Guide to Statistics and Methods

Practical Guide to Meta-analysis

Shipra Arya, MD, SM; Todd A. Schwartz, DrPH; Amir A. Ghaferi, MD, MS

Meta-analysis is a systematic approach of synthesizing, combining, and analyzing data from multiple studies (randomized clinical trials[1] or observational studies[2]) into a single effect estimate to answer a research question. Meta-analysis is especially useful if there is debate around the research question in the literature published to date or the individual published studies are underpowered. Vital to a high-quality meta-analysis is a comprehensive literature search, prespecified hypothesis and aims, reporting of study quality, consideration of heterogeneity and examination of bias. In the hierarchy of evidence, meta-analysis appears above observational studies and randomized clinical trials because it rigorously collates evidence across a larger body of literature; however, meta-analysis is largely dependent on the quality of the primary data.

Related articles pages 432, 434, and 436

Use of Methods

Meta-analysis (or the analysis of analyses) is a subset of systematic reviews (ie, reviews that follow a prespecified protocol, including eligibility criteria, research question and methods to collect evidence from multiple sources in published literature).[3] While systematic reviews can clarify the nature and causes of disagreement and identify areas needing more research, meta-analysis is particularly useful compared with systematic review in which clinical trials provide conflicting results or are underpowered to address a research question but have adequate data with less heterogeneity. Meta-analysis can establish associations between effect estimates and study-level variables (meta-regression) and explore outcomes in subsets of patients.

Conducting meta-analysis includes 5 standard specific steps (**Box**).[3] First, a research question must be established a priori, following the Population, Intervention/exposure, Comparison, Outcomes, Study designs model. Second, meta-analysis must use a prespecified comprehensive search strategy. The goal is to include all relevant evidence to date and to avoid publication bias. Key words are specified and noted for reproducibility using Medical Subject Heading (MeSH) terms, and a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram should be used to guide search results by 2 independent investigators. Relevant studies are identified from multiple electronic data sources (MEDLINE, PubMed, Embase, Cochrane Central Register of Controlled Trials, Web of Science, and Google Scholar) and augmented by searching reference lists, abstracts for meetings, conference proceedings, books, translating non-English references (if possible), and contacting experts. Third, titles, abstracts, and complete manuscripts must be screened based on prespecified inclusion and exclusion criteria; individual studies must be assessed for quality and scored to evaluate quality of included studies (eg, Jadad Score for randomized clinical trials and Newcastle-Ottawa Scale for nonrandomized studies[4]), and sources of heterogeneity must be assessed. Data must be extracted, determining meta-analysis eligibility, population variables, hierarchy of study designs, methods quality, interventions or exposure, outcomes, analytical techniques, and results; independent assessment must be performed by 2 data abstractors (report interrater agreement) with differences resolved by discussion or expert opinion. Fourth, data must be analyzed, synthesizing summary results to generate an effect estimate and corresponding confidence interval. The quality of a meta-analysis is only as high as the included individual studies because meta-analysis does not remove sources of bias from individual studies. Summary estimates may be quite precise but biased. Fifth, it is important to evaluate for sources of heterogeneity, report publication bias estimates, and perform sensitivity analyses to assess the robustness of findings (eg, omitting low-quality studies, and trimming studies with extreme findings).

The most common measures of effect used for categorical response variables are the risk ratio (or relative risk) and the odds ratio. Continuous response variables are usually synthesized using the standardized mean difference. All methods allow for the weighting of studies so that the evidence of a specific study is reflected in the summary estimate. The most common method of weighting is to use the inverse of the variance which incorporates information on sample size and variance. Standard statistical software packages are available such as SAS (SAS Institute Inc), Stata (StataCorp LLC), and R (The R Project for Statistical Computing), as well as specialized software (eg, Comprehensive Meta-Analysis [Biostat Inc], RevMan [The Cochrane Collaboration], MetAnalysis [computer software], MetaWin, MIX [Excel], and WEasyMA [Clininfo]).[5]

Statistical Considerations

There are several important considerations when conducting meta-analysis. The statistical considertions include modeling, heterogeneity, and publication bias.

Random-Effects vs Fixed-Effects Modeling

There are 2 approaches to modeling in meta-analysis—random effects and fixed effects—depending on how different the underlying studies are and the assumptions made while synthesizing the data. In a random-effects approach, the assumption is that there is no single true effect estimate, but rather an underlying distribution of effects. Variability is thought to arise from both within and between studies. This is a more conservative approach than others and could be considered a starting point in meta-analysis because it provides effect size variation across studies. A commonly used method in random-effects modeling is the DerSimonian and Laird[6] method. In a fixed-effects approach, the assumption is that all studies included are estimating the same underlying true effect. Variability (error) is assumed to originate from within each study, not between the studies. Commonly used methods for fixed-effects modeling are the Mantel-Haenszel method and the Peto method.

