Introduction to Meta-Analysis: A Guide for the Novice

September 30, 2014

Free Meta-Analysis Software and Macros

MetaXL (Version 2.0)

RevMan (Version 5.3)

Meta-Analysis Macros for SAS, SPSS, and Stata

Opposing theories and disparate findings populate the field of psychology; scientists must interpret the results of any single study in the context of its limitations. Meta-analysis is a robust tool that can help researchers overcome these challenges by assimilating data across studies identified through a literature review. In other words, rather than surveying participants, a meta-analysis surveys studies. The goal is to calculate the direction and/or magnitude of an effect across all relevant studies, both published and unpublished. Despite the utility of this statistical technique, it can intimidate a beginner who has no formal training in the approach. However, any motivated researcher with a statistics background can complete a meta-analysis. This article provides an overview of the main steps of basic meta-analysis.

Meta-analysis has many strengths. First, meta-analysis provides an organized approach for handling a large number of studies. Second, the process is systematic and documented in great detail, which allows readers to evaluate the researchers' decisions and conclusions. Third, meta-analysis allows researchers to examine an effect within a collection of studies in a more sophisticated manner than a qualitative summary.

However, meta-analysis also involves numerous challenges. First, it consumes a great deal of time and requires a great deal of effort. Second, meta-analysis has been criticized for aggregating studies that are too different (i.e., mixing "apples and oranges"). Third, some scientists argue that the objective coding procedure used in meta-analysis ignores the context of each individual study, such as its methodological rigor. Fourth, when a researcher includes low-quality studies in a meta-analysis, the limitations of these studies impact the mean effect size (i.e., "garbage in, garbage out"). As long as researchers are aware of these issues and consider the potential influence of these limitations on their findings, meta-analysis can serve as a powerful and informative approach to help us draw conclusions from a large literature base.

Identifying the Right Question

Similar to any research study, a meta-analysis begins with a research question. Meta-analysis can be used in any situation where the goal is to summarize quantitative findings from empirical studies. It can

be used to examine different types of effects, including prevalence rates (e.g., percentage of rape survivors with depression), growth rates (e.g., changes in depression from pretreatment to posttreatment), group differences (e.g., comparison of treatment and placebo groups on depression), and associations between variables (e.g., correlation between depression and self-esteem). To select the effect metric, researchers should consider the statistical form of the results in the literature. Any given meta-analysis can focus on only one metric at a time. While selecting a research question, researchers should think about the size of the literature base and select a manageable topic. At the same time, they should make sure the number of existing studies is large enough to warrant a meta-analysis.

Determining Eligibility Criteria

After choosing a relevant question, researchers should then identify and explicitly state the types of studies to be included. These criteria ensure that the studies overlap enough in topic and methodology that it makes sense to combine them. The inclusion and exclusion criteria depend on the specific research question and characteristics of the literature. First, researchers can specify relevant participant characteristics, such as age or gender. Second, researchers can identify the key variables that must be included in the study. Third, the language, date range, and types (e.g., peer-reviewed journal articles) of studies should be specified. Fourth, pertinent study characteristics, such as experimental design, can be defined. Eligibility criteria should be clearly documented and relevant to the research question. Specifying the eligibility criteria prior to conducting the literature search allows the researcher to perform a more targeted search and reduces the number of irrelevant studies. Eligibility criteria can also be revised later, because the researcher may become aware of unforeseen issues during the literature search stage.

Conducting a Literature Search and Review

The next step is to identify, retrieve, and review published and unpublished studies. The goal is to be exhaustive; however, being too broad can result in an overwhelming number of studies to review.

Online databases, such as PsycINFO and PubMed, compile millions of searchable records, including peer-reviewed journals, books, and dissertations. In addition, through these electronic databases, researchers can access the full text of many of the records. It is important that researchers carefully choose search terms and databases, because these decisions impact the breadth of the review. Researchers who aren’t familiar with the research topic should consult with an expert.

Additional ways to identify studies include searching conference proceedings, examining reference lists of relevant studies, and directly contacting researchers. After the literature search is completed, researchers must evaluate each study for inclusion using the eligibility criteria. At least a subset of the studies should be reviewed by two individuals (i.e., double coded) to serve as a reliability check. It is vital that researchers keep meticulous records of this process; for publication, a flow diagram is typically required to depict the search and results. Researchers should allow adequate time, because this step can be quite time consuming.

Calculating Effect Size

Next, researchers calculate an effect size for each eligible study. The effect size is the key component of a meta-analysis because it encodes the results in a numeric value that can then be aggregated. Examples of commonly used effect size metrics include Cohen's *d* (i.e., group differences) and Pearson's *r* (i.e., association between two variables). The effect size metric is based on the statistical form of the results in the literature and the research question. Because studies that include more participants provide more accurate estimates of an effect than those that include fewer participants, it is important to also calculate the precision of the effect size (e.g., standard error).

Meta-analysis software guides researchers through the calculation process by requesting the necessary information for the specified effect size metric. I have identified some potentially useful resources and programs below. Although meta-analysis software makes effect size calculations simple, it is good practice for researchers to understand what computations are being used.

Analysis

The effect size and precision of each individual study are aggregated into a summary statistic, which can be done with meta-analysis software. Researchers should confirm that the effect sizes are independent of each other (i.e., no overlap in participants). Additionally, researchers must select either a fixed effects model (i.e., assumes all studies share one true effect size) or a random effects model (i.e., assumes the true effect size varies among studies). The random effects model is typically preferred when the studies have been conducted using different methodologies. Depending on the software, additional specifications or adjustments may be possible.

During analysis, the effect sizes of the included studies are weighted by their precision (e.g., inverse of the sampling error variance) and the mean is calculated. The mean effect size represents the direction and/or magnitude of the effect summarized across all eligible studies. This statistic is typically accompanied by an estimate of its precision (e.g., confidence interval) and *p*-value representing statistical significance. Forest plots are a common way of displaying meta-analysis results.

Depending on the situation, follow-up analyses may be advised. Researchers can quantify heterogeneity (e.g., Q, t^2, I^2), which is a measure of the variation among the effect sizes of included studies. Moderator variables, such as the quality of the studies or age of participants, may be included to examine sources of heterogeneity. Because published studies may be biased towards significant effects, it is important to evaluate the impact of publication bias (e.g., funnel plot, Rosenthal's Fail-safe *N*). Sensitivity analysis can indicate how the results of the meta-analysis would change if one study were excluded from the analysis.

If properly conducted and clearly documented, meta-analyses often make significant contributions to a specific field of study and therefore stand a good chance of being published in a top-tier journal. The biggest obstacle for most researchers who attempt meta-analysis for the first time is the amount of work and organization required for proper execution, rather than their level of statistical knowledge.

Recommended Resources

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*.

Hoboken, NJ: Wiley.

Cooper, H., Hedges, L., & Valentine, J. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, California: Sage Publications.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Hoboken, NJ: Wiley.