Letter in Response to the Article: Exploring Discordance in Evidence from Meta-Analyses and Subsequent Large-Scale Randomized Controlled Trials in Perioperative Medicine"

Ravi Y.[1], Jatin K.[2]

[1]*Department of Anesthesia & Operation Theater Technology, School of Health Sciences & Technology, Medhavi Skills University, Singtam, Sikkim, India;* [2]*Department of Anesthesia & Operation Theater Technology, School of Allied Health Sciences, Noida International University, Noida, UP, India.*

Corresponding author: Ravi Yadav, Assistant Professor, Department of Anesthesia & Operation Theater Technology, School of Health Sciences & Technology, Medhavi Skills University, Singtam, Sikkim, India.
E-mail: Yravi6945@gmail.com

Dear Editor,

We read the original study by Vanoverschelde et al. with considerable interest, which aims to investigate the discordance between meta-analyses and subsequent large-scale randomized controlled trials (RCTs; ≥1,000 participants) in perioperative medicine, as published in Acta Anaesthesiologica Belgica[1]. We commend the authors for tackling this critical issue and using trial sequential analysis (TSA) to assess the quality of evidence from meta-analyses. However, the concordance rate of 78.3% between meta-analyses and large-scale RCTs, together with the reclassification of some studies as inconclusive or false-positive using TSA, should be interpreted in the context of the methodological issues associated with the study.

Firstly, the search strategy adopted for the retrieval of large RCTs was limited to MEDLINE, using wide search terms such as "anesthesia" and "perioperative medicine," without exploring other databases (e.g., Embase, CENTRAL) or the grey literature. This is likely to result in the exclusion of relevant RCTs, leading to selection bias[2,3]. Furthermore, the search terms used, the MeSH terms, and filters were not provided, which is essential for reproducibility, a basic requirement of the PRISMA statement[4]. The procedures adopted in the meta-analysis also discuss the results of supplementary searches conducted in PubMed and the bibliography, but the details were not provided, which may have resulted in an incomplete database.

Secondly, the retrieval of large RCTs was conducted using a custom R-based text scoping approach, where the titles and abstracts were screened based on the presence of RCT-related keywords and sample-size indicators (e.g., "n=1000"). Although the validation of the search results during the development stage is acknowledged, the sensitivity and specificity of the search results were not provided, nor were examples of potential errors, which are essential for completeness.

Third, the entire screening and eligibility assessment was done by a single reviewer, with a second reviewer consulted in "uncertainties or borderline cases." In the standard method, it is recommended that there be dual independent screening to reduce the risk of error and bias, as single-reviewer approaches tend to miss 5 to 9 percent on average of the total eligible studies, which is a significant change in the results of the meta-analysis[5,6]. In this study, the risk of selection bias is increased, especially since there is a limited number of randomized controlled trials included in the study (n=10).

Fourth, the choice of the "most recent or highest quality" meta-analysis per randomized controlled trial was done without the use of any set criteria or a set instrument, such as the AMSTAR-2, which again increases the risk of bias, especially since there is a level of subjectivity in the results[7]. In addition, the results were matched pragmatically, with primary and secondary endpoints with somewhat similar definitions, and the results of the 23 comparisons were treated as independent, even though they are clustered in pairs with the randomized controlled trial and the meta-analysis, and there was no mention of the confidence intervals used in the determination of the 78.3 percent concordance. In addition, the agreement was based solely on the results being statistically significant, with either both p-values less than 0.05 or neither p-value less than 0.05, without regard to the effect size and the overlap of the confidence intervals, although the authors criticized the use of dichotomous p-values in other studies[8].

Fifth, the boundaries used by TSA were unidirectional and based on anticipated benefits, whereas in the case of the interventions, a balanced assessment of harms and null effects is necessary[9]. Assumed relative risk reductions of 20 and 30 percent were given, but the results showed "clinical risk differences" of 70 percent and 80* with inconsistent terminology and without clear clinical justification or sensitivity analyses on estimators like the Hartung-Knapp method. Cochrane Risk of Bias was used to assess the controls' event rates, but it was not fully integrated with the overall framework of the meta-analysis, and publication bias was not assessed despite the risk of small study effects[10].

Furthermore, there is a lack of registration of a prospective protocol, e.g., PROSPERO, and the full and strict application of the PRISMA and PRISMA-P checklists, which increases the risk of post hoc decisions and selective outcome reporting[11,12]. Also, the presence of extra text, e.g., "Manuscript will be submitted to a peer-reviewed journal as a thesis," is suggestive of incomplete editing.

In conclusion, although the authors are to be complimented on highlighting the limitations of the results of the meta-analytic studies, clinicians are advised to view the results of the 78.3 percent concordance and the TSA results with a critical eye in light of the methodological limitations. Future studies could include a randomized method of assessing the reliability of the results of the meta-analytic studies in the field of perioperative medicine.

References

1. Vanoverschelde H, De Pauw E, Kerremans O, Wyffels P, Wouters P. Exploring discordance in evidence from meta-analyses and subsequent large-scale randomized controlled trials in perioperative medicine. Acta Anaesth Belg 2025; 76:247–260.
2. Korevaar DA, Hooft L, Ter Riet G. Systematic reviews and meta-analyses of preclinical studies: Publication bias in laboratory animal experiments. Lab Anim 2011; 45:225–230.
3. Savoie I, Helmer D, Green CJ, Kazanjian A. Beyond Medline: reducing bias through extended systematic review search. Int J Technol Assess Health Care 2003; 19:168–178.
4. Page MJ, Moher D, Bossuyt PM et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021; 372:n160.
5. Waffenschmidt S, Knelangen M, Sieben W, Bühn S, Pieper D. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. BMC Med Res Methodol 2019; 19:132.
6. Gates A, Gates M, Duarte G et al. Evaluation of the reliability, usability, and applicability of AMSTAR, AMSTAR 2, and ROBIS: protocol for a descriptive analytic study. Syst Rev 2018; 7:85.
7. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358:j4008.
8. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature 2019; 567:305–307.
9. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol 2008; 61:64–75.
10. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315:629–634.
11. Booth A, Clarke M, Dooley G et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev 2012; 1:2.
12. Page MJ, Shamseer L, Altman DG et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: A cross-sectional study. PLoS Med 2016; 13:e1002028.

doi.org/10.56126/77.2.03