journal title, 2022, 1–12
doi: DOI HERE
Advance Access Publication Date: Day Month Year
Paper

`simmr`: An open-source tool to perform simulations in Mendelian Randomization

Noah Lorincz-Comi,[1,*] Yihe Yang[1] and Xiaofeng Zhu[1]

[1]Department of Population Quantitative Health Sciences, Case Western Reserve University, Street, Postcode, Cleveland, Ohio, USA

[*]Corresponding author. noahlorinczcomi@gmail.com

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Mendelian Randomization (MR) has become a popular tool for inferring causality of risk factors on disease. There are currently over 50 different methods available to perform MR, reflecting the extremely active research area. It would be desirable to have a standard simulation environment to objectively evaluate the existing and future methods. We present `simmr`, an open-source software for performing simulations to evaluate the performance of different MR methods. Researchers can directly modify the `simmr` source code and parameters with the intended purpose of eventually arriving at a widely accepted protocol for researchers to evaluate the performance of MR methods in different real-world conditions. (I don't understand the last sentence

Key words: Mendelian Randomization, bias, causal inference, instrumental variables, simulation, statistical genetics

Introduction

Mendelian Randomization is a genetic instrumental variable method that has become a popular statistical tool for estimating causal effects of risk factors on disease [47, 45]. The validity of inferences made using MR relies heavily on the satisfaction of three primary assumptions [16]: the genetic instruments are (i) strongly associated with the exposure(s) [14, 49], (i) not associated with the outcome conditional on the exposures [27], and (iii) not associated with any confounders of the exposure(s)-outcome relationship(s) [36]. Violation of any of these assumptions can lead to bias in causal estimation, false positive, and/or false negative causal findings [2]. Three additional challenges are present in MR analyses: (a) some individuals may have been present in both the exposure(s) and outcome GWAS [6], (b) applying strict IV selection criteria to satisfy assumption (i) above can introduce a 'winner's curse' bias [30], and (c) strong LD between IVs [21], and its imprecise estimation using reference panels, can drastically inflate false positive rates [23].

Many MR methods have been introduced in the literature to address different subsets of assumptions (i)-(iii) and scenarios (a)-(c) in either the single-exposure (univariable) or multiple-exposure (multivariable) settings [1]. These methods include IVW [5], dIVW [60], pIVW [57], MR-Egger [3], MR-RAPS [63], MRAID [61], MRMix [41], MR-cML [58], MVMR-cML [33], MR-PRESSO [53], IMRP [64], MR-Median [3], MR-MaxLike [7], MR-Corr [11], MR-Robust [44], MR-Lasso [31], MR-Conmix [8], CAUSE [36], MR-CUE [13], MR-Horse [25], MR-BMA [65], MR-Robin [22], EMIC [29], MR-Mode [26], MRBEE [35], MR-Lap [37], the Wald test [39], JAM [38], MR using factor analysis [40], mixIE [32], MRMO [18], BMRMO [17], BWMR [62], moPMR-Egger [34], sisVIVE [31], MR-LDP [12], MR-CIP [56], MR-PATH [28], MR-Clust [20], BayesMR [4], BMRE [50], MR-link [52], OMR [55], CoJo [59], MR using PCA [10], and GRAPPLE [54]. Clearly, the literature is saturated with MR estimators. In each of these studies introducing the methods, simulation was performed to evaluate the performance of new and existing methods.

The variability of simulation settings reported in the literature is vast. Some use 50 or fewer SNPs to explain most of phenoptypic [53, 64] or genetic [33] variance while others have used hundreds [35] or even thousands [37] of SNPs. Additionally, some have used GWAS sample sizes no larger than 500 [18] while others have used GWAS sample sizes no less than 100k [56]. Nearly all multivariable MR simulation sets do not reflect the reality. For example, some have 50 or less IVs in simulation [33, 46, 48], but experience with real data analyses with multiple exposures suggest that hundreds or thousands of SNPs meeting the valid IV conditions may be identified [15, 35]. Since weak IV bias generally becomes more severe as more IVs are used [14, 35], MVMR simulations in which very few IVs are used may unrealistically reflect optimal performance of MVMR methods, which may break down in practice. Some authors have additionally modelled highly phenotypically correlated phenotypes that are completely independent genetically [33, 24], a scenario unlikely to be encountered in practice.

