

Toward a transparent meta-analysis

Denise Moore Danos PhD

I am planning to perform a meta-analysis to compare sodium-glucose cotransporter-2 (SGLT2) inhibitors and dipeptidyl peptidase-4 (DPP-4) inhibitors as add on therapy to metformin in treating type II diabetes patients. I have a concern that if a meta-analysis is not appropriately performed, it could lead to a biased conclusion. How does one improve the quality of a meta-analysis?

First applied by Karl Pearson, and coined by Gene Glass, a meta-analysis is a statistical analysis that summarizes the results from a number of individual studies.³ Meta-analysis is a useful tool across several research fields, including biomedical sciences, and is increasingly preferred in systematic literature reviews. Generally speaking, meta-analysis can be performed by using either individual subject data or analytical results from individual studies. Although there are a number of advantages using individual subject data, very often those data are not readily available due to either patient information confidentiality or institutional financial considerations. Therefore, in this article, we will focus primarily on meta-analysis of aggregated data.

Meta-analysis is a quantitative analysis for systematic literature reviews, in which results from individual studies are summarized numerically. Although meta-analysis increases the statistical power of hypothesis testing and improves the precision of point estimates, due to often substantially increased sample size, it does rely on certain assumptions implicitly made in making statistical inferences. In fact, many of the critiques of meta-analysis are related to violations of these assumptions, attributable to decision making in the analysis process.

Corresponding author: Denise Moore Danos
Contact Information: ddanos@lsuhsc.edu
DOI: 10.12746/swrccc.v8i33.641

From the analytical point of view, either a fixed or random effect statistical model can be used in a meta-analysis.¹ While a fixed effect model assumes that all studies are estimating the same (fixed) treatment effect, a random effect model assumes that individual studies selected for a meta-analysis are random samples from a larger population of studies. We will briefly discuss which model is more appropriate after a discussion of the limitations of a meta-analysis. Note that these limitations are naturally associated with a meta-analysis and should be addressed appropriately in the planning stage.

SUBJECTIVE DATA QUALITY ASSESSMENT

If the data quality of individual studies is poor due to, for example, poor randomization or inadequate blinding, then regardless of the statistical model used, the conclusion of a meta-analysis can be misleading. This is known as “garbage in–garbage out.”^{3,4,5,7}

To address this issue, it is highly recommended to include only randomized clinical trials in a meta-analysis whenever feasible because non-randomized studies are more susceptible to bias. Among randomized trials, it is critical to formally assess the data quality on the basis of randomization, blindness, compliance, drop-out and withdrawal, intention-to-treat, etc., for each study. Should randomized trials not be available, observational studies, such as cohort and case-control studies, can be included, and it is essential to assess whether such studies have selection biases, and whether confounders have been adequately controlled.

There are several options for calculating study quality scores, along with recommendations on how to incorporate such scores in data analyses. The most straightforward and easy-to-implement option is to define an acceptable quality score and include only studies with higher scores than this threshold in the meta-analysis. The overall goal of data quality

assessment is to make this subjective decision making process as transparent and reproducible as possible, and it should be planned in advance, and clearly reported in the methods of analysis.

PUBLICATION BIAS

Under both the fixed and random effect models, the included studies should represent random samples from a homogeneous population. However, it is common that if a study has a small sample size and negative findings, then it is less likely to be accepted for publication, compared to studies with larger sample sizes or positive findings. Therefore, many small studies with negative findings would be missed in literature searches and thus systematically excluded from a meta-analysis. In addition, studies published in a language other than English might not be considered as part of the literature search.⁵ This exclusion of studies from a meta-analysis is called “publication bias.”¹⁻⁷ Though it is hard to address publication bias, literature searches should be comprehensive and sensitive, and it is beneficial to perform a thorough review of supplementary materials of available literature for “unpublishable” results.

Nevertheless, publication bias can be assessed post hoc by using a funnel plot with effect estimates of the studies on the x axis and either the sample size of the studies or the effect variability on the y axis.^{2,6} Should there be no serious publication bias, the plot would have a funnel shape. Otherwise, one lower side of the funnel would be blank.

HETEROGENEITY

There are two types of heterogeneity: methodological and clinical heterogeneity. Methodological heterogeneity is attributable to variations in data quality. Examples include differences in the quality of randomization, the degree of blindness, and the control of covariates. Because this type of heterogeneity is a reflection of difference in individual study quality, but not the difference in intervention effect, it should be avoided whenever possible.

On the other hand, clinical heterogeneity reflects real differences in patients, interventions, and outcome measurements, and thus is often considered a strength of a meta-analysis. Specifically, because a meta-analysis allows the inclusion of more heterogeneous groups of subjects than an individual study, it facilitates the generalization of results to a larger population. That said, it is challenging to define unambiguously the larger population represented by subjects in a meta-analysis since subjects from individual studies are often convenience samples.

