Evaluating Meta-analytic Methods to Detect Outcome Reporting Bias in the Presence of Dependent Effect Sizes

Melissa A. Rodgers & James E. Pustejovsky

Large-scale meta-analyses in psychology, education, and other fields often involve studies that contribute multiple, statistically dependent effect sizes resulting from multiple treatment groups, multiple outcome measures, or longitudinal designs. An array of methods for handling dependent effect sizes are now available to applied researchers. Another challenge for meta-analysis is outcome reporting bias, which results from selective reporting and publication based on the statistical significance of study findings. The presence of selective outcome reporting is well-recognized as a critical threat to the validity of research findings synthesis; thus, inclusion of one or more methods to detect its presence is an essential component in a reliable, comprehensive synthesis. However, currently available methods for detecting selective outcome reporting all assume independent effect sizes. Despite the breadth of methodological research on how to address these challenges separately, there is limited exploration of how to handle both together—that is, how to address dependence when assessing the presence of selective outcome reporting. There is therefore a need to identify, evaluate, and disseminate methods that simultaneously address both of these challenges.

In current practice, applied meta-analyses tend to use ad hoc solutions to handle dependence when assessing selective outcome reporting, by either ignoring or aggregating effect sizes and then applying univariate tests. The aim of this study is to evaluate the performance of these approaches. We selected tests to reflect the common approaches used in recent applied research synthesis. Specifically, Trim & Fill (Duval & Tweedie, 2000), a variant of Egger's Regression (Pustejovsky & Rodgers, 2018) and a 3 Parameter Selection Model (3PSM; Hedges & Vevea, 2005) are used to detect selective outcome reporting when either a) ignoring dependency or b) taking a simple average of the effect sizes and their variances within each study. We also evaluated an Egger's Regression variant with a multivariate meta-regression model and robust variance estimation to account for dependence.

We conducted a Monte Carlo simulation study to assess the performance of the methods for Type I error rates in the absence of outcome reporting bias and power to detect outcome reporting bias when selection based on statistical significance is introduced at varying levels of censoring. Preliminary results of the simulation study suggest that ignoring dependence inflates Type I error rates with all detection methods. Aggregation also results in Type I error rates above the nominal alpha level ($\alpha = 0.05$) for Trim & Fill and 3PSM methods, while Egger's Regression variants maintains the correct Type I error. This is also true for when dependence is modeled using RVE the multivariate meta-regression variant of Egger's Regression. Further exploration of the simulation design conditions, including power to detect outcome reporting bias, will be included in the final paper. Results of this study will provide guidance to applied researchers who wish to apply valid and powerful methods to detect selective outcome reporting when synthesizing dependent effects.