Beyond Cochrane's I^2: Diverse Methods for Assessing Heterogeneity in Meta-Analysis

Dear Editor,

We appreciate the interest shown by Cherif and Dziri in our meta-analysis on venous thromboembolism (VTE) in patients with amyotrophic lateral sclerosis (ALS) and thank them for their valuable feedback [1]. We have carefully considered each of their comments and offer the following responses to clarify our methodological choices, while also exploring additional heterogeneity assessment methods that may benefit future meta-analyses.

Response to Commented Methodological Aspects

Inclusion of Retrospective Studies

The inclusion of retrospective studies in our meta-analysis was a deliberate decision. ALS is a rare disease, and prospective studies on VTE incidence in ALS patients are even rarer. We agree that retrospective studies carry inherent limitations, such as selection bias and potential data gaps. However, omitting them would significantly reduce the sample size, potentially limiting the power and reliability of our findings. Including high-quality retrospective studies enabled us to construct a more comprehensive picture of VTE incidence in ALS, even if it required balancing methodological rigor with practical constraints. Although Cherif and Dziri suggest focusing on prevalence rather than incidence [1], our objective was to estimate the frequency of new VTE events, necessitating an emphasis on incidence. Our methods section clarifies the definition of incidence, noting that it requires follow-up in longitudinal studies [2]. Additionally, we excluded a cross-sectional study measuring prevalence, as indicated in our PRISMA flow diagram [2].

Heterogeneity Assessment Using I^2

We appreciate the suggestion to broaden our heterogeneity analysis by incorporating Prediction Intervals (PIs) and Tau squared (τ^2), acknowledging that these could add valuable context regarding the variability of effect sizes across studies [3]. We selected I^2 for our study due to its established prevalence in systematic reviews and ease of interpretation for readers [4]. While I^2 alone does have limitations, it is widely used and generally sufficient for evaluating variability attributable to heterogeneity [5]. We also performed sensitivity analyses to assess the potential influence of individual studies on overall findings. Additionally, our software was not equipped to perform PI analysis, which restricted our ability to include it in this instance. However, we agree that PIs could enhance interpretability by predicting the range of true effect sizes in future studies and will consider this in future analyses.

Recalculated Forest Plot with Prediction Intervals

The recalculated forest plot provided by the Cherif and Dziri, with a 95% PI of 1.3% to 6.7%, highlights that the extent of variability is minimal in our analysis. Although we found I^2 sufficient to address our research objectives, we recognize that including PIs might have further clarified the robustness and interpretability of our results. Future analyses can benefit from this more nuanced approach, especially in contexts with high heterogeneity or more complex study populations.

Different Approaches for Heterogeneity Assessment in Meta-Analysis

In response to the feedback received, we wish to outline a few additional methods for heterogeneity assessment that go beyond the Cochrane I^2 and could be beneficial for researchers seeking to enhance the depth of meta-analytic interpretation. The choice of heterogeneity metric should be informed by study objectives, data structure, and the desired balance between sensitivity and interpretability. Table 1 summarizes several heterogeneity metrics and their respective applications.

Conclusion

In summary, we are grateful for the feedback provided, which underscores the importance of methodological rigor and the evolving nature of meta-analytic tools. By exploring additional heterogeneity metrics beyond I^2, we aim to provide a resource for researchers who wish to strengthen the interpretative framework of their meta-analyses. We recognize the value of incorporating PIs and τ^2 in particular, as these metrics enhance the interpretability of study variance and predict future study outcomes. Moving forward, our work will reflect a continued commitment to robust methodologies that adapt to both traditional and emerging analytic needs in systematic review research.

LA TUNISIE MEDICALE-2024; Vol 102 (12): 983-984 DOI: 10.62438/tunismed.v102i12.5561

Table 1. Presents an overview of various metrics and methods used to assess heterogeneity in meta-analyses

