Lipids in Health and Disease

COMMENT

Open Access

Reclaiming mendelian randomization from the deluge of papers and misleading findings

Stefan Stender[1,2*], Helene Gellert-Kristensen[1,2] and George Davey Smith[3,4]

Abstract

Mendelian randomization (MR) is a powerful epidemiological method for causal inference. However, its recent surge in popularity has brought two concerning trends. First, the public availability of summary results from genome-wide association studies has led to an explosion of low-quality two-sample mendelian randomization (2SMR) studies. These studies add minimal – if any – value and overwhelm reviewers and journals. Second, the availability of large datasets with individual-level genotype data, like UK Biobank, has spurred the development and use of novel MR methods. However, some methods are being applied without proper testing, leading to misleading results, as exemplified by recent spurious findings that are being retracted and/or corrected relating to vitamin D. What can editors and peer reviewers do to handle the deluge of 2SMR studies and the premature application of highly complex MR methods? We advise editors to simply reject papers that only report 2SMR findings, with no additional supporting evidence. For reviewers receiving such papers, we provide a template for rejection. In addition, reviewers should demand rigorous testing of novel methods, including through the use of positive and negative controls before they are applied. Rejecting non-contributory 2SMR papers and imposing intensive scrutiny to novel methods is crucial if the scientific community is to reclaim MR.

Keywords Mendelian randomization, 2SMR, Epidemiology, Nonlinear MR, UK Biobank, Vitamin D

Mendelian randomization (MR) is an epidemiological method that uses genetic variation to infer causal effects of modifiable exposures on outcomes [1]. The method relies on the use of genetic variants that associate with an exposure of interest as proxies for that exposure. If the genetic variants associated with the exposure can be robustly linked to an outcome, this strengthens evidence that the exposure itself plays a causal role in that outcome. This inference is predicated on the fact that genetic variants are less prone to confounding and are not influenced by reverse causation—sources of error that are difficult to fully address in conventional epidemiological studies.

One of the most notable successes of MR has been in confirming and refuting causal relationships between various circulating biomarkers and atherosclerosis. MR studies have accurately recapitulated the efficacy of low-density lipoprotein cholesterol (LDL-C) lowering drugs such as statins and PCSK9 inhibitors, and correctly predicted the lack of effect for other targets like vitamin D and high-density lipoprotein cholesterol (HDL-C) [2, 3].

*Correspondence:
Stefan Stender
Stefan.stender@regionh.dk
[1]Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[2]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[3]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[4]Population Health Sciences, University of Bristol, Bristol, UK

The MR toolkit has expanded in parallel with the documented success of the methodology. Among the many methods is two-sample MR (2SMR), where the exposure and outcome are derived from summary results from two separate studies. Two-sample MR was applied in 2003 in the first extended exposition of MR [4]. Since its inception the approach has been greatly facilitated by the public sharing of summary results from thousands of genome-wide association studies (GWAS). These data have been collated into accessible databases, allowing 2SMR analyses to be performed online with a few clicks of the mouse (e.g., https://app.mrbase.org/) [5]. Additionally, the data can be easily imported and analyzed using user-friendly packages in R, such as TwoSampleMR [5] and MendelianRandomization [6]. In short, 2SMR has democratized the use of MR, making the method accessible to a broader audience, including non-experts.

Unfortunately, this accessibility has led to an explosion in the number of 2SMR studies, often characterized by poor quality and a lack of scientific rigor, which was predicted in the early years of the take-off [7]. Some of these studies are being produced in papermill-like factories that help generate – at a cost – 2SMR studies, apparently for accreditation and career advancement purposes.

The problem with the abundance of low-quality 2SMR studies is now glaringly apparent [8]. A decade ago, there were about 100 MR papers published annually. In 2023, this number exceeded 3,000, and it will reach almost 5,000 in 2024 (Fig. 1). Most of these are 2SMR studies. This surge is overwhelming peer reviewers, editors, and journals. For example, *Lipids in Health and Disease* alone receives about 60 2SMR studies monthly, and between them the authors of this comment receive more than 30 requests for peer review of this type of study weekly.

