Teaching Meta-Analysis: Concepts, Controversies, and Resources

Dawson, DV[1,2], Blanchette, DR[1]
[1]Division of Biostatistics and Research Design, College of Dentistry, University of Iowa, 801 Newton Rd, Iowa City, IA 52246
[2] Department of Biostatistics, College of Public Health, University of Iowa, 145 N. Riverside Drive 100 CPHB, Iowa City, IA 52242

Abstract
Countless meta-analyses are published annually, representing a substantial subset of the literature consumed by researchers, health professionals and, especially, by trainees in the health and life sciences. Such publications represent potentially useful tools for both the researcher and the busy practitioner, provided that the reader is able to understand their implications – and their limitations. Approaches to the presentation of meta-analysis to students in the health and life sciences are outlined, emphasizing conceptual understanding and interpretation, as well as an appreciation for the more controversial aspects of meta-analysis. Interactive classroom approaches are discussed, and resources to enhance consideration of this important topic are shared. These include bibliographies, helpful web sites, consideration of graphics, software, and freely available instructional materials. Developing approaches to the systematic evaluation of the quality of published meta-analyses are described. Finally, some of the more recent innovations within meta-analytic methodology are referenced.

Key Words: Meta-analysis, statistical education, health sciences, teaching resources

1. Introduction

The term *meta-analysis* was coined by Glass in 1976 to refer to "the statistical analysis of a large collection of analysis results from individual studies for purposes of integrating the findings" (Glass, 1976). Since then, meta-analysis has come to be considered an important tool for assembling, synthesizing, and evaluating the evidence critical to the pursuit of evidence-based clinical practice, and has found wide usage. Given the proliferation of meta-analyses in the biomedical literature, where they are regularly accessed by trainees and practitioners in the health sciences, we submit that it is important that such individuals be able to understand and critically assess meta-analyses. In this paper, we will discuss the motivation and utility of meta-analysis, and the key didactic elements that we believe should be communicated to trainees and practitioners in the health sciences. We will review current efforts to formalize the evaluation of meta-analyses. We will also address several controversies associated with the use of meta-analytic procedures. For each of these considerations, pedagogical approaches and exercises that may be used in a formal teaching environment will be suggested, and

resources for those wishing to impart a basic understanding of meta-analysis to non-statisticians will be described.

2. Motivation and Key Didactic Elements

2.1 What can Meta-Analysis Do for Us?

Meta-analyses are intended to synthesize results from individual studies, most commonly from all of the suitable extant literature addressing a particular scientific question. Meta-analyses are intended to quantitatively synthesize the results of multiple studies in an objectively verifiable manner, using statistical methods to combine and summarize the results of these component studies. Meta-analysis increases the statistical power of primary end points – including those for subgroups – due to an increased sample size, and results in increased precision of risk estimates. Meta-analysis can be applied to clinical or intervention trials, or to observational studies, such as case-control and cohort studies used in epidemiologic investigations, or studies of genetic associations with disease. Meta-analytic approaches are able to assess and describe apparent conflicts in the literature, and may be able to resolve uncertainty when reports disagree. Meta-analysis therefore has the potential to assist the field in reaching conclusions about treatment, as well as to assist in the exploration of inconsistent results, of therapeutic efficacy in subgroups, and of effect modification (statistical interaction). They may help determine harm as well as efficacy, by furthering the study of adverse events, and may assist with exploration of covariate effects, with identification of gaps in the literature, and with the planning of future scientific investigations.

2.2 Methodological Considerations

In order to be able to interpret and evaluate a meta-analysis, the health care professional should have some appreciation for the procedures involved in the conduct of a meta-analysis, in order to properly understand and interpret the results, and to have the means of assessing the quality and value of a particular meta-analysis. We note that while meta-analysis can technically be used to summarize any assemblage of studies, meta-analysis is most often intended as a summary of the extant literature pertinent to a particular scientific question. In this context, the protocol by which the individual studies to be subjected to meta-analysis are identified is critical and should be assembled through the careful methodology of a systematic review. The consumer of meta-analyses should therefore be conversant with the procedures of systematic reviews, with protocols related to data abstraction for the meta-analysis, with the statistical methodology of meta-analysis, including graphics, and with considerations related to quality assessment.

