4-26-2019

Variance heterogeneity in psychological research: A Monte Carlo study of the consequences for meta-analysis

Bruce E. Blaine
St. John Fisher College, bblaine@sjfc.edu

Variance heterogeneity in psychological research: A Monte Carlo study of the consequences for meta-analysis

Abstract

Variance heterogeneity is common in psychological research. Surveys of psychological research show that variance ratios (VRs) in two-group studies average around 2.5, with a substantial minority of studies having much higher VRs. Research has established that variance heterogeneity disturbs Type I error rates of parametric tests in primary research. Fixed-effects meta-analysis is a common statistical method in psychology for synthesizing primary research, and plays an important role in cumulative science and evidence-based practice. Little is known about the consequences of variance heterogeneity for meta-analytic estimates. The present research reports a Monte Carlo study in which the results of k = 8 or 20 primary studies were generated from each of the distributions N(100, 15) and N(106, 15), for δ = 0.40 (effect size). Variance heterogeneity was created by contaminating the second distribution with elements from a N(106, 45) distribution in proportions ranging from 0.00 to 0.25, to achieve VRs ranging from 1.0 to 3.0. Each simulated fixed-effects meta-analysis (5000 replications) yielded the following estimates: Hedges' g = $CI_{95\%}$ coverage, and I^2. In the baseline (VR = 1.0) simulation, g = 0.40 and $CI_{95\%}$ coverage = 0.950. In general, larger VRs at the primary-study level were associated with smaller Hedges' gs and poorer $CI_{95\%}$ coverage at the meta-analytic level. For example, at VR = 2.6, g = 0.30 and $CI_{95\%}$ coverage = 0.801. In other words, a meta-analysis of studies that simulated the average VR in psychological research substantially underestimated the true effect and inflated the Type I error rate. Study-level variance heterogeneity also inflated estimates of between-study variance (I^2), which has implications for meta-regression modeling. This study demonstrates that widely used meta-analytic methods do not produce accurate parameter estimates in the presence of study-level variance heterogeneity.

Disciplines

Statistics and Probability

Comments

Poster presented at the Eighth Annual Conference of the Upstate Chapters of the American Statistical Association in Rochester, New York, on April 26, 2019.

Variance heterogeneity in psychological research: A Monte Carlo study of the consequences for meta-analysis

Bruce E. Blaine, Statistics and Data Sciences Program, St. John Fisher College

Introduction / Background

- Variance heterogeneity is common in psychological research. Variance ratios (VRs) in two-group studies average around 2.5, with a substantial minority of studies having much higher VRs (Ruscio & Roche 2012).
- At the primary study level, variance heterogeneity disturbs the power and Type I error rates of parametric tests, and these disturbances are worse when combined with small or unequal sample sizes (Grissom 2000).
- Meta-analysis is a common method in psychology for synthesizing primary research, but is plagued by methodological deficiencies (Hohn et al 2019; Shercliffe et al 2009). Little is known about the consequences of primary-study variance heterogeneity for meta-analytic estimates.
- This simulation study explored the effects of variance heterogeneity on meta-analytic estimates of an effect by manipulating VRs at the primary-study level in fixed effect (FE) meta-analyses. Outcomes of interest were the estimated treatment effect (Hedges g), confidence interval coverage for the parameter, and I^2.

References

Grissom, Robert. 2000. "Heterogeneity of variance in clinical data." *Journal of Consulting and Clinical Psychology*, 68: 155-165.

Hohn, Richard, Kathleen Slaney, and Donna Tafreshi. 2019. "Primary study quality in psychological meta-analyses: An empirical assessment of recent practice." *Frontiers in Psychology*, 9:2667.

Ruscio, John and Brendan Roche. 2012. "Variance heterogeneity in published psychological research: A review and a new index." *Methodology*, 8:1, 1-11.

Shercliffe, Regan, William Stahl, and Megan Tuttle. 2009. "The use of meta-analysis in psychology: A superior vintage or the casting of old wine in new bottles?" *Theory & Psychology*, 19:3, 413–430.

Tukey, John. 1960. "A survey of sampling from contaminated distributions." *Contributions to Probability and Statistics*, 448-485.

Method

- Primary studies were generated from distributions N~(100,15) and N~(106,15), establishing a treatment effect of δ=.40.
- Following Tukey (1960), variance heterogeneity was created by contaminating the second distribution with elements from a N~(106,45) distribution in proportions ranging from .00 to .25, to achieve populations in which mean VRs ranged from 1.0 to 3.0.
- Each simulated meta-analysis created primary studies by randomly sampling from these distributions (with n_1=n_2 and constrained to be between 20-40), calculating d_{Hedges} for each study, repeating k times, and doing a FE meta-analysis of the k studies. For each meta-analysis the weighted mean treatment effect, the lower and upper values of the 95% CI for estimating δ, and I^2 were retrieved.
- The simulation described above was repeated 5000 times in each of two size conditions. Small and typical-sized meta-analyses were conducted (k=20 and 40, respectively), reflecting norms in psychological meta-analytic research. Hedges' g expected value, the $CI_{95\%}$ coverage, and the I^2 expected value were retrieved from each simulation.

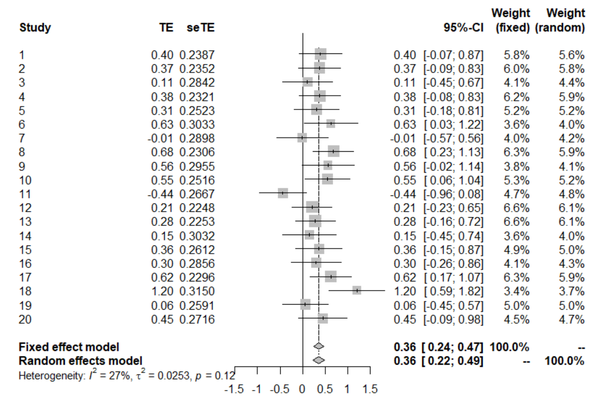
Results

- In the baseline condition (VR=1), the FE meta-analytic estimate of δ was accurate, and CI coverage (and Type I error rates) were at nominal levels.
- Increasing variance heterogeneity at the primary study level produced more biased estimates of δ, and poorer CI coverage of the parameter.
- Estimation of δ did not improve with larger meta-analyses. Moreover, larger meta-analyses produced poorer CI coverage of the parameter due to more precise interval estimates.

Table 1. Treatment effect estimates, CI coverage, and I^2 from FE meta-analyses of primary studies with variance heterogeneity.

		k=20			k=40		
contamination	VR	d_{Hedges}	95%CI coverage of δ	I^2	d_{Hedges}	95%CI coverage of δ	I^2
0	1	0.40	0.950	0.09	0.40	0.950	0.07
0.05	1.4	0.36	0.901		0.36	0.860	
0.1	1.8	0.34	0.817		0.34	0.677	
0.15	2.2	0.32	0.708		0.32	0.493	
0.2	2.6	0.30	0.603		0.30	0.322	
0.25	3	0.29	0.498	0.11	0.29	0.219	0.08

Figure 1. Forest plot of a simulated meta-analysis in the VR=1, k=20 condition



Conclusions

- Variance heterogeneity at the primary study level substantially biased FE meta-analytic estimates of a treatment effect.
- The biasing effect depends on three factors: the size of the primary study VR, which group (treatment or control) has the larger variance, and whether the contamination reflects an symmetrical or asymmetrical pattern. Future research will explore these factors.
- Asymmetrical contamination (e.g., high outliers in one group) will affect both the mean difference and the pooled variance. This should disturb I^2, further complicating the interpretation of meta-analytic estimates of a standardized mean difference.