

<https://www.bmj.com/content/388/bmj-2024-079839/rr-0>

Rapid response to:

Risk of Bias in Network Meta-Analysis (RoB NMA) tool

BMJ 2025; 388 doi: <https://doi.org/10.1136/bmj-2024-079839> (Published 18 March 2025) Cite this as: BMJ 2025;388:e079839

25 March 2025

Du Zhicheng

PhD candidate

Institute of Biopharmaceutical and Health Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China.

Shenzhen

Rapid Response:

Rethinking the Tools of Science: When "Subjective Judgments" Become the Invisible Enablers of Bias

Dear Editor

Rethinking the Tools of Science: When "Subjective Judgments" Become the Invisible Enablers of Bias

Medical research is often likened to a compass that guides medical decisions. But if the compass itself is misaligned, the consequences can be fatal. Carole Lunny, et al. recently introduced RoB NMA, a new tool for assessing the risk of bias in network meta-analyses (NMA), which claims to identify "invisible vulnerabilities" in studies (1). However, as I perused the article, I found that a key issue was understated: the tool was highly dependent on the subjective judgment of the evaluator, which could turn "correcting" into "creating" bias.

The article emphasizes that the use of the RoB NMA requires "collaboration between clinical and methodological experts" and acknowledges the need for "a combination of professional judgment" in the assessment. This seems reasonable, but in fact it hides a contradiction. For example, when assessing whether an intervention is a reasonable combination, experts need to empirically judge the similarity of different treatment options. But in reality, experts' definitions of "similar" can vary wildly—some strictly by dose, others by mechanism. This subjectivity is a double-edged sword: it gives tools flexibility, but it also sows the seeds of disagreement.

More critically, the authors do not delve into how this subjectivity affects the reliability of the conclusions. Imagine that the same study is evaluated by different teams, and the conclusion may be "low risk" or "high risk," depending on the background of the experts. The paper mentioned that the median evaluation took 79 minutes, but

it didn't say how to train the evaluators to reduce bias. Without uniform standards, tools can become bias amplifiers, turning supposedly objective scientific assessments into contests of expert intuition (2).

The root of such problems lies in the design logic of scientific tools. RoB NMA is like an "open recipe," with clear steps but customized condiments. For example, in deciding whether a study has publication bias, the evaluator must decide whether missing evidence affects the results. However, without quantitative thresholds or case-base support, such judgments are highly susceptible to being dominated by personal experience. When a tool relies too much on qualitative descriptions rather than quantitative indicators, its scientific nature may be diluted by subjectivity.

The solution is not to negate the value of subjective judgments, but to create guardlines for them. For example, a companion decision flow chart or a base of common dispute cases can be developed to help evaluators anchor judgment criteria. At the same time, the evaluation results of different teams on the same study are disclosed, and the tool is continuously optimized through data feedback. The progress of medical research needs tools, but the vitality of tools lies in "more accurate" rather than "once and for all."

The birth of RoB NMA is an important step, but if we ignore the "human" variable, even the perfect tool may fail in reality. The essence of science is to reduce uncertainty, not to wrap uncertainty into authority. Only when tools and human intelligence form a closed loop of "calibration-feedback" can we truly polish the compass of medical research.

References

1. Lunny C, Higgins J P T, White I R, Dias S, Hutton B, Wright J M et al. Risk of Bias in Network Meta-Analysis (RoB NMA) tool BMJ 2025; 388 :e079839 doi:10.1136/bmj-2024-079839
2. Hess, Konstantin, et al., "Efficient and Sharp Off-Policy Learning under Unobserved Confounding." arXiv preprint arXiv:2502.13022 (2025).

Competing interests: No competing interests