doi: 10.6084/m9.figshare.31292455

Letter on “Assessment of adverse effects attributed to statin therapy in product labels: a meta-analysis of double-blind randomised controlled trials”.

Alessandro Rovetta,[1] Steven D Stovitz,[2] Leonardo Biscetti,[3] Mohammad Ali Mansournia.[1,4]

1. International Committee Against the Misuse of Statistical Significance (ICAMSS), Bovezzo, Italy.
2. Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, Minnesota, USA.
3. Neurology Unit, IRCCS INRCA, Ancona, Italy.
4. Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.

Correspondence

Alessandro Rovetta
Research and Disclosure,
International Committee Against the Misuse of Statistical Significance,
25073, Bovezzo (BS), Italy.
Email: alessandrorovetta@icamss.com
Phone number: +39 3927112808
ORCID: https://orcid.org/0000-0002-4634-279X

Citation

Rovetta, A., Stovitz, S. D., Biscetti, L., & Mansournia, M. A. (2026). Letter on "Assessment of adverse effects attributed to statin therapy in product labels: A meta-analysis of double-blind randomised controlled trials". Figshare.
https://doi.org/10.6084/m9.figshare.31292455.v1

doi: 10.6084/m9.figshare.31292455

Dear Editor,

Reith et al. report an individual participant data meta-analysis of double-blind randomised trials aimed at assessing whether adverse outcomes listed in statin product labels are causally attributable to statin therapy [1]. Their analysis aimed at controlling the false discovery rate (FDR) at 5% across 66 prespecified labelled adverse outcomes and concluded that blinded trial data do not support causal relationships for most labelled conditions. While this paper is not intended to question the cardiovascular benefits of statin-based treatments, we are concerned that the above statistical strategy and its interpretation risk obscuring clinically relevant uncertainty, particularly in the context of drug safety.

First, the adoption of stringent multiplicity control prioritises the reduction of false positive findings at the expense of increased false negatives. In pharmacovigilance and safety surveillance, this trade-off is not neutral [2]. A highly sensitive screening approach will inevitably generate some false positive signals; however, such signals are typically subjected to further investigation and can be rapidly identified as spurious. In contrast, highly specific screening approaches produce fewer associations that are flagged for follow-up, so that a greater number of true positives may remain undetected. In settings where the costs of missed harms fall on patients, clinicians, and health systems, screening strategies should be calibrated to protect these more vulnerable stakeholders.

Second, the proposed interpretation of statistically “non-significant” findings often conflates statistical uncertainty with overall uncertainty and with the absence of a causal effect [3]. Indeed, conclusions framed as “not supporting a causal association” rely on the fact that the null-hypothesis P-value does not cross a prespecified threshold after FDR adjustment. Yet a P-value is not intrinsically tied to the data or the null hypothesis: it can and *should* be computed for hypotheses that are contextually relevant, as it measures the degree of compatibility (consistency, agreement) between those hypotheses and the observed data given the chosen method and its assumptions [4]. Therefore, inference should prioritise hypotheses involving clinically important adverse effects in order to safeguard the most exposed stakeholders. Moreover, focusing on the null-hypothesis P-value discards information about how compatible the data are with a range of non-null effects under the method, including effects that may be more compatible with the data than the null itself, and encourages a misleading present/absent interpretation of causality.

This issue is particularly evident in the use of 95% interval estimates. A 95% “confidence” interval should not be treated as a binary device for declaring effects present or absent. Rather, it summarises the set of effect sizes that would yield P-values above 0.05 (equivalently, are at least reasonably compatible with the observed data under the 0.05 criterion) according to the chosen method [4]. Importantly, hypotheses within the interval are not equally compatible with the data under the assumptions: compatibility is highest near the point estimate, with P-values approaching 1, and decreases as one moves toward the interval limits, where P approaches 0.05. Furthermore, hypotheses just outside the interval have P-values similar to those just inside the interval [3]. Thus, an interval that includes the null value does not indicate “no effect”, but that the data are at least reasonably compatible with the null only to the extent that they are as or more

doi: 10.6084/m9.figshare.31292455

compatible with other effect sizes within the interval, according to the method used to compute it [4]. Wide intervals indicate large statistical imprecision, meaning that the data remain consistent with effects spanning from clinically negligible to clinically important. Concluding “no causal association” from wide intervals that contain the null therefore mistakes the absence of clear evidence for clear evidence of absence [3-5].

