

# Introduction

Meta-analysis is a critical tool for synthesizing evidence across studies to draw generalizable conclusions in fields such as education, psychology, and medicine. However, the validity of meta-analytic findings depends on the completeness and representativeness of the available data. A persistent threat to validity is selective reporting, where effect size estimates are more likely to be published, and thus included in a meta-analysis, if they are statistically significant or consistent with researchers’ hypotheses (Carter et al. 2019). Extensive evidence from diverse domains suggests that such reporting biases are widespread (e.g., O’Boyle Jr, Banks, and Gonzalez-Mulé 2017; Chan et al. 2004; Franco, Malhotra, and Simonovits 2016; John, Loewenstein, and Prelec 2012; Lancee et al. 2017; Pigott et al. 2013), resulting in distorted estimates of intervention effects and potentially misleading conclusions.

To address selective reporting bias, a variety of statistical approaches have been developed. Among the most promising are  $p$ -value selection models, which assume that the probability of an effect being reported depends on its statistical significance. These models are appealing because they are based on explicit assumptions about the selection mechanism and can be integrated into conventional meta-analytic frameworks (e.g., meta-regression). The step-function selection model, originally developed by Hedges (1992) and later extended by Vevea and Hedges (1995), assumes a piece-wise constant selection probability across different  $p$ -value thresholds (e.g.,  $\alpha = .05$ ). This approach captures plausible patterns of selective reporting and has been shown to outperform simpler diagnostics or regression-based adjustments, particularly when effect sizes are heterogeneous (Carter et al. 2019; Terrin et al. 2003).

Most existing methods for assessing and/or correcting for selective reporting—including  $p$ -value selection models—have been developed under the assumption that each study contributes a single, independent effect size. However, this assumption is increasingly unrealistic. Many meta-analyses now include multiple, dependent effect sizes per study, such as estimates derived from different outcomes, time points, or treatment comparisons. Our recent work (Pustejovsky, Citkowicz, & Joshi, 2025) addressed this gap by integrating the step-function selection model with robust variance estimation (RVE) and bootstrap methods to account for dependent effect sizes. Through simulation studies, we demonstrated that this approach reduces bias in the estimate of the overall effect size, but that there is a bias variance trade-off relative to the unadjusted meta-analytic model. Moreover, cluster bootstrapping leads to confidence intervals with coverage rates that are close to the nominal level of 0.95.

While the step-function selection model provides a structured and intuitive way to characterize selective reporting, it relies on the meta-analyst to specify “psychologically salient”  $p$ -value thresholds, such as 0.05 or 0.01. In practice, the true pattern of selection may not conform neatly to such step-wise forms. To address this limitation, the present paper introduces a beta-function selection model that allows the selection probability to vary smoothly as a function of the  $p$ -value by using the beta density to model the selection process. This model builds on earlier work by Citkowicz and Vevea (2017), extending it to the context of dependent effect sizes by incorporating RVE and bootstrap methods. The beta-function selection model both offers greater flexibility in capturing diverse forms of selection in meta-analyses with dependent effect sizes and allows meta-analysts to assess whether the form of selection matters.

We begin the paper by formally describing the beta-function selection model and outlining our proposed estimation and inference procedures. We then illustrate the model using a previously published meta-analysis, highlighting how it can reveal patterns of selective reporting not captured by the step-function selection model. Next, we report findings from an extensive simulation study evaluating the model’s performance under a variety of conditions. We conclude with summary findings, limitations, directions for future research, and implications for practice.

- Carter, Evan C, Felix D Schönbrodt, Will M Gervais, and Joseph Hilgard. 2019. "Correcting for Bias in Psychology: A Comparison of Meta-Analytic Methods." *Advances in Methods and Practices in Psychological Science* 2 (2): 115–44.
- Chan, An-Wen, Asbjørn Hróbjartsson, Mette T Haahr, Peter C Gøtzsche, and Douglas G Altman. 2004. "Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials: Comparison of Protocols to Published Articles." *Jama* 291 (20): 2457–65.
- Citkowitz, Martyna, and Jack L Vevea. 2017. "A parsimonious weight function for modeling publication bias." *Psychological Methods* 22 (1): 28–41. <https://doi.org/10.1037/met0000119>.
- Franco, Annie, Neil Malhotra, and Gabor Simonovits. 2016. "Underreporting in Psychology Experiments: Evidence from a Study Registry." *Social Psychological and Personality Science* 7 (1): 8–12. <https://doi.org/10.1177/1948550615598377>.
- Hedges, Larry V. 1992. "Modeling Publication Selection Effects in Meta-Analysis." *Statistical Science* 7 (2): 246–55.
- John, Leslie K, George Loewenstein, and Drazen Prelec. 2012. "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling." *Psychological Science* 23 (5): 524–32.
- Lancee, M, CMC Lemmens, RS Kahn, CH Vinkers, and JJ Luykx. 2017. "Outcome Reporting Bias in Randomized-Controlled Trials Investigating Antipsychotic Drugs." *Translational Psychiatry* 7 (9): e1232–32.
- O'Boyle Jr, Ernest Hugh, George Christopher Banks, and Erik Gonzalez-Mulé. 2017. "The Chrysalis Effect: How Ugly Initial Results Metamorphosize into Beautiful Articles." *Journal of Management* 43 (2): 376–99.
- Pigott, Therese D, Jeffrey C Valentine, Joshua R Polanin, Ryan T Williams, and Dericka D Canada. 2013. "Outcome-Reporting Bias in Education Research." *Educational Researcher* 42 (8): 424–32.
- Terrin, N, C H Schmid, J Lau, and I Olkin. 2003. "Adjusting for Publication Bias in the Presence of Heterogeneity." *Statistics in Medicine* 22 (13): 2113–26. <https://doi.org/10.1002/sim.1461>.
- Vevea, Jack L, and Larry V Hedges. 1995. "A General Linear Model for Estimating Effect Size in the Presence of Publication Bias." *Psychometrika* 60 (3): 419–35. <https://doi.org/10.1007/BF02294384>.