Published in final edited form as:

Anesth Analg. 2016 October ; 123(4): 812–813. doi:10.1213/ANE.0000000000001596.

Publication Bias: The Elephant in the Review

Jarrod E. Dalton, PhD[*],

Assistant Professor of Medicine, Departments of Quantitative Health Sciences and Outcomes Research, Cleveland Clinic Lerner College of Medicine at Case Western Reserve University

Shari D. Bolen, MD, MPH, and

Associate Professor of Medicine, Epidemiology and Biostatistics, Division of Internal Medicine, Center for Health Care Research and Policy, MetroHealth Medical Center/Case Western Reserve University

Edward J. Mascha, PhD

Associate Professor of Biostatistics and Anesthesiology, Departments of Quantitative Health Sciences and Outcomes Research, Cleveland Clinic Lerner College of Medicine at Case Western Reserve University

Systematic reviews and meta-analyses serve an important role in summarizing the results of multiple investigations – particularly, rigorously-designed trials aimed at estimation of specific treatment effects. It is widely believed that, when properly conducted, results of meta-analyses can be stronger than single investigations due to increased sample sizes and diversity of settings. However, proper conduct of these investigations is simultaneously difficult and labor intensive(Blangiardo and Cameletti 2015). Clearly, the conclusion reached by a particular meta-analysis or systematic review depends on many methodological factors, including choice of search criteria, choice of literature sources, study inclusion and exclusion criteria, procedures for abstracting treatment effects, and, importantly, the extent to which things like confounding might differentially affect the results from included studies.

One key factor which may affect the conclusions reached by many such reviews is the hidden elephant of publication bias. In this issue of *Anesthesia & Analgesia*, Hedin et al. provide an assessment of the extent to which systematic reviews and meta-analyses reported in major anesthesia journals include evaluations of publication bias[1]. Describing publication bias as the tendency to publish “only results that are statistically or clinically significant”, they found that, among 207 systematic reviews meeting inclusion criteria, only 114 (55%) discussed it and 89 (43%) evaluated it. Furthermore, they found that only 68 (33%) of the reviews reported following the PRISMA guidelines (a number which may be artificially low as an estimate of the proportion following reporting guidelines in these studies, given that guidelines other than PRISMA were predominant prior to 2009), which clearly recommend the assessment of publication bias as a means for avoiding situations in which “[t]he absence of information from some studies may pose a serious threat to the validity of a review.”[2]

[*]**Correspondence:** Cleveland Clinic, 9500 Euclid Ave. (JJN3), Cleveland, OH, 44195, daltonj@ccf.org, Phone: (216) 444-9924, Fax: (216) 444-6135.

While the prevalence of publication bias in major anesthesia journals has been described previously[3], a dedicated evaluation of current practice in assessing and reporting publication bias within systematic reviews and meta-analyses in these journals was both lacking and needed. The finding that the prevalence of evaluating publication bias may be as low as 43% is troubling. However, as publication bias is difficult to evaluate among reviews of 10 or fewer studies (due to lack of power) and among reviews of non-randomized studies (due to confounding issues), perhaps a more relevant denominator might be the 25 studies which involved meta-analysis of more than 10 randomized trials. The prevalence of evaluating publication bias using the denominator of systematic reviews containing meta-analyses would then be much higher. In our experience, the degree to which publication bias may affect results is lower among this subset. Indeed, based on our review of Table 4, it appears to us that a rather small number (say, 2 to 5) of the included reviews had meaningful differences in pooled effect estimates after accounting for publication bias. Regardless of this point, however, publication bias does exist and can lead to large differences in conclusions when not adequately assessed and addressed.

Hedin et al briefly distinguished systematic review and meta-analyses in the title and in the Methods section, but subsequently referred to all articles as “systematic reviews” in the Results and Discussion sections. Systematic reviews are a reproducible method for identifying and collating all empirical evidence on a specific topic. These may or may not include a meta-analyses, which is a quantitative method for combining individual studies to estimate an overall treatment effect. The distinction is important when evaluating publication bias because systematic reviews by definition do not typically contain an assessment of overall treatment effect. We cannot verify the authors’ claims that the proportion of anesthesia reviews evaluating publication bias is lower than that of other medical fields, or that the prevalence of publication bias is higher than that of other medical fields, since their estimates are not restricted to reviews that included meta-analyses.

Appreciating the mechanisms by which publication bias and other forms of reporting bias might arise is a prerequisite to minimizing their ultimate impact on our interpretation of the literature, and correspondingly needs to be thoroughly understood by the anesthesia research community. Such mechanisms include confirmation bias (selective preference for new results which agree with prior evidence), improper study design (e.g., lack of power for detecting meaningful differences, improper specification of the relevant patient population for the intervention of interest), improper hypothesis testing practice (discontinuation of the research/manuscript development process once negative results are established in the analytic phase), lack of appropriate avenues for reporting negative studies beyond “grey” literature (such as dedicated journal sections which report negative findings), selective outcomes reporting bias (reporting only outcomes with positive findings), selective analyses reporting bias (only reporting analyses which led to positive findings), and lack of incentives for completion of mechanistic research studies (such as requirements of advance registration by journals).

