Re-evaluating the level of evidence in research: Incorporating trial sequential analysis and fragility index

Dear Editor,

One of the fundamental principles of evidence-based medicine (EBM) is the hierarchical classification of evidence, with systematic reviews and meta-analysis (SRMA) of randomised controlled trials (RCTs) at the top and expert opinion at the bottom.[1,2] This system is widely used to assess the reliability of clinical evidence. Recently, there has been an increase in the quantum of discussion on the clinical applicability[3] as well as the need to expand this classification by including new statistical measures such as the fragility index (FI) and the trial sequential analysis (TSA). One of the primary limitations of traditional evidence grading is its reliance on study design without adequately considering sample size and statistical power.[4,5] The traditional hierarchy often gives precedence to RCTs over other study designs, leading to publication bias. Positive RCTs are more likely to be published, while negative or inconclusive RCTs are often left unpublished. This can result in an overestimation of the efficacy of certain interventions.

TSA is a statistical method developed to address some of the limitations of conventional meta-analysis

in assessing the robustness of the results from clinical trials.[6] It considers the issue of multiple testing and the risk of type I and type II errors. TSA helps researchers determine whether the available evidence is sufficient to draw meaningful conclusions or if more trials are needed. TSA also calculates the required information size (RIS) based on a predefined level of statistical significance and the estimated treatment effect. If the Z curve in TSA, which is a measure of treatment effect, surpasses the RIS, a firm conclusion can be drawn, and this finding further strengthens the significant findings of SRMA. However, if the RIS is not met even with statistically significant results in SRMA, TSA suggests further trials before making conclusive statements about treatment effects.

FI is another statistical tool that has recently gained attention in the context of EBM.[7] It quantifies the robustness of a study's findings by determining how many additional events it would require to change the statistical significance of the results. In other words, it assesses the impact of individual events or outcomes on the overall conclusion of a study. A higher FI (more than 10) suggests more robust findings, while a lower FI (less than 5) indicates greater fragility in the study's conclusions. Findings of statistically significant SRMA with lower FI may be interpreted cautiously. Similarly, statistically significant SRMA findings and higher FI would further substantiate our results.

Clinicians can make better decisions by incorporating TSA and FI into the level of evidence framework. It can provide a more precise view of the quality and reliability of the evidence. If TSA suggests insufficient evidence based on the Z curve, research resources can be allocated more efficiently to the required domain. Similarly, studies that are unlikely to change clinical practice even after the addition of any further clinical research based on TSA findings can be deprioritised in favour of more informative research. Including TSA and FI may provide the advantage of a more cautious and evidence-based approach to decision-making in situations where decisions have significant repercussions for public health. Although there is a debate for the inclusion of TSA and FI in the level of evidence framework, there are aspects like educational efforts to understand these new statistical tools and integration with existing levels of evidence. Also, determining the appropriate Type I error, Type II error, and relative risk reduction which involves some subjectivity could introduce bias into the assessment in TSA and can be a disadvantage to it.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

ORCID
Nitinkumar B. Borkar: https://orcid.org/0000-0002-9624-9999
Abhijit Nair: https://orcid.org/0000-0003-2506-0301
Shilpa Meshram: https://orcid.org/0009-0009-9592-2019
Tridip D. Baruah: https://orcid.org/0000-0002-8755-4362

Nitinkumar B. Borkar, Abhijit Nair[1], Shilpa Meshram[2], Tridip D. Baruah[3]

Departments of Paediatric Surgery and [3]General Surgery, All India Institute of Medical Sciences, Raipur, Chhattisgarh, India, [1]Department of Anaesthesiology, Ibra Hospital, Ministry of Health-Oman, Ibra, Sultanate of Oman, [2]Department of Anaesthesia, BALCO Medical Center, Raipur, Chhattisgarh, India

Address for correspondence:
Dr. Nitinkumar B. Borkar,
Department of Pediatric Surgery
AIIMS Raipur, Tatibandh CG India-492 099.
E-mail: drnitinborkar25@gmail.com

Submitted: 10-Jan-2024
Revised: 03-Feb-2024
Accepted: 07-Feb-2024
Published: 13-Mar-2024

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. BMJ 1996;312:71-2.
2. Centre for Evidence-Based Medicine (CEBM). Oxford Centre for Evidence-Based Medicine-Levels of Evidence. Available from: http://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/. Last accessed on 20 December 2023.
3. Tugwell P, Knottnerus JA. Is it time to retire the hierarchy of evidence? A discussion paper. J Clin Epidemiol 2010;63:429-35.
4. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, *et al*. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis-A simulation study. PLoS One 2011;6:e25491. doi: 10.1371/journal.pone. 0025491.
5. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, *et al*. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. J Clin Epidemiol 2020;122:129-41.

6. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial Sequential Analysis reveals insufficient information size and potentially false positive results in many meta-analyses. J Clin Epidemiol 2008;61:763-9.
7. Walsh M, Srinathan SK, McAuley DF, Mrkobrada M, Levine O, Ribic C, *et al*. The statistical significance of randomized controlled trial results is frequently fragile: A case for a Fragility Index. J Clin Epidemiol 2014;67:622-8.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Access this article online

Quick response code

Website: https://journals.lww.com/ijaweb

DOI: 10.4103/ija.ija_35_24

How to cite this article: Borkar NB, Nair A, Meshram S, Baruah TD. Re-evaluating the level of evidence in research: Incorporating trial sequential analysis and fragility index. Indian J Anaesth 2024;68:403-5.

© 2024 Indian Journal of Anaesthesia | Published by Wolters Kluwer - Medknow