Box. Best Practices for Meta-analysis

1. State research question and study protocol/design a priori.
2. Perform reproducible comprehensive literature search.
3. Abstract study data using two independent abstractors and collect detailed data on study quality, sources of heterogeneity.
4. Analyze effect estimate and consider random-effects model as a starting point.
5. Perform sensitivity analyses (eg, on high quality studies) and explore sources of heterogeneity. Report publication bias.

Reporting Heterogeneity Among Studies

Heterogeneity is the presence of variation in the effect sizes of underlying studies. It is an important concept to measure, report, and include in interpretation of meta-analysis findings. Heterogeneity arises from systematic differences between studies included in the meta-analysis, such as study design or sample characteristics. Common heterogeneity statistics reported are Cochran *Q* statistic (a measure of weighted squared deviations), tau-squared (between-study variance), and the inconsistency index (I^2) (the ratio of true heterogeneity to total observed variation; preferred because estimates consistency of evidence not associated with chance).[7] When any of these measures of heterogeneity is high, summary estimates should be interpreted with caution. Additional investigation of the sources of heterogeneity should be undertaken by performing sensitivity analyses on high-quality studies.[8]

Publication Bias

Negative studies are less likely to be submitted or published; hence publication bias can occur suggesting beneficial or large effect estimates are published. Popular methods of assessing publication bias include funnel plots, Begg rank correlation test, and Egger test.[2]

Where to Find More Information

Several textbooks on systematic review and meta-analysis methods provide further detail. The Cochrane collaboration (https://www.cochranelibrary.com), PRISMA[4] (http://prisma-statement.org), meta-analysis of observational studies in epidemiology (MOOSE) Guidelines,[2] the Equator network (http://www.equator-network.org), and PROSPERO registration (https://www.crd.york.ac.uk/PROSPERO/) provide guidance for conducting a systematic review or meta-analysis. Additionally, published studies or protocols offer examples of meta-analysis.[9,10]

ARTICLE INFORMATION

Author Affiliations: Stanford-Surgery Policy Improvement, Research and Education (S-SPIRE) Center, Palo Alto, California (Arya); Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill (Schwartz); Department of Surgery, University of Michigan, Ann Arbor (Ghaferi).

Corresponding Author: Shipra Arya, MD, SM, Stanford-Surgery Policy Improvement, Research and Education (S-SPIRE) Center, 1070 Arastradero Rd, Palo Alto, CA 94306 (sarya1@stanford.edu).

Published Online: January 29, 2020. doi:10.1001/jamasurg.2019.4523

Conflict of Interest Disclosures: Dr Ghaferi reported grants from Agency for Healthcare research and Quality, grants from the Patient-Centered Outcomes Research Institute, grants from Blue Cross Blue Shield of Michigan, and grants from National Institutes of Health outside the submitted work. No other disclosures were reported.

REFERENCES

1. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535. doi:10.1136/bmj.b2535

2. 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(15):2008-2012. doi:10.1001/jama.283.15.2008

3. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2*. www.cochrane-handbook.org. Updated September 2009. Accessed December 15, 2019.

4. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality; 2012.

5. Bax L, Yu L-M, Ikeda N, Moons KGM. A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC Med Res Methodol*. 2007;7:40-40. doi:10.1186/1471-2288-7-40

6. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2

7. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Identifying and quantifying heterogeneity. In: *Introduction to Meta-Analysis*. Hoboken, NJ: John Wiley & Sons Ltd; 2009.

8. Gagnier JJ, Morgenstern H, Altman DG, et al; Ann Arbor Clinical Heterogeneity Consensus Group. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. *BMC Med Res Methodol*. 2013;13:106. doi:10.1186/1471-2288-13-106

9. Arya S, Pipinos II, Garg N, Johanning J, Lynch TG, Longo GM. Carotid endarterectomy is superior to carotid angioplasty and stenting for perioperative and long-term results. *Vasc Endovascular Surg*. 2011;45(6):490-498. doi:10.1177/1538574411407083

10. Sahebally SM, McKevitt K, Stephens I, et al. Negative pressure wound therapy for closed laparotomy incisions in general and colorectal surgery: a systematic review and meta-analysis. *JAMA Surg*. 2018;153(11):e183467. doi:10.1001/jamasurg.2018.3467