

6.4 | Unavailability of information

The best-case scenario for meta-analyses is the availability of individual participant data. However, most individual research reports only contain summary results, such as the mean, standard deviation, proportions, relative risk and odds ratio. In addition to the possibility of reporting errors, the lack of information can severely limit the types of analyses and conclusions that can be achieved in a meta-analysis. For example, the unavailability of information from individual studies may preclude the comparison of effects in predetermined subgroups of participants. Therefore, if feasible, the researchers could contact the author of the primary study for individual participant data.

6.5 | Heterogeneity

Although the studies included in a meta-analysis have the same research hypothesis, there is still the potential for several areas of heterogeneity.[86] Heterogeneity may exist in various parts of the studies' design and conduct, including participant selection, interventions/exposures or outcomes studied, data collection, data analyses and selective reporting of results.[87] Although the difference of the results can be overcome by assessing the heterogeneity of the studies and performing subgroup analyses,[88] the results of the meta-analysis may become meaningless and even may obscure the real effect if the selected studies are too heterogeneous to be comparable. For example, Nicolucci et al.[89] conducted a review of 150 published randomized trials on the treatment of lung cancer. Their review showed serious methodological drawbacks and concluded that heterogeneity made the meta-analysis of existing trials unlikely to be constructive.[89] Therefore, combining the data in meta-analysis for studies with large heterogeneity is not recommended.

6.6 | Misleading funnel plot

Funnel plots are appealing because they are a simple technique used to investigate the possibility of publication bias. However, their objective is to detect a complex effect, which can be misleading. For example, the lack of symmetry in a funnel plot can also be caused by heterogeneity.[90] Another problem with funnel plots is the difficulty of interpreting them when few studies are included. Readers may also be misled by the choice of axes or the outcome measure.[91] Therefore, in the absence of a consensus on how the plot should be constructed, asymmetrical funnel plots should be interpreted cautiously.[91]

6.7 | Inevitable subjectivity

Researchers must make numerous judgments when performing meta-analyses,[92] which inevitably introduces considerable subjectivity into the meta-analysis review process. For example, there is often a certain amount of subjectivity when deciding how similar studies should be before it is appropriate to combine them. To minimize subjectivity, at least two researchers should jointly conduct a meta-analysis and reach a consensus.

7 | SUMMARY

The explosion of medical information and differences between individual studies make it almost impossible for healthcare providers to make the best clinical decisions. Meta-analyses, which summarize all eligible evidence and quantitatively synthesize individual results on a specific clinical question, have become the best available evidence for informing clinical practice and are increasingly important in medical research. This article has described the basic concept, common methods, principles, steps, strengths and limitations of meta-analyses to help clinicians and investigators better understand meta-analyses and make clinical decisions based on the best evidence.

AUTHOR CONTRIBUTIONS

CM designed and directed the study. XMW and XRZ had primary responsibility for drafting the manuscript. CM, ZHL, WFZ and PY provided insightful discussions and suggestions. All authors critically reviewed the manuscript for important intellectual content.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article because no datasets were generated or analyzed during the current study.

REFERENCES

1. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*. 2000; 283:2008-2012.
2. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine – new approaches and challenges. *Acta Informatica Medica*. 2008;16:219-225.
3. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*. 1997;126: 376-380.
4. Glass GV. Primary, secondary, and meta-analysis of research. *Educational Researcher*. 1976;5:3-8.
5. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ (Clinical Research Ed)*. 1997;315:1533-1537.
6. Chalmers TC, Matta RJ, Smith H Jr, Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med*. 1977;297:1091-1096.
7. Haidich A-B. Meta-analysis in medical research. *Hippokratia*. 2010;14 (Suppl 1):29-37.
8. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-based medicine working group. *JAMA*. 1995;274:1800-1804.
9. Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. *BMC Med Res Methodol*. 2005; 5:14
10. Kovalchik SA. Survey finds that most meta-analysts do not attempt to collect individual patient data. *J Clin Epidemiol*. 2012;65: 1296-1299.
11. Rathi V, Dzara K, Gross CP, et al. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ (Clinical Research Ed)*. 2012; 345:e7570.

12. Petushek EJ, Sugimoto D, Stoolmiller M, Smith G, Myer GD. Evidence-based best-practice guidelines for preventing anterior cruciate ligament injuries in young female athletes: a systematic review and meta-analysis. *Am J Sports Med*. 2019;47:1744-1753.

13. Schmid CH, Stark PC, Berlin JA, Landais P, Lau J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. *J Clin Epidemiol*. 2004;57: 683-697.

14. Higgins JP, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions. Chichester, West Sussex, UK: John Wiley & Sons; 2019.

15. Thomas D, Radji S, Benedetti A. Systematic review of methods for individual patient data meta- analysis with binary outcomes. *BMC Med Res Methodol*. 2014;14:79.

16. Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature vs individual patient data. Paternal cell immunization for recurrent miscarriage. *JAMA*. 1995;274:830-836.

17. McCormack K, Grant A, Scott N. Value of updating a systematic review in surgery using individual patient data. *Br J Surg*. 2004;91: 495-499.

18. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ (Clinical Research Ed)*. 2010;340:c221.

19. Debray TP, Moons KG, Abo-Zaid GM, Koffijberg H, Riley RD. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? *PLoS ONE*. 2013;8(4):e60650.

20. Cheng LL, Ju K, Cai RL, Xu C. The use of one-stage meta-analytic method based on individual participant data for binary adverse events under the rule of three: a simulation study. *PeerJ*. 2019;7: e6295.

21. Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. *Res Synth Methods*. 2010;1:2-19.

22. Grams ME, Surapaneni A, Ballew SH, et al. APOL1 kidney risk variants and cardiovascular disease: an individual participant data meta-analysis. *J Ame Soc Nephr: JASN*. 2019;30:2027-2036.

23. Thomas D, Platt R, Benedetti A. A comparison of analytic approaches for individual patient data meta-analyses with binary outcomes. *BMC Med Res Methodol*. 2017;17(1):28.

24. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med*. 2008;27:1870-1893.

25. Leimu R, Koricheva J. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *Proceedings Bio Scie*. 2004;271:1961-1966.

26. Feng H, Zhao Y, Jing T, et al. Traditional and cumulative meta-analysis: Chemoradiotherapy followed by surgery versus surgery alone for resectable esophageal carcinoma. *Molecular and Clin Onc*. 2018;8:342-351.

27. Maheshwari A, Pandey S, Amalraj Raja E, Shetty A, Hamilton M, Bhattacharya S. Is frozen embryo transfer better for mothers and babies? Can cumulative meta-analysis provide a definitive answer? *Hum Reprod Update*. 2018;24:35-58.

28. Siontis GCM, Nikolakopoulou A, Efthimiou O, Räber L, Windecker S, Jüni P. Evaluation of cumulative meta-analysis of rare events as a tool for clinical trials safety monitoring. *JAMA Netw Open*. 2020;3(9): e2015031.

29. Kaderli RM, Spanjol M, Kollár A, et al. Therapeutic options for neuroendocrine tumors: a systematic review and network meta-analysis. *JAMA Oncol*. 2019;5:480-489.

30. Mayo-Wilson E, Dias S, Mavranezouli I, et al. Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry*. 2014;1: 368-376.

31. Chaimani A, Caldwell DM, Li T, Higgins JP, Salanti G. Undertaking network meta-analyses. *Cochrane Hand for Syste Revi Interve*. 2019; 285-320.

32. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med*. 2011;9:79.

33. Rouse B, Chaimani A, Li T. Network meta-analysis: an introduction for clinicians. *Intern Emerg Med*. 2017;12:103-111.

34. Dias S, Caldwell DM. Network meta-analysis explained. *Arch Dis Child Fetal Neonatal Ed*. 2019;104:F8-F12.

35. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ (Clinical Research Ed)*. 2005;331:897-900.

36. Yuan J, Zhang R, Yang Z, et al. Comparative effectiveness and safety of oral phosphodiesterase type 5 inhibitors for erectile dysfunction: a systematic review and network meta-analysis. *Eur Urol*. 2013;63: 902-912.

37. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods*. 2012;3:80-97.

38. Thompson M, van den Bruel A. *Diagnostic Tests Toolkit*. Vol. 13. Chichester, UK: John Wiley & Sons; 2011.

39. Bachmann LM, Puhan MA, ter Riet G, Bossuyt PM. Sample sizes of studies on diagnostic accuracy: literature survey. *BMJ (Clinical Research Ed)*. 2006;332:1127-1129.

40. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health*. 2015;18:103-109.

41. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA*. 2018;319: 388-396.

42. McInnes MD, Bossuyt PM. Pitfalls of systematic reviews and meta-analyses in imaging research. *Radiology*. 2015;277:13-21.

43. Tansarli GS, Chapin KC. Diagnostic test accuracy of the BioFire® FilmArray® meningitis/encephalitis panel: a systematic review and meta-analysis. *Clinical Microbio Infection: Official Publi Euro So Clinical Microbio Infect Dise*. 2020;26:281-290.

44. Thomas J, Kneale D, McKenzie JE, Brennan SE, Bhaumik S. Determining the scope of the review and the questions it will address. *Cochrane Handbo Syst Rev Inter*. 2019;13-31.

45. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:16.

46. Bettany-Saltikov J. *How to do a systematic literature review in nursing: a step-by-step guide*. UK: McGraw-Hill Education; 2012.

47. Cleyle S, Booth A. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*. 2006;24:355-368.

48. Muka T, Glisic M, Milic J, et al. A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *Eur J Epidemiol*. 2020;35:49-60.