What can editors and peer reviewers do to handle the deluge of 2SMR studies? We recommend that editors simply reject submitted papers that only report 2SMR with no other supportive data. One rationale behind this recommendation is that all combinations of exposure and outcome results based on data available in IEU openGWAS (https://gwas.mrcieu.ac.uk/) can be browsed online on epigraphDB.org [9, 10]. In other words, these results are, in effect, already published. Reporting them again in a scientific paper adds nothing to what can be looked up online in minutes. For peer reviewers receiving such 2SMR papers for review, we recommend using a template review for quickly dealing with them. For example:

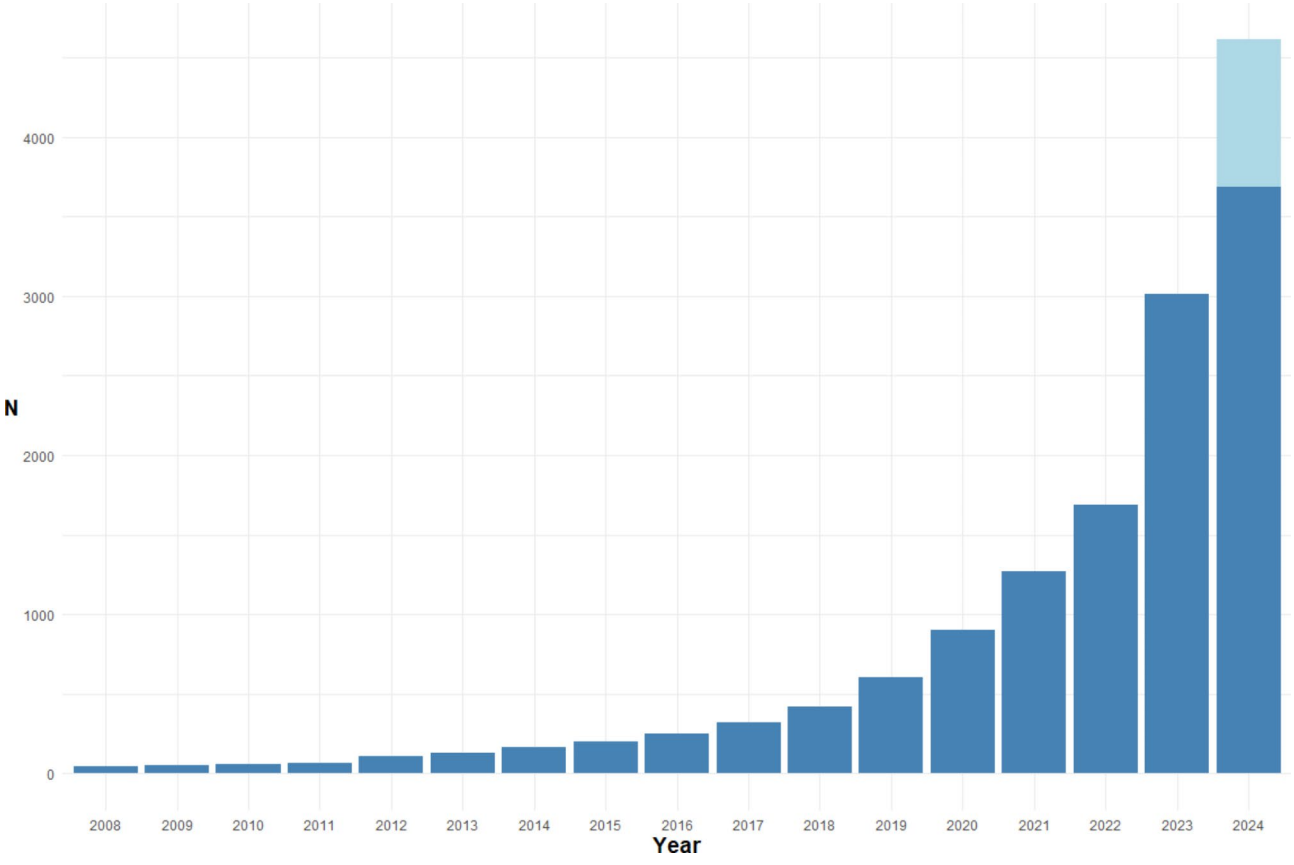

Fig. 1 Number of papers identified using the search term 'Mendelian randomization' in PubMed, stratified by year. The light blue bar for 2024 indicates the projected number for the last four months of the year

