Comments on Meta-Analyses in General and in Stem Cell Research: An Overview and Cautionary Advice

Ronald L. Thomas and Graham C. Parker

Meta-analysis, a tool for contrasting and combining results from different research studies, has been around now for over 40 years. Journal editors are eager to publish the results from meta-analyses, as they propose to represent the integration of best research evidence with clinical expertise and patient values. There are guidelines available, most notably through the Cochrane Collaborative, for investigators to follow in conducting a responsible and, therefore publishable, meta-analysis. Despite the burgeoning popularity of this powerful analytical tool, the procedure is not without its pitfalls. In this study, we advise the readership to familiarize themselves with the most common shortcomings in an effort to help elevate our ability to critically appraise the results of these analyses.

META-ANALYSIS, A TOOL for contrasting and combining results from different research studies, has been around now for over 40 years. Meta-analysis, as we know it today, was formulated back in 1974 by Gene Glass [1], a psychologist, who was relentless in repudiating the position that psychotherapy was an ineffectual approach. Its actual gestation although was many years before that and can be dated to the 1930s to research pioneers Fisher [2], Pearson [3], Cochran [4], and Tippet [5], all of whom sought a way to synthesize and combine research data from differing agricultural studies [6].

The first published meta-analysis in medicine is attributed to American Chalmers et al. [7] in 1975. In the years since, meta-analysis has grown enormously from a small group of statistical researchers interested in integrating studies to both an academic institution and commercial enterprise. As a single proof, The Cochrane Collaborative [8] organization lists among its membership 31,000 volunteers, with centers in the United Kingdom, South Asia, Brazil, Australia, and the United States. It has 6 separate review groups, 15 method groups, and formed an official relationship in 2011 with the World Health Organization (WHO) as a partner with a seat on the World Health Assembly to provide input into WHO resolutions.

There are several commercial and free software programs available to conduct meta-analyses. As of this writing, a single keyword search on PUBMED of “meta-analysis” returned 91,664 listings. The same keyword search for the year 2014 returned 15,983 listings. There is no doubt that

hundreds of researchers at universities and commercial industries all over the world, at this very moment, are conducting a meta-analysis. It is easy to see why. Journal editors are eager to publish the results from meta-analyses, as they propose to represent the integration of best research evidence with clinical expertise and patient values. Meta-analyses are, arguably, at the top of the evidence pyramid, as they provide much more evidence than the outcomes of one randomized controlled trial (RCT), cohort study, case report, or an expert opinion. Meta-analysis has found its way into, or will find its way into, every medical, clinical, and biomedical field—including stem cell research [9–12][13, this issue].

Methodology

Meta-analysis is a term which describes going “back” in a systematic fashion over all identified studies to examine a specific aim or hypothesis by pooling data from each study. It proceeds carefully as a “study of studies,” as each aspect of the study conducted is examined for comparability. These include, but certainly are not limited to, comparing identical hypotheses, operational definitions of independent and dependent variables, data collection procedures and abstraction, types of statistical analyses, checks for heterogeneity, publication bias, and construction of plots to represent the results of the meta-analysis.

More formally, similar to individual primary research studies, a meta-analysis is conducted in successive steps or stages. It begins with a specific aim or formulation of a

Carman and Ann Adam Department of Pediatrics, Children’s Research Center of Michigan, Wayne State University School of Medicine, Detroit, Michigan.

testable hypothesis, then moves through identification of relevant studies to include and compare, followed by data collection and abstraction from them, and statistical analysis of the collected data follows (eg, weighted odds ratios, risk ratios, means). The process concludes with a presentation of the results, most famously, as a forest plot [14]. Meta-analyses, done correctly, are laborious undertakings. Within each of these steps or stages along the way, there are additional questions to be asked and addressed (eg, inclusion and exclusion criteria), which can affect the validity of the combined review. There are guidelines available with clear recommendations for investigators to follow in conducting a responsible and, therefore, publishable meta-analysis. The obvious best practices have arisen through the Cochrane Collaborative for systematic and meta-analytic reviews, which standardize the methodology and dissemination of results and PRISMA [15]. As its website states, PRISMA stands for preferred reporting items for systematic reviews and meta-analyses. It is an evidence-based minimum set of 27 items for reporting in systematic reviews and meta-analyses. It began in 1996 to address the suboptimal reporting of meta-analyses. An international group developed a guide called the QUOROM (quality of reporting of meta-analyses) statement, which focused on the reporting of meta-analyses of randomized controlled trials. In 2009, the guideline was updated to address several conceptual and practical advances in the science of systematic reviews, and was thence renamed PRISMA.