Metric	Description	Use Cases	Strengths	Limitations	Possible Software
Cochrane's Q Test [6]	Tests if observed variations across study results exceed what could be expected by chance alone.	Commonly used in meta-analyses for an initial heterogeneity test.	Simple to calculate; widely recognized and understood.	Sensitive to the number of studies; low power with few studies and overly sensitive with many studies.	RevMan, MetaXL, Comprehensive Meta-Analysis (CMA)
I^2 Statistic [5]	Quantifies the percentage of variation across studies that is due to heterogeneity rather than chance.	Used in most systematic reviews and meta-analyses to quantify heterogeneity.	Widely understood; allows for comparisons across meta-analyses.	Does not convey the direction or impact of heterogeneity; interpretation can be subjective.	RevMan, MetaXL, CMA, Stata
Tau² (τ^2) [7]	Estimates the absolute amount of between-study variance in a random-effects model, indicating the level of heterogeneity.	Suitable for meta-analyses with high between-study variability.	Provides an absolute measure of heterogeneity.	Not intuitive to interpret; sensitive to the scale of outcome measures.	Stata, R (metafor package), CMA
Prediction Interval (PI) [7]	Provides an interval predicting the range of true effect sizes in future studies, accounting for between-study heterogeneity.	Useful when projecting how findings might apply in new contexts.	Offers an interpretable range, accounting for real-world variability.	Not widely reported in meta-analyses; requires larger sample sizes to be stable.	R (metafor), Stata, CMA
H^2 Statistic [8]	Ratio of total variation in effect sizes to within-study variation.	Useful for comparison between total and within-study variability.	Straightforward calculation; can complement I^2.	May be redundant if I^2 is already reported; interpretation similar to I^2.	RevMan, CMA, Stata
R^2 Statistic [4]	Calculates the ratio of between-study variance to within-study variance, showing relative inflation in confidence interval width under a random-effects model compared to a fixed-effect model.	Useful for comparing the impact of between-study variance relative to within-study variance.	Helps interpret how heterogeneity inflates confidence intervals in random-effects models.	Less common; may be confusing without familiarity; limited software support.	R (metafor)
Meta-Regression	Models relationships between study-level covariates (moderators) and outcomes, helping identify sources of heterogeneity.	Used when there are potential moderators of effect size across studies.	Provides insight into how study characteristics may impact effect sizes.	Requires larger number of studies for robust results; complex interpretation.	Stata, R (metafor, meta packages), CMA
Subgroup Analysis	Divides studies into subgroups by certain characteristics (e.g., age, intervention type) and compares effect sizes, assessing heterogeneity by subgroup.	Suitable when specific study characteristics are expected to explain heterogeneity.	Simple to interpret; offers a straightforward view of heterogeneity by specific factors.	Prone to multiple testing issues; may not reveal all sources of heterogeneity.	RevMan, Stata, CMA, R
Visual Methods (Plots)	Includes various plots (e.g., Baujat, Galbraith) to visually assess heterogeneity by illustrating study results and identifying outliers or patterns.	Useful for detecting outliers, assessing study influence, and visualizing variability.	Provides intuitive visual assessment; can complement quantitative metrics for a fuller picture of heterogeneity.	Interpretations can be subjective; may not quantify heterogeneity without accompanying statistical measures.	R (metafor), Stata

Abdullah Ashraf Hamad

Faculty of Medicine, Menoufia University, Menoufia, Egypt. Medical Research Group of Egypt, Negida Academy, Arlington, MA, USA

Email: abdullah.hamad744@gmail.com

Ibraheem M Alkhawaldeh

Faculty of Medicine, Mutah University, Al-Karak, Jordan.

References

1. CHERIF H, DZIRI C.We need a Prediction Interval to evaluate the heterogeneity of meta-analyses. Tunis Med.
2. Hamad AA, Alkhawaldeh IM, Abbas A, et al. Incidence and risk factors of venous thromboembolism in patients with amyotrophic lateral sclerosis: a systematic review and meta-analysis. Tunis Med 2024; 102: 610–615.
3. Dziri C. How to assess heterogeneity for a meta-analysis? Tunis Med 2022; 100: 353-undefined.
4. Dwivedi SN. Which is the Preferred Measure of Heterogeneity in Meta-Analysis and Why? A Revisit. Biostat Biometrics Open Access J 2017; 1.. DOI: 10.19080/BBOAJ.2017.01.555555.
5. Higgins JPT. Measuring inconsistency in meta-analyses. BMJ 2003; 327.. DOI: 10.1136/bmj.327.7414.557.
6. Cochran WG. The Combination of Estimates from Different Experiments. Biometrics 1954; 10. DOI: 10.2307/3001666.
7. Borenstein M, Higgins JPT, Hedges L V., et al. Basics of meta-analysis: I 2 is not an absolute measure of heterogeneity. Res Synth Methods 2017; 8. DOI: 10.1002/jrsm.1230.
8. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21. DOI: 10.1002/sim.1186.