'A general comment regarding this type of 2-sample Mendelian randomization (2SMR) studies is that they have become very easy to do, owing to publicly available GWAS results and ready-to-use R-packages.
The result is a tsunami of 2SMR studies being produced and submitted for publication. Many of these stem from research papermills that churn out papers using identical methods and analytical pipelines, only changing the exposure and outcome between each paper.
These papers add little - if anything - of scientific value, and they are a huge burden on journals, editors, and peer reviewers. I recommend that the editors of [insert journal name] reject these papers without sending them for peer review in the future.'

The reader is, of course, free to use or modify this template, or parts of it, in her or his own future reviews of low-quality 2SMR papers. This could be helpful to journals in alerting their editors to this problem, although the publication of many 2SMR papers in predatory or near-predatory journals mitigates against the success of this approach.

The ready availability of individual level genotyped data from large population studies – in particular UK Biobank (UKB) – is leading to a second epidemic of papers, which can in principle investigate more interesting hypotheses than can stand-alone 2SMR studies. This has led to new methods emerging for extending MR analyses, which may become valuable tools for causal inference in population biology. Each approach will require additional assumptions, beyond those now well described in conventional MR [1]. The methods should be thoroughly tested before being rolled out, including through the use of negative and positive controls, when possible. A cautionary tale relates to vitamin D, shown by many randomized controlled trials and substantial conventional MR evidence to be unlikely to majorly modify risk of most common complex diseases at a population level [2, 11]. Applying a method purporting to be able to identify the non-linear causal effect of vitamin D modification across the range of vitamin D within UKB suggested substantial benefit in relation to cardiovascular and all cause mortality in several widely cited studies [12, 13]. Unfortunately, these findings were literally impossible [14, 15], and the first paper has been retracted and replaced with a null paper in agreement with the prior conventional MR studies [16] and for the second an editorial expression of concern has been published [17]. The approach used appears to often simply replicate observational associations, and unsurprisingly a paper applying it has appeared suggesting that raising HDL-C would reduce coronary heart disease risk in 70% of the UKB population [18]. The spurious non-linear MR papers of vitamin D have received an order of magnitude more citations since their publication than have conventional MR papers of vitamin D reporting null effects, an example of what has been called "the natural selection of bad science" [19].

The issue with MR studies applying inadequately tested complex methods is very different to that of the papermill produced 2SMR studies, as there will be few reviewers who could truly evaluate them. The simple application of negative controls would have thrown serious doubt on the reliability of the method – applying it to UKB suggests that vitamin D has a causal effect on the age and chromosomal sex of participants, an obviously nonsensical finding [20]. Body mass index – the topic of another misleading non-linear MR paper [21] – demonstrates similar nonsensical apparent causal effects when applying a non-linear MR approach [20, 22]. Such testing of the method should have been applied before it was rolled out as it was. Reviewers should ask searching questions of papers that extend MR to new domains and consider whether adequate methodological work has been done before they are implemented.

Sadly, MR has run off the rails. What is at heart a powerful and elegant scientific method for assessing causality in epidemiology is now being exploited for mass production of low-quality research, and is also reporting misleading findings, including ones that falsely repudiate the very valuable findings of earlier, veridical, MR studies. Rejecting non-contributory 2SMR papers and imposing intensive scrutiny to novel methods is imperative if the scientific community is to reclaim MR [8].

Abbreviations

MR	Mendelian randomization
GWAS	Genome-wide association study
2SMR	Two-sample Mendelian randomization
UKB	UK Biobank

Acknowledgements
Not applicable.

Author contributions
SS wrote the first draft of the comment and revised it. HGK revised the comment. GDS wrote the first draft of the comment and revised it. All authors accepted the final version of the comment for publication.