For example, in Figure 1 (effect of statins versus placebo on events listed in statin SmPCs), several adverse outcomes – including hyperglycaemia, hypoglycaemia, acute kidney injury, thrombocytopenia, muscle-related conditions, vasculitis, immune conditions affecting muscle, Lupus-like syndrome, other skin conditions, laryngeal pain, gynaecomastia, non-specific muscle disorders, cholestasis and jaundice, hepatic failure or damage, and hepatitis – have 95% interval estimates that include relative risks exceeding 1.25, 1.50, or even 2 (i.e., Lupus-like syndrome) [1]. Similar concerns apply to Figure 2 (effect of more intensive versus less intensive statin therapy), where many interval estimates indicate a high compatibility between the data and non-negligible increases in risk under the chosen method [1]. In safety assessment, such imprecise estimates should be interpreted in light of biological plausibility and the broader clinical context, rather than being dismissed solely based on statistical “non-significance” [2-5].

Third, restricting the analysis to randomized controlled trials can be effective for shedding light on “nocebo effects”, but it is not well suited to detecting adverse drug effects in everyday contexts, where clinical and epidemiological features differ substantially from those of controlled trials [4,6]. In addition, intention-to-treat estimates in safety analyses of placebo-controlled randomized trials are biased toward the null – relative to the per-protocol effect – under non-adherence: participants assigned to treatment can receive little or no actual exposure, which dilutes risk contrasts for adverse effects between groups [7]. These aspects should be explicitly acknowledged to ensure appropriate dissemination of the findings and better inform future research directions.

Overall, these choices have consequences beyond academic debate, as the public dissemination of too simplified messages such as “statins do not cause most labelled adverse effects” overstates the decisiveness of the evidence, given both declared methodological limitations and the criticalities discussed above. We reiterate that this letter does not question the cardiovascular benefits of statin-based therapies; rather, it emphasises the importance of ensuring that their use follows careful, individualised clinical evaluation and appropriate monitoring, as should be the case for any therapeutic intervention. We thus call for a presentation that places effect sizes and their overall uncertainty at the centre of interpretation (considering both relative and absolute risks), rigorously assesses the sensitivity of conclusions to alternative error trade-offs, risk stratifications, and dosages, and clearly distinguishes lack of statistical decisiveness from lack of biological plausibility or clinical relevance. Making these judgements explicit is essential to ensure that conclusions align with the principles of public health and stakeholders' protection [2].

Declaration of interests

The authors declare no competing interests.

Funding

doi: 10.6084/m9.figshare.31292455

This work received no specific funding from any public, commercial, or not-for-profit funding agencies.

Role of the funding source
Not applicable.

Ethical approval
Not required.

Patient and public involvement
Patients and the public were not involved in the design, conduct, reporting, or dissemination of this correspondence.

Data availability
No new data were generated or analysed for this correspondence.

Contributors
All authors contributed to the conception, critical analysis, and drafting of the manuscript, and approved the final version for submission.

References

1. Cholesterol Treatment Trialists' (CTT) Collaboration. Assessment of adverse effects attributed to statin therapy in product labels: a meta-analysis of double-blind randomised controlled trials. Lancet. 2026 Feb 5. doi: 10.1016/S0140-6736(25)01578-8.

2. Greenland S. Analysis goals, error-cost sensitivity, and analysis hacking: Essential considerations in hypothesis testing and multiple comparisons. Paediatr Perinat Epidemiol. 2021 Jan;35(1):8-23. doi: 10.1111/ppe.12711.

3. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019 Mar;567(7748):305-307. doi: 10.1038/d41586-019-00857-9.

4. Rovetta A, Mansournia MA, Stovitz SD, Adams WM, Greenland S. Interpreting p values and interval estimates based on practical relevance: guidance for the sports medicine clinician. Br J Sports Med. 2025 Dec 29;59(24):bjsports-2024-109357. doi: 10.1136/bjsports-2024-109357.

5. Mansournia MA, Nazemipour M. Recommendations for accurate reporting in medical research statistics. Lancet. 2024 Feb 17;403(10427):611-612. doi: 10.1016/S0140-6736(24)00139-9.

6. Dahabreh IJ, Hernán MA. Extending inferences from a randomized trial to a target population. Eur J Epidemiol. 2019 Aug;34(8):719-722. doi: 10.1007/s10654-019-00533-2.

doi: 10.6084/m9.figshare.31292455

7. Mansournia MA, Higgins JP, Sterne JA, Hernán MA. Biases in Randomized Trials: A Conversation Between Trialists and Epidemiologists. Epidemiology. 2017 Jan;28(1):54-59. doi: 10.1097/EDE.0000000000000564.