While making a decision on whether to include a large study into a meta-analysis could be less subjective, it is often not so for a small study. In fact, subjects recruited in a small study are usually more homogeneous than those in a large study, and thus often a larger effect can be observed if the effect is real. Consequently, the inclusion of one or more small studies could substantially affect the overall effect estimate of a meta-analysis. Analytically, small studies are sometimes considered to be potential outliers, and there are options for removing such studies from a meta-analysis. However, this process can be subjective if the total number of studies is small.

Similar to data quality assessment, the decision on study inclusion/exclusion on the basis of heterogeneity consideration can also be subjective. Broader inclusion increases heterogeneity, but results might not apply to any specific group of subjects; narrower inclusion increases homogeneity, but results would not apply to a large population. To minimize subjectivity and improve reproducibility, it is highly recommended that researchers define the study inclusion criteria in advance and report clearly this decision making process.

Post hoc evaluation of heterogeneity can be performed both numerically and graphically. Numerically, the commonly used test statistic is the Cochran's Q calculated by summing the squared deviations of each study's estimate from the overall estimate, weighted by each study's contribution.² To evaluate the degree of heterogeneity, the calculated Q statistic is compared with a Chi-squared distribution with $k-1$ degrees of freedom, where k is the number of studies. If the test statistic is small, it indicates the studies are

homogeneous; otherwise, if it is greater than a critical value, it indicates that the studies are heterogeneous. Unfortunately, like many other methods for testing heterogeneity, the statistical power for Cochran's Q is low if the number of studies is small. Graphically, heterogeneity can be evaluated by using a forest plot,¹ in which treatment effect along with the confidence interval are plotted by trial. The variation of individual study estimate can be visually inspected.

Depending on whether the studies are homogeneous, a fixed or random effect model can be used to estimate the overall effect.⁷ If the estimates for individual studies are similar, and the results are not expected to be generalized to different populations, a fixed effect model can be applied; otherwise, a random effect model is preferred. In reality, due to the nature of meta-analysis data collection, it is often unrealistic to assume that all the studies included have similar effect size; therefore, a random effect model is generally recommended. Note that should there be minimal heterogeneity between studies, the fixed and random effect models are virtually the same.

Transparency and reproducibility are critical considerations in a meta-analysis. If all subjective decision making processes are well planned and reported, then it is possible that results from a meta-analysis can be reproduced. Therefore, researchers who have the same research interest would have the option of evaluating the impact of the subjective decision making processes by re-analyzing the meta-analytic data using modified criteria, and making their own conclusions that are more pertinent to their interests. To facilitate transparency and reproducibility, many software packages have been developed, e.g., the RevMan (<https://community.cochrane.org/help/tools-and-software/revman-5>) software for preparing and maintaining a Cochrane Review, and the GRADEpro software (www.gradepr.org) for judging the quality of a study.¹

In general, a meta-analysis performs an exhaustive literature search on an outcome and an intervention and then provides a numerical synthesis of

evidence collected from all eligible studies. With a negligible cost compared to a large randomized trial, if appropriately performed, a meta-analysis could attract a large number of citations, and provide a valuable reference for future research and practices in a specific field.

Keywords: meta-analysis, heterogeneity, publication bias

Article citation: Danos DM. Toward a transparent meta-analysis. *The Southwest Respiratory and Critical Care Chronicles* 2020;8(33):60–62

From: School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA

Submitted: 1/2/2020

Accepted: 1/14/2020

Reviewer: Shengping Yang PhD

Conflicts of interest: none

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

REFERENCES

1. Basu A. How to conduct meta-analysis: a basic tutorial 2017. <https://peerj.com/preprints/2978v1.pdf>.
2. Christensen P. Topics in meta-analysis. A literature survey. Institute of Transport Economics 2003.
3. Devinney T, Tang RW. Do we really understand a research topic? Finding answers through meta-analyses. *SSRN Electronic Journal* 2014;26.
4. Garg AX, Hackam D, Tonelli M. Systematic review and meta-analysis: when one study is just not enough. *Clin J Am Soc Nephrol* 2008;3:253–260.
5. Greco T, Zangrillo A, Biondi-Zoccai G, et al. Meta-analysis: pitfalls and hints. *Heart, Lung and Vessels* 2013;5(4): 219–225.
6. Lakens D, Hilgard J, Staaks J. On the reproducibility of meta-analyses: six practical recommendations. *BMC Psychology* 2016;4:24.
7. van Wely M. The good, the bad and the ugly: meta-analyses. *Human Reproduction* 2014;29(8):1622–1626.