2.2.1 Protocol development

When, as is usually the case, the meta-analysis is intended as a summary of the extant literature pertinent to a particular scientific question, a detailed research protocol should be developed *a priori*. It should carefully define the scientific question, the endpoints of interest, and a procedure for a comprehensive search to identify relevant studies, with specific inclusion and exclusion criteria. The protocol should be so clearly stated as to be readily reproducible by a third party. The aim is to comprehensively search and represent the entire body of pertinent evidence via a *systematic review*, defined as “a high-level overview of primary research on a particular research question that tries to identify, select, synthesize and appraise all high quality research evidence relevant to that question in order to answer it” (Cochrane, 1972). Further details on the methodology of systematic

reviews are found in the overview of Needleman (2002) and in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green, Version 5.1.0, updated March 2011). This handbook may be found online through the Cochrane Community web site of the Cochrane Collaboration at *http://community.cochrane.org/handbook* or at *http://handbook.cochrane.org/*. The goal of these rigorous methods for appraisal of the literature is to limit bias in order to improve the reliability and accuracy of conclusions. For this reason, identification of relevant studies in the so-called "grey literature" outside mainstream published journals and monographs is frequently attempted, including searches involving dissertations, conference abstracts or proceedings, and private sector or governmental research.

Another aspect of protocol construction is the development of the procedures and instruments for the data abstraction process. Abstraction of key data such as the study-specific characteristics, sample sizes, effect estimates, and information that can be used to obtain the study-specific standard errors should ideally be done by more than one individual, and the results of replicate abstractions compared. Protocols should be in place for the conduct and evaluation of this comparison, and the resolution of discrepancies. It is also helpful to recognize that the nature and format of the information as it is presented in the individual publication will affect the ability to include a particular study in the planned meta-analysis. The simple exercise of having individuals in the class fill out a provided abstraction form for a few judiciously chosen papers from the literature can bring home this message. This exercise can lead naturally to a discussion of the impact of reporting; helpful information to support such a discussion may be found in the reporting guidelines, including basic statistical reporting, for various types of studies found online, supplied by the Cochrane Collaboration at *http://www.cochrane.org/about-us/evidence-based-health-care/webliography/books/reporting*. It may also be useful for students to become acquainted with the CONSORT (Consolidated Standards of Reporting Trials) guidelines which were developed to alleviate the problems arising from inadequate reporting of randomized controlled trials: http://www.consort-statement.org/.

2.2.2 Analytic considerations

While a detailed understanding of statistical techniques used in meta-analysis is not essential for the consumer of the biomedical literature to read a meta-analysis profitably, familiarity with a few concepts should be helpful. These include an understanding of the variety of effect sizes that may be the focus of a particular meta-analysis (e.g., differences between mean treatment response, odds ratio as measures of association, and so forth), and the interpretation of the combined effect size and its associated (typically 95%) confidence interval derived from the meta-analysis. Another important analytic aspect is the assessment of homogeneity among the component studies. Considerations of homogeneity vs. heterogeneity have implications for the specific analytic model used, i.e., *fixed effect* vs. *random effect* meta-analytic models, and the interpretation of results (Borenstein et al, 2009; Sutton et al, 2000; DerSimonian & Laird, 1986). Graphical display of both the results of individual studies and the summary from the meta-analysis (i.e., the combined effect estimate and its associated confidence interval) is accomplished via a *forest plot (http://www.cebi.ox.ac.uk/for-practitioners/what-is-good-evidence/how-to-read-a-forest-plot.html).*

Meta-analyses seeking to summarize the existing evidence are particularly concerned with limiting bias, and typically include investigations of the possibility of publication bias, both through formal testing and graphical displays such as funnel plots (Begg et al, 1994; Egger et al, 1997; Macaskill et al, 2001; Tang et al, 2000; Ioannidis et al, 2007).