Ultimately, both investigators and editorial boards bear responsibility for ensuring proper conduct so that systematic reviews and meta-analyses achieve the ideal of providing the highest-quality evidence on efficacy. As Hedin and colleagues detail, requirements for

reporting systematic reviews are inconsistent and generally scant in anesthesia journals. We agree with Hedin and colleagues that anesthesia journals should require that authors follow the PRISMA[4] (or, similarly, MOOSE[5]) guidelines when submitting systematic reviews, both in their instructions for authors and in implementation of the peer review process. At a minimum, authors of systematic reviews should be required to include a PRISMA checklist and flow diagram. We suggest that anesthesia journals might establish sections for negative studies, but at least dedicate sufficient space for publishing them, assuming they are of sufficient quality.

Investigators should also improve the quality of systematic reviews and meta-analyses. First, they should include a "grey" literature search to assist with reducing publication bias, although care should be taken when incorporating non-peer-reviewed sources such as conference abstracts which can oftentimes be works in progress, pilot studies, or in the worst case, incorrect. If evaluating pharmaceuticals or medical devices, we recommend searching the Food and Drug Administration's website in addition to clinicaltrials.gov and other sources mentioned by the authors.

Second, investigators should use appropriate techniques to assess publication bias such as Egger's regression or symmetry of funnel plots whenever there are greater than 10 studies combined in a meta-analysis (with less than 10 the assessment methods are not very reliable). One caveat in assessing publication bias is to ensure that the nature of the observed publication bias is well understood. For instance, an asymmetric funnel plot could be due to missing small positive studies (instead of the typical small negative studies), and would likely strengthen any positive conclusions as opposed to weakening the conclusion.

And third, investigators should measure the effect of publication bias on the estimated treatment effect(s) whenever publication bias is suspected. Particularly, we recommend the use of Duval and Tweedle's trim and fill technique[6] – which aims to adjust pooled treatment effect estimates to account for funnel plot asymmetry and which the authors helpfully demonstrate. When heterogeneity in treatment effects across studies is suspected (or indicated by the I^2 statistic[7]), publication bias should be evaluated within meaningful *a priori* defined subpopulations across which treatment effects might vary.

Another technique for assessing publication bias, known as the "fail-safe N method" or "Rosenthal analysis"[8], involves identifying the number of additional negative studies that would be needed to increase the P-value in a meta-analysis to above 0.05. The fail-safe N method is popular because it is simple to apply and purports to measure the fragility of reported findings. However, it is highly dependent on the treatment effects assumed for unobserved studies, i.e., it can vary greatly based on what is assumed, and is thus a fragile number itself. More importantly, it inherently places too much emphasis on the P-value and its arbitrary threshold (usually $P<0.05$) instead of focusing on the estimated treatment effects and confidence intervals, which give considerably more information. We therefore concur with the authors of the Cochrane handbook in recommending against the use of the fail-safe N method[9].

More broadly, our research community needs to move past our collective tendency to give preferential treatment to positive results. This aggregate behavior might be due to a false sense of belief that – assuming studies are well-designed and well-executed – there is more to learn from positive findings than from negative findings. This may be due to the long history of positive research findings in medicine being translated to improved quality of life and longevity, or due to the fact that positive studies are more likely to be published in high-impact journals[2] and cited more often (ultimately resulting in increased impact factors for the journals). Regardless, we should continue to work toward creating (and acculturating to) a professional environment which focuses on the quality of research questions and the quality of research methods, independent of study findings. Assessment of publication bias should become more routine as we move forward, and finding it should become less routine.

Acknowledgments

Funding Disclosure:

Financial support for this study was provided in part by the Clinical and Translational Science Collaborative of Cleveland, KL2TR000440 from the National Center for Advancing Translational Sciences component of the National Institutes of Health (NIH) and the NIH roadmap for Medical Research. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing and publishing the report.

References

1. Hedin RJ, Umberham BA, Detweiler BN, Kollmorgen L, Vassar M. Publication bias and nonreporting found in majority of systematic reviews and meta-analyses in anesthesiology journals. Anesthesia & Analgesia. 2016; 123(4):1018–1025. [PubMed: 27537925]
2. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. Annals of Internal Medicine. 2009; 151(4):W–65.
3. De Oliveira GS Jr, Chang R, Kendall MC, Fitzgerald PC, McCarthy RJ. Publication bias in the anesthesiology literature. Anesth Analg. May; 2012 114(5):1042–1048. [PubMed: 22344237]
4. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. Aug 18; 2009 151(4):W65–94. [PubMed: 19622512]
5. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. Apr 19; 2000 283(15):2008–2012. [PubMed: 10789670]
6. Duval S, Tweedie R. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000; 56(2):455–463. [PubMed: 10877304]
7. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003; 327(7414):557–560. [PubMed: 12958120]
8. Rosenthal R. The file drawer problem and tolerance for null results. Psychological Bulletin. 1979; 86:638–641.
9. Higgins, JP., Green, S. Cochrane handbook for systematic reviews of interventions. Vol. 5. Wiley; Online Library: 2008.