Strengths and Weaknesses

Despite the burgeoning popularity and clear recommendations for investigators to follow in conducting a responsible meta-analysis, the procedure is not without its criticisms. The general strengths and weaknesses of meta-analysis have been well documented in the literature, and therefore will not be extensively elaborated on here. Noble [6] listed 16 strengths and 17 weaknesses taken from 11 publications. It is, however, important to indicate that the purported strength of all meta-analyses is predicated on the fact that the measures and constructs, as outlined in the individual studies, can be validly combined to evaluate the strength of evidence. In the perfect sense, all original data from each individual study would be combined and analyzed. This would go a long way in reducing possible mistakes and invalidity of data extraction and coding from outcome variables. Unfortunately, an examination of all raw data from individual studies is rarely accomplished, which becomes the major weakness of a meta-analysis.

Perhaps the least obvious, but clinically relevant shortcoming is the lack of analyses, including moderating and mediating variables. Interaction effects on the primary outcome (main effect) variable are seldom reported between clinically relevant third variables. Such difficulty in isolating and interpreting possible interactions could make interpretation of the results misleading. In an extreme example, the lack of such an examination might preclude detection of a serious adverse event. This represents a dangerous omission or possible misindication. The lack of examination of interaction effects cannot always be corrected satisfactorily in the statistical analyses [16]. In addition, examining observed differences across studies still has the

inherent possibility of being confounded by third variables (even if they were randomized) that may vary between studies. This essentially precludes the detection of potentially important cause-and-effect relationships. To be clear, the failure of such an analysis to detect an interaction is truly a shortcoming rather than a fault. In our example, the adverse event would not be revealed by the individual study, which contributed to the meta-analysis, beyond perhaps an anecdotal statement. Also, surely, the whole point of the meta-analysis is to elevate an evidence-based medicine above the anecdotal?

A second related concern is the interpretation of the quality of different conclusions drawn from the same meta-analysis. There may be a very strong effect on one outcome that is heavily supported by the outcomes of several studies with relatively large numbers of well-controlled contributing data points. How then do we properly interpret negative results, which are actually only determined by less well-controlled aspects, of the same meta-analysis? As it is reported within the same article, the casual reader may well be tempted to give the negative conclusions, in our example, the same weight as the well-supported positive conclusion. In this instance, the negative results of a well-controlled single RCT should be given more weight than the meta-analysis, however well intentioned. Responsible descriptions of the heterogeneity within each study and indeed the number of relevant studies contributing to each reported result within a meta-analysis are crucial to assessing the validity of its conclusions.

The next concern might seem almost a tautology. The data sets that contribute to a given meta-analysis are almost, without exception, published studies. Journals are still resistant to publishing “negative results. How can a meta-analysis report on the results of all relevant studies if the negative results have not been published? More troubling is the question of how many relevant studies have been undertaken and then not released due to the negative nature of the results? For every study we see published with a negative result or indeterminate clinical indication, can we make an estimation of how many more have been undertaken and not reported?

An oft-stated strength of meta-analysis is its ability to make use of studies that were performed or even designed with insufficient power to allow firm conclusions to be drawn from the resulting data set. However, underpowered studies with small sample sizes have the potential to overstate the effect size estimated from the meta-analysis. Also, as previously indicated, there is a major absence in small sample trials of assessing the influence of potential moderating and mediating variables. Clinical trials are largely a balance of benefits and risks. The lack of examination of clinically important patient characteristics (moderators) and/or trial conditions (mediators) certainly weaken the beneficial tradeoff of the trial.