For a general discussion of the problem of publication bias, see Dwan et al (2013) and Kicinski et al (2015). Other potential sources of bias are often explored in the context of a meta-analysis: Dwan et al (2013) also discuss outcome reporting bias, and Higgins et al (2011) discuss various sources of bias in randomized trials, describing a Cochrane Collaboration tool for assessing risk of bias.
Depending on time constraints, the instructor may wish to discuss at least the existence of such techniques as evaluation of covariate effects, trends, and dose-response, as well as the possibility of exploring and formally assessing differences in subgroups of the component studies.

Useful references for the instructor include Borenstein et al (2009) and Sutton et al (2000). A useful web site providing basic definitions and illustrations is supported by the University of Oxford Centre for Evidence Based Interventon (CEBI): http://www.cebi.ox.ac.uk/for-practitioners/what-is-good-evidence.html

The Cochrane Collaboration includes detailed information on the conduct of meta-analysis, including the previously mentioned handbook. The Cochrane Training site provides additional links to other resources, including Spanish language versions, a handbook for systematic reviews of diagnostic test accuracy, summaries, glossaries, and various updates. These can be accessed via the website:
http://community.cochrane.org/about-us/evidence-based-health-care/webliography/books/sysrev.

Also available through the Cochrane Collaboration is the downloadable REVMAN (Review Manager) software used for preparing and maintaining Cochrane reviews and performing simple-meta-analyses and associated graphical representations of the results: http://tech.cochrane.org/revman. Other proprietary software packages are also available.

The Cochrane site also provides access to a vast number of systematic reviews and meta-analyses completed using the rubrics of the Collaboration. These can be used for illustration and as a basis for classroom exercises. Periodic updates help convey the ongoing nature of scientific investigation, the development of our understanding, and the accumulation of evidence to the students. These meta-analyses are sometimes quite formidable and lengthy, and a number of strategies have been found to be helpful in guiding students through understanding, interpretation and critical review of meta-analyses. One strategy is to have group presentations of meta-analyses, or of the several aspects of a single meta-analysis. Providing an outline of elements for inclusion with questions for consideration is often helpful. When a particular analysis is particularly extensive, selection of a few key tables or graphics for dissection and discussion can assist the students in working through the meta-analysis. Finally, summaries of published meta-analyses found in scientific journals may also support the learning of students new to meta-analysis. For example, Niederman (2003) provided a summary of the Cochrane review and meta-analysis comparing the performance of manual versus powered toothbrushes, summarizing and critically reviewing the original Cochrane report (Heanue et al, 2003).

3. Evaluation of Meta-Analysis

In some respects, critical review of a meta-analysis is no different from that of any other publication. The reader should consider questions such as the following:

Was the research question well defined? Was the primary outcome and the populations selected for study well chosen? Were the statistical methods appropriately chosen and were the analyses properly carried out? Were the interpretations of the findings appropriate, i.e., did the authors draw valid conclusions based upon the analysis of the data?

There are however, other aspects particularly germane to meta-analysis that should also be critically evaluated. These include the protocol for the selection of the individual component studies, which should be detailed, well-reasoned and reproducible. If the intent is to represent all available literature related to the question under consideration, there should be evidence that the authors of the meta-analysis attempted to avoid publication bias. Such efforts could include a thorough, carefully specified and systematic search protocol, consideration of the grey literature, and attempts to contact authors where usable information for purposes of meta-analysis was not present in the publication. Inclusion of reports translated from other languages is also considered a positive effort in the quest to eliminate bias. In addition, the reader should consider the suitability of the procedures for data abstraction and the verification of those abstractions; another consideration is whether formal evaluation of agreement among abstractors was included. Appropriate meta-analytic models and statistical methods, with reasonable assumptions regarding heterogeneity, should have been employed, and appropriate graphical displays should be present. Additional features should include evaluations of potential bias, particularly the possibility of publication bias.