Funding
SS has received grants from Independent Research Fund Denmark and the Novo Nordisk Foundation. GDS works within the MRC Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_00032/01). None of the funders had a specific role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
GDS co-wrote the first extended exposition of Mendelian randomization and therefore has considerable intellectual investment in the approach. He has received funding for MR studies over many years and directs an MRC Unit that conducts a substantial amount of MR research.

Received: 26 August 2024 / Accepted: 2 September 2024
Published online: 07 September 2024

References

1. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. Nat Reviews Methods Primers. 2022;2:6.
2. Fang A, Zhao Y, Yang P, Zhang X, Giovannucci EL. Vitamin D and human health: evidence from mendelian randomization studies. Eur J Epidemiol. 2024;39:467–90.
3. Davey Smith G, Phillips AN. Correlation without a cause: an epidemiological odyssey. Int J Epidemiol. 2020;49:4–14.
4. Davey Smith G, Ebrahim S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32:1–22.
5. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. Elife. 2018;7:449.
6. Patel A, Ye T, Xue H, Lin Z, Xu S, Woolf B et al. MendelianRandomization v0.9.0: updates to an R package for performing mendelian randomization analyses using summarized data. Wellcome Open Res. 2023;8.
7. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol. 2016;45:1717–26.
8. Davey Smith G, Ebrahim S. Mendelian randomisation at 20 years: how can it avoid hubris, while achieving more? Lancet Diabetes Endocrinol. 2024;12:14–7.
9. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, et al. The MRC IEU OpenGWAS data infrastructure. bioRxiv. 2020. https://doi.org/10.1101/2020.08.10.244293.
10. Liu Y, Elsworth B, Erola P, Haberland V, Hemani G, Lyon M, et al. EpiGraphDB: a database and data mining platform for health data science. Bioinformatics. 2021;37:288.
11. Gallagher JC, Rosen CJ, Vitamin D. 100 years of discoveries, yet controversy continues. Lancet Diabetes Endocrinol. 2023;11:362–74.
12. Sofianopoulou E, Kaptoge SK, Afzal S, Jiang T, Gill D, Gundersen TE, et al. RETRACTED: estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and mendelian randomisation analyses. Lancet Diabetes Endocrinol. 2021;9:837–46.
13. Zhou A, Selvanayagam JB, Hyppönen E. Non-linear mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. Eur Heart J. 2022;43:1731–9.
14. Davey Smith G. Non-linear mendelian randomization publications on vitamin D report spurious findings and require major correction. Eur Heart J. 2024;45:2677–8.
15. Davey Smith G. Mendelian randomisation and vitamin D: the importance of model assumptions. Lancet Diabetes Endocrinol. 2023;11:14.
16. The Editors Of The Lancet Diabetes Endocrinology. Retraction and republication-estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and mendelian randomisation analyses. Lancet Diabetes Endocrinol. 2024;12:8.
17. Expression of Concern. Non-linear mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. Eur Heart J. 2024;45:2305.
18. Chen JX, Li Y, Zhang YB, Wang Y, Zhou YF, Geng T, et al. Nonlinear relationship between high-density lipoprotein cholesterol and cardiovascular disease: an observational and mendelian randomization analysis. Metabolism. 2024;154:155817.
19. Smaldino PE, McElreath R. The natural selection of bad science. R Soc Open Sci. 2016;3:160384.
20. Hamilton FW, Hughes DA, Spiller W, Tilling K, Davey Smith G. Non-linear mendelian randomization: detection of biases using negative controls with a focus on BMI, vitamin D and LDL cholesterol. Eur J Epidemiol. 2024;39:451–65.
21. Sun Y-Q, Burgess S, Staley JR, Wood AM, Bell S, Kaptoge SK, et al. Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear mendelian randomisation analyses. BMJ. 2019;364:l1042.
22. Wade KH, Hamilton FW, Carslake D, Sattar N, Davey Smith G, Timpson NJ. Challenges in undertaking nonlinear mendelian randomization. Obes (Silver Spring). 2023;31:2887–90.

Publisher's note