Future Directions and Advice

The Cochrane organization, along with other professional organizations such as PRISMA, has performed at an exemplary level in keeping up with the explosion of published meta-analyses. Their efforts are vital, as the careful interpretation of well-performed meta-analyses and systematic

reviews is a critical boon to the decision-making process made by healthcare professionals and policy makers.

Given the rapid rise of meta-analyses and systematic reviews being conducted, it is more than advisable to the readership to familiarize themselves with the methods and interpretative techniques necessary to critically appraise the results from these reviews—instead of simply trusting the published results and conclusions. Proper skills are critical in performing these types of reviews, and the same skills are necessary to determine whether or not the results are reliable, valid, and presented free of bias, conscious or otherwise. Investigators should acquire these skills through workshops, formal training, and in collaboration with statistical experts before accepting the results as the best evidence available and allowing them to inform clinical practice.

Acknowledgment

G.C.P. received support from the Environmental Health Sciences CURES Center grant P30 ES020957.

Author Disclosure Statement

No competing financial interests exist.

References

1. Hunt M. (1997). *How Science Takes Stock: The Story of Meta-analysis*. Russell Sage Foundation, New York.
2. Fisher RA. (1932). *Statistical Methods for Research Workers*, 4th edn. Oliver and Boyd, London.
3. Pearson K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika* 25:379–410.
4. Cochran WG. (1937). Problems arising in the analysis of a series of similar experiments. *J Royal Stat Soc* 100 (Suppl 4):102–118.
5. Tippet LHC. (1931). *The Methods of Statistics*. Williams and Norgate, London.
6. Noble JH Jr. (2006). Meta-analysis: strengths, weaknesses, and political uses. *J Lab Clin Med* 147:7–20.
7. Chalmers TC, RJ Matta, H Smith and AM Kunzler. (1977). Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med* 297: 1091–1096.
8. The Cochrane Collaboration Steering Group and Subgroups. Accessed July 10, 2015.
9. Janowski M, P Walczak and I Date. (2010). Intravenous route of cell delivery for treatment of neurological disorders: a meta-analysis of preclinical results. *Stem Cells Dev* 19:5–16.
10. Ben-Yosef D, A Amit, M Malcov, T Frumkin, A Ben-Yehudah, I Eldar, N Mey-Raz, F Azem, G Altarescu, et al. (2012). Female sex bias in human embryonic stem cell lines. *Stem Cells Dev* 21:363–372.
11. Abdul Wahid SF, NA Ismail, MR Mohd-Idris, FW Jamaluddin, N Tumian, EY Sze-Wei, N Muhammad and NL Nai. (2014). Comparison of reduced-intensity and myeloablative conditioning regimens for allogeneic hematopoietic stem cell transplantation in patients with acute myeloid leukemia and acute lymphoblastic leukemia: a meta-analysis. *Stem Cells Dev* 23:2535–2552.
12. Wang K, X Chen and J Ren. (2015). Autologous bone marrow stem cell transplantation in patients with liver failure: a meta-analytic review. *Stem Cells Dev* 24:147–159.
13. Riecke J, KM Johns, C Cai, FS Vahidy, Parsha K, E Furr-Stimming, M Schiess and SI Savitz. (2015). Meta-analysis of mesenchymal stem cells in animal models of Parkinson's disease. *Stem Cells Dev* 24:2082–2090.
14. Lewis S and M Clarke. (2001). Forest plots: trying to see the wood and the trees. *BMJ* 322:1479–1480.
15. Moher D, A Liberati, J Tetzlaff, DG Altman; The PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 8:336–341.
16. Abrams KR, L Gillies and PC Lambert. (2005). Meta-analysis of heterogeneously reported trials assessing change from baseline. *Stat Med* 24:3823–3844.

Address correspondence to:

Dr. Ronald L. Thomas

Carman and Ann Adam Department of Pediatrics

Children's Research Center of Michigan

Wayne State University School of Medicine

3901 Beaubien Boulevard

Detroit, MI 48201

E-mail: rthomas@med.wayne.edu

Received and accepted for publication August 11, 2015