Lastly, we note that the quality of the component studies is an important aspect of the quality of the meta-analysis based upon them. For this reason, journals will often request an evaluation of study characteristics, including such dimensions as possible conflict of interest, handling of missing data, *a priori* considerations of power and sample size requirements, and procedural aspects such as masking (blinding to treatment), randomization, and calibration. The Cochrane tool (Higgins, Altman & Gøtzsche, 2011) is sometimes requested for this evaluation, and may also be used in a classroom setting, if only to indicate to students how many of these critical elements are not explicitly reported or discernible in the literature.

A number of recent initiatives have focused on formal quality assessment and the development of procedures and instruments for such assessment. One example of a formal statement regarding meta-analysis quality is the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement, which includes a 27 item checklist of recommended items for inclusion in the report of a systematic review or meta-analysis (Moher et al, 2009; Liberati et al, 2009). Further details are accessible at http://www.prisma-statment.org. A more recent elaboration focused on protocol aspects, and associated with a shorter checklist, is found in the PRISMA-P guidelines (Shamseer et al, 2015). Other statements, with accompanying checklists for use by editors, reviewers and readers – and

potentially by students – are the QUORUM (Quality of Reporting of Meta-analyses, Moher et al, 1999) statement for meta-analysis of randomized clinical trials, MOOSE (Meta-analysis Of Observational Studies in Epidemiology, Stroup et al, 2000), and AMSTAR (Assessment of Multiple Systematic Reviews, Shea et al, 2007). Application of assessments such as these can be a sobering exercise for students (DeBuitrago et al, 2013).

4. Controversies Associated with Meta-Analysis

Interestingly enough, one controversy centers around exactly what constitutes a meta-analysis. Some practitioners consider a meta-analysis to be a systematic review which includes a quantitative synthesis, as exemplified by this definition by Carr (2002): "When the systematic review includes a statistical synthesis of the numerical results of several trials that examined the same question it is termed a meta-analysis." Others insist that the term meta-analysis should be used exclusively to denote the statistical techniques involved in the quantitative synthesis. While it is true that the statistical techniques associated with meta-analysis can – and have – been applied outside the context of a systematic review, meta-analysis is most typically conducted in association with a systematic review for the assembly of the individual studies to be subjected to analysis. Clearly, when a meta-analysis is used for the purpose of summarizing all information in the literature pertaining to a particular research question, it is essential that it be conducted in the context of a systematic review, and that a well-conceived and well-conducted systematic review constitutes an important design aspect of the meta-analysis.

Another area associated with some controversy is the assessment of publication bias. Approaches used to assess publication bias may be misleading; for example, factors other than publication bias can produce patterns in funnel plots similar to those that would be expected to arise from publication bias. There is a considerable literature addressing these concerns (Begg et al, 1994; Egger et al, 1997; Macaskill et al, 2001; Tang et al, 2000; Ioannidis et al, 2007).

Finally, we observe that meta-analysis has not always been accepted wholeheartedly in all quarters, as illustrated by these quotations from the titles of two journal article titles: ""Meta-analysis/shmeta-analysis . . . " (Shapiro, 1994) and "Statistical alchemy for the 21st century . . ." (Feinstein, 1995). In particular, the problem of publication bias constitutes a formidable and pervasive challenge (Dwan et al, 2013; Kicinski et al, 2015). Further, it is clearly unreasonable to expect that the synthesis provided by meta-analysis will be able to somehow overcome the shortcomings and biases of the component studies. As eloquently expressed below:

"Meta-analysis cannot transcend the limitations of the data upon which it is based. It can but hold a mirror to the scientific community, summarizing the

conclusions and the quality of the available evidence concerning the substantive questions at issue."