Some of these issues are from researchers reproducing the unrealiastic simulation settings of others (e.g., [64] reproduced [53]; [33] reproduced [24]. Other challenges may exist because of a lack of a unified framework for performing simulations in MR. Both explanations have produced a literature of results potentially sensitive to the unique simulation conditions that may or may not mirror reality. Indeed, some independent simulations intended to mimic the same reality with nearly identical reported simulation settings have even produced conflicting results. As examples, the literature contains some conflicting evidence about the bias in dIVW in the absence of sample overlap and horizontal pleiotropy ([37] and [60]); the power of MR-Robust in the presence of 50% invalid IVs, being either less than 20% or greater than 80% ([33] and [51]); or MR-RAPS having Type I error greater than 80% or controlled at 5% in the presence of horizontal pleiotropy ([42] and [58]).

We introduce `simmr`, an open-source and easily accessible tool for generating data to use in MR simulations. `simmr` can be used in both the univariable and multivariable MR settings and users can simultaneously specify different properties of uncorrelated horizontal pleiotropy, correlated horizontal pleitoropy, weak instrument bias, GWAS sample size and overlap, winner's curse, and linkage disequilibrium between instrumental variables. With no standard set in the literature, the purpose

of the publicly modifiable `simmr` software is to begin the construction of a standard simulation environment in which fair comparisons can be made between competing MR estimators in different settings that are encountered in practice.

Implementation

Simulation models

The simulation data-generating process is based on the directed acyclic graph (DAG) in Figure 1. Simulation data are generated under the directed acyclic graph in Figure 2 and the following models:

$$U = \mathbf{g}^\top \boldsymbol{\gamma}^C + \epsilon_U, \tag{1}$$

$$\mathbf{x} = \mathbf{B}^\top \mathbf{g} + \pi_{\mathbf{x}} U + \boldsymbol{\epsilon}_{\mathbf{x}}, \tag{2}$$

$$Y = \mathbf{x}^\top \boldsymbol{\theta} + \pi_Y U + \mathbf{g}^\top \boldsymbol{\gamma}^U + \epsilon_Y \tag{3}$$

$$= \boldsymbol{\alpha}^\top \mathbf{g} + \widetilde{\epsilon}_Y \tag{4}$$

where U is a confounder of the relationship between exposure(s) $\mathbf{x}$ and outcome Y, $\boldsymbol{\theta}$ represents the corresponding causal effect(s), $\mathbf{B}$ represents true associations between $\mathbf{g}$ and $\mathbf{x}$, and $\boldsymbol{\gamma}^C$ and $\boldsymbol{\gamma}^U$ are respectively CHP and UHP effects. Users of `simmr` can fix the variance explained in each phenotype by the others and partition the variance in $\mathbf{x}$ explained by $\mathbf{g}$ into effects from UHP, CHP, valid, and weak SNPs using the procedure in Algorithm 1:

Figure 3 shows an example of data generated under the above models alongside computation times for different combinations of GWAS sample overlap, GWAS sample sizes, and numbers of exposure SNPs. This set of SNPs can then be reduced by the user-specified significance and LD pruning thresholds [19], both of which are implemented in the `simmr` software. Users also have the option to perform IV pruning such that a specific F-statistic for instrument strength [9] is achieved for each exposure. Some MR methods use the correlation matrix $\boldsymbol{\Omega}$ among GWAS estimation errors across cohorts to correct for weak instrument bias (e.g., [33], [35], [12], [37]). These methods calculate $\boldsymbol{\Omega}$ using non-significant GWAS summary statistics [35, 37] [cite ref 63 and Zhu et al AJHG 2015] or LD score regression [33, 12]. `simmr` will directly calculate $\boldsymbol{\Omega}$ from realized values of $(\mathbf{x}, Y)$ without any estimation error and return it to the user.

Software

`simmr` uses the popular R software v4.0.0 or later [43] and only requires downloading 3 files: `basicfunctions.R`, `set_params.R`, and `generate_data.R`. Users of `simmr` change the parameters in the `set_params.R` file to their desired quantities. These parameters are displayed in blue text in Figure 2. Next, another file named `generate_data.R` is sourced and the simulation data is automatically loaded into the user's R environment where simulations can be performed. The `simmr` software is available at `https://github.com/noahlorinczcomi/simmr`, where a tutorial is present and researchers can directly change any of the source code to improve it. Changes are tracked automatically and earlier versions can be restored at any time.

Algorithm 1 Pseudo-code of simulated data generation using `simmr`

Require: $\text{Var}(Y|x_k, U)$ for $k = 1, ..., p$, $\text{sign}(\text{Cov}[Y, x_k])$, $\text{Cov}(\mathbf{x}|U)$, $\text{Cov}(\mathbf{x}|\mathbf{g})$, phenotypic and genetic correlation between exposures $\mathbf{S}_P$ and $\mathbf{S}_G$, exposures and outcome GWAS sample size n_1 and n_0 and proportion of overlap p_{01}, number of causal exposure SNPs M and their LD structure $\mathbf{R}$, number of UHP and CHP causal exposure SNPs m^U and m^C, respective variance in Y and U explained by UHP and CHP causal exposure SNPs $\sigma^2_{\gamma^U}$ and $\sigma^2_{\gamma^C}$, LD pruning threshold for IV selection κ^2, simulation type (either winner's curse or weak IVs), type of standardization to apply to the MR summary data (Z-score standardization or no standardization), P-value threshold for IV selection if performing winner's curse simulations τ, F-statistic for IV strength of association with exposures (F) if performing a weak instrument simulation.

Preliminaries: All phenotypes will have mean 0 and variance 1. UHP, CHP, and SNPs that are otherwise valid IV candidates are put into mutually exclusive groups (see Figure 2).

1. Draw $n_0 + n_1$ genotypes from $g_j \sim \text{Binomial}(2, 0.3)$ and standardize to $\text{E}(g_j) = 0, \text{Var}(g_j) = 1$; $j = 1, ..., M$

2. Draw m^C CHP effects from $\gamma^C_j \sim \text{Uniform}(-1, 1)$ and re-scale to explain $\sigma^2_{\gamma^C}$ variance in U

3. Draw m^U CHP effects from $\gamma^U_j \sim \text{Uniform}(-1, 1)$ and re-scale to explain $\sigma^2_{\gamma^U}$ variance in Y

4. Draw un-transformed true SNP associations with the exposures ($\widetilde{\mathbf{B}}$) from a matrix normal distribution with row-wise covariance $\mathbf{R}$ and column-wise covariance $\mathbf{S}_G$ and re-scale the columns such that $\left[\text{diag}(\widetilde{\mathbf{B}})\right]_k = 1 - \left[\text{diag}(\text{Cov}\{\mathbf{x}|\mathbf{g}\})\right]_k$

5. Transform $\widetilde{\mathbf{B}}$ using the copula method to convert to a multivariate uniform distribution: $\left[\mathbf{B}\right]_{jk} = \Phi\Big(\left[\tilde{B}\mathbf{S}_G^{-1/2}\right]_{jk}\Big)$

6. Set $\pi_Y = -1$ and $\theta_k = \text{sign}(\text{Cov}[Y, x_k])$ and re-scale each to achieve $\text{Var}(Y|x_k, U)$ for $k = 1, ..., p$

7. Define sets $\mathcal{P}_\mathbf{x}$ and $\mathcal{P}_Y$ of simulated individuals to respectively use in the exposure and outcome GWAS, of which $p_{01} \times \min(n_0, n_1)$ elements will overlap

8. Perform GWAS using OLS on phenotypes $Y_{\mathcal{P}_Y}$ and $\mathbf{x}_{\mathcal{P}_\mathbf{x}}$ using OLS

9. Standardize all GWAS summary data according to user input (one of 'z' or 'none')

10. Perform IV selection according to κ^2, τ, and F to obtain IV set $\mathcal{S}$

Ensure: Standardized GWAS estimates for the set of IVs: $\widehat{\mathbf{B}}_\mathcal{S}$, $\widehat{\text{SE}}(\widehat{\mathbf{B}}_\mathcal{S})$, $\widehat{\boldsymbol{\alpha}}_\mathcal{S}$, and $\widehat{\text{SE}}(\widehat{\boldsymbol{\alpha}}_\mathcal{S})$ as the corresponding R objects `bx`, `bxse`, `by`, and `byse`; their unstandardized correlates `bx_unsted`, `bxse_unstd`,`by_unstd`,`byse_unstd`; indications for each IV if it was generated under a vertical pleiotropy, UHP, or CHP model as `IVType`; $\boldsymbol{\theta}$ as `theta`; $\boldsymbol{\Omega}$ as `RhoME`; $\mathbf{R}_\mathcal{S}$ as `LDMatrix`.

Input and output

Users of `simmr` input all parameters that are present in the `set_params.R` file. The output of executing the `generating_data.R` file after all parameters from `set_params.R` are present in the global environment is the following summary data:

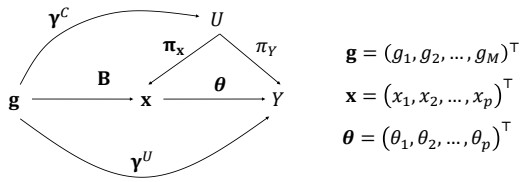

Fig. 1: This DAG is used to produce the simulation models described in the main text. Users of `simmr` can modify all parameters that are present in this DAG.

(i) standardized SNP association estimates between all IVs and each exposure in the matrix (stored in the object `bx`), (ii) their corresponding standardized standard errors in (`bxse`), (iii) standardized estimates of SNP association between each IV and the outcome in the vector (`by`), (iv) their corresponding standardized standard errors in (`byse`), (v) all corresponding unstandardized estimates of SNP association between the exposures and outcome and their corresponding standard errors (`bx_unstd,bxse_unstd,by_unstd,byse_unstd`), (vi) the number of selected IVs (`mIVs`), (vii) an indication for each IV if it was generated under a non-UHP/CHP, UHP, or CHP model (`IVType`), (viii) the matrix of LD correlations between all IVs in (`LD`), (ix) the true causal effects of each exposure on the outcome in the vector (`theta`), (x) and the matrix of correlations between GWAS estimation errors for the outcome and all exposures in (`RhoME`).

Conclusion

There is currently no standard simulation framework for performing simulations in Mendelian Randomization research. Different researchers have independently performed simulations designed to reflect the similar real world conditions, but the performance of the same methods can vary greatly. The MR literature is replete with MR estimators, with each at some point having a demonstrated advantage over others in simulation. The transferability of their performance to real world settings may be in question if the advantages can only be reached in specific conditions. We present `simmr`, an open-source software for performing simulations to evaluate the performance of Mendelian Randomization methods. Researchers can directly modify the `simmr` software. It is our intention that the community will use this opportunity to establish an accepted procedure for performing simulations using MR. As `simmr` is refined and expanded, we expect it will provide a useful tool to facility future MR method developments and evaluations. [For a software, you may need to tell what are inputs and what are output so users can use it]

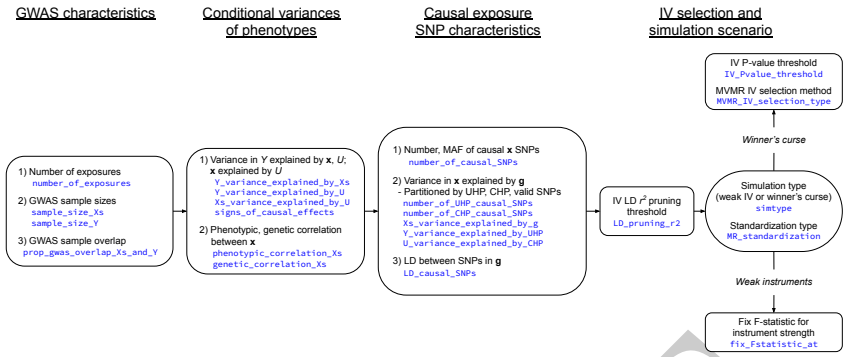

Fig. 2: This figure provides a conceptual overview of the `simmr` data-generating process. First, users specify the sizes of exposure and outcome GWAS sample sizes and their degree of overlap. Next, users specify the true causal model by changing variances in each phenotype explained by each other and by the SNPs that are causally related to the exposures. Next, the user specifies the particular characteristics of the causal SNPs, inputting the degree of UHP, CHP, and LD between them. Finally, the user can perform IV selection to evaluate winner's curse or weak instrument bias. `simmr` commands are presented in blue text underneath the description for each parameter.

Availability and requirements

Project name: `simmr`
Project home page: `https:github.com/noahlorinczcomi/simmr`
Operating system(s): Platform independent
Programming language: R
Other requirements: R 4.1.1 or higher; `mvnfast` package
License: MIT
Any restrictions to use by non-academics: License needed

Competing interests

No competing interest is declared.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

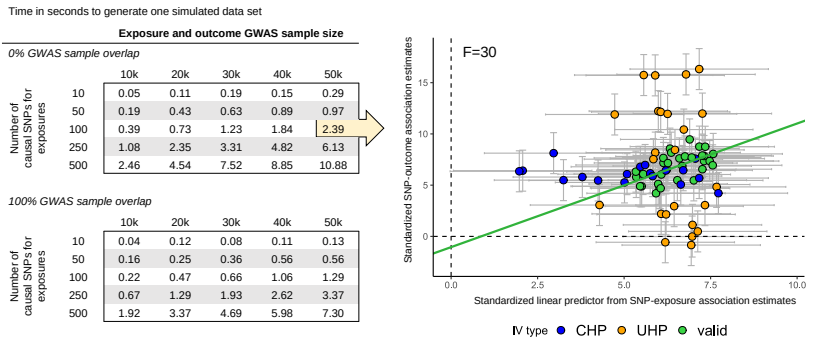

Fig. 3: This figure shows computation times to generate one simulated data set using `simmr` (left) and an example of the generated summary data for evaluation of MR methods (right). The data in the right panel is only of the final selected and pruned set of instruments. The original number of causal SNPs was 100, but the selected number of IVs to achieve an F-statistic of 30 was 91. The full code used to produce this figure is available at `https://github.com/noahlorinczcomi/simmr`. The x-axis in the right panel is the standardized linear predictor used in multivariable MR ($\hat{\mathbf{B}}\boldsymbol{\theta}$) and the y-axis is the estimated association of the IVs with the outcome in standardized scale. The green line is the linear fit to the green points representing valid instrumental variables. The yellow arrow in the left panel indicates that the time it took to generate the data in the right panel was 2.4 seconds when there is no GWAS sample overlap and GWAS sample sizes are 50k. When the exposure and outcome GWAS sample overlap is 100%, computation time was reduced to 1.39 seconds.

Availability of data and materials

The `simmr` software