49. Forero DA, Lopez-Leon S, González-Giraldo Y, Bagos PG. Ten simple rules for carrying out and writing meta-analyses. *PLoS Comput Biol*. 2019;15(5):e1006922.

50. Leenaars M, Hooijmans CR, van Veggel N, et al. A step-by-step guide to systematically identify all relevant animal studies. *Lab Anim*. 2012; 46:24-31.

51. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *FASEB Journal: Official Publi Fed Ame Soci Experi Bio*. 2008;22:338-342.

52. 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.

53. Basu A. How to conduct meta-analysis: a basic tutorial. *PeerJ Preprints*. 2017:e2978v2971-e2978v2971.

54. Tawfik GM, Dila KAS, Mohamed MYF, et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tro Med and Health*. 2019;47:46.

55. Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research Ed)*. 2011;343:d5928.

56. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25:603-605.

57. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529-536.

58. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. *Stat Med*. 1991;10:1665-1677.

59. Melsen WG, Bootsma MC, Rovers MM, Bonten MJ. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. *Clinical Microbiology and Infection: The Official Publication of the European Soc Clini Microbio Infec Dis*. 2014;20:123-129.

60. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed)*. 2003;327:557-560.

61. Bowden J, Tierney JF, Copas AJ, Burdett S. Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. *BMC Med Res Methodol*. 2011;11:41.

62. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med*. 1987;316:450-455.

63. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;1:97-111.

64. Zintzaras E, Lau J. Synthesis of genetic association studies for pertinent gene-disease associations requires appropriate methodological and statistical approaches. *J Clin Epidemiol*. 2008;61:634-645.

65. Glasziou PP, Sanders SL. Investigating causes of heterogeneity in systematic reviews. *Stat Med*. 2002;21:1503-1511.

66. Patsopoulos NA, Evangelou E, Ioannidis JP. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. *Int J Epidemiol*. 2008;37:1148-1157.

67. Morton SC, Adams JL, Suttorp MJ, Shekelle PG, AHRQ Technical Reviews. *Meta-regression Approaches: What, Why, When, and How?*. Rockville, MD: Agency for Healthcare Research and Quality (US); 2004.

68. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol*. 2001;54:1046-1055.

69. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ (Clinical Research Ed)*. 2006;333:597-600.

70. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. *BMJ (Clinical Research Ed)*. 2001;322:1479-1480.

71. Jackson D, Turner R. Power analysis for random-effects meta-analysis. *Res Synth Methods*. 2017;8:290-302.

72. Liu P, Ioannidis JPA, Ross JS, et al. Age-treatment subgroup analyses in Cochrane intervention reviews: a meta-epidemiological study. *BMC Med*. 2019;17(1):188.

73. Salanti G, Sanderson S, Higgins JP. Obstacles and opportunities in meta-analysis of genetic association studies. *Gen Med: Official J Amer Col Medical Gen*. 2005;7:13-20.

74. Lee YH. An overview of meta-analysis for clinicians. *Korean J Intern Med*. 2018;33:277-283.

75. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ (Clinical Research Ed)*. 1994;309:1286-1291.

76. Steinbrook R. Searching for the right search--reaching the medical literature. *N Engl J Med*. 2006;354:4-7.

77. Lemeshow AR, Blum RE, Berlin JA, Stoto MA, Colditz GA. Searching one or two databases was insufficient for meta-analysis of observational studies. *J Clin Epidemiol*. 2005;58:867-873.

78. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol*. 2000;53:207-216.

79. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS ONE*. 2013;8(7):e66844.

80. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann N Y Acad Sci*. 1993;703:135-146.

81. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358:252-260.

82. Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol*. 2002;31:115-123.

83. Sharpe D. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away. *Clin Psychol Rev*. 1997;17:881-901.

84. De Luca G, Suryapranata H, Stone GW, et al. Coronary stenting versus balloon angioplasty for acute myocardial infarction: a meta-regression analysis of randomized trials. *Int J Cardiol*. 2008;126:37-44.

85. Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths and limitations. *Cleve Clin J Med*. 2008;75:431-439.

86. Dreier M. *Quality Assessment in Meta-analysis. Methods of Clinical Epidemiology*. Berlin, Heidelberg: Springer Series on Epidemiology and Public Health; 2013.

87. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539-1558.

88. Bailar JC 3rd. The promise and problems of meta-analysis. *N Engl J Med*. 1997;337:559-561.

89. Nicolucci A, Grilli R, Alexanian AA, Apolone G, Torri V, Liberati A. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. *JAMA*. 1989;262:2101-2107.

90. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med*. 2003;22:2113-2126.

91. Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol*. 2000;53:477-484.

92. Garg AX, Hackam D, Tonelli M. Systematic review and meta-analysis: when one study is just not enough. *Clin J Amer Soc Nephrology: CJASN*. 2008;3:253-260.