(Shapiro and Shapiro, 1983, page 43)

5. Summary and Conclusions

When carefully applied, meta-analytic approaches have the potential to supply useful syntheses of the results from individual reports. As with any analytic approach, there are assumptions and pitfalls associated with its application, but a well-conducted meta-analysis can integrate this collective information in a valid manner, and can make it possible to resolve conflicting results in the literature, and potentially answer questions not addressable through individual studies in the extant literature. There is considerable value in understanding the state of the existing literature, even when the meta-analysis primarily serves to underscore the gaps and deficiencies in that literature. The ubiquity, widespread citation, and potential utility of meta-analysis all support the need for education regarding meta-analysis among clinicians, biomedical researchers and trainees. It is critical that such individuals have a basic understanding of the methods of meta-analyses, and to be able to interpret and critically assess their results.

Acknowledgements

This work was partially supported by the endowment of the Morris Bernstein Professorship in Dentistry.

References

Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994:1088-101.

Borenstein M, Hedges L, Higgins J, Rothstein H. Introduction to meta-analysis. Chichester: John Wiley & Sons, Ltd; 2009.

Carr AB. Systematic reviews of the literature: the overview and meta-analysis. Dental Clinics of North America 2002;46(1):79-86.

Cochrane AL. Effectiveness and Efficiency: Random Reflections on Health Services. London: Nuffield Provincial Hospitals Trust, 1972.

Cochrane Community. Reporting guidelines: http://www.cochrane.org/about-us/evidence-based-health-care/webliography/books/reporting. Accessed March 19, 2015.

De Buitrago JG, Avila-Ortiz G, Elangovan S. Quality assessment of systematic reviews on alveolar ridge preservation. JADA 2013;144:1349-57.

DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.

Dwan K, Gamble C, Williamson PR, Kirkham JJ; Reporting Bias Group Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013;8(7):e66844. doi: 10.1371/journal.pone. 0066844.)

Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.

Feinstein AR. Meta-analysis: statistical alchemy for the 21st century. J Clin Epidemiol

1995;48(1):71-79.
Glass GV. Primary, secondary, and meta-analysis of research. Educational Researcher 1976;5(1):3-8.
Heanue M, Deacon SA, Deery C, Robinson PG, Walmsley AD, Worthington HV, Shaw WC. Manual versus powered toothbrushing for oral health. The Cochrane Database of Systematic Reviews 2003, Issue 1. Art. No.: CD002281. doi: 10.1002/14651858. CD002281.
Higgins, JPT, Green S. Cochrane handbook for systematic review of interventions, Version 5.1.0. Cochrane Collaboration 2011. http://www.cochrane-handbook.org/
Higgins J, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343: doi: 10.1136/bmj.d5928.
Ioannidis JP, Trinkalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. CMAJ 2007;176(8):1091-6.
Kicinski M, Springate DA, Kontopantelis E. Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews. Stat Med 2015;34(20):2781-93. doi: 10.1002/sim.6525.)
Liberati A, Altman D, Tetzlaff J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7):e1000100.
Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta‐analysis. Stat Med. 2001;20(4):641-54.
Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999;354(9193):1896-900.
Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.
Needleman IG. A guide to systematic reviews. J Clin Periodontol. 2002;29(s3):6-9.
Niederman R. Manual versus powered toothbrushes: The Cochrane review. JADA 2003;134(9):1240-1244.
Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ, 2015;349:g7647.
Shapiro DA, Shapiro D. Comparative therapy outcome research: Methodological implications of meta-analysis. J Consult Clin Psychol. Journal of Consulting and Clinical Psychology 1983;51(1):42-53.
Shapiro S. Meta-analysis/Shmeta-analysis. Am J Epidemiol. 1994;140(9):771-78.
Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Research Methodology 2007;7:10.
Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. 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.
Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for meta-analysis in medical research. Chichester: J. Wiley & sons Ltd. 2000.
Tang J-L, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. J Clin Epidemiol. 2000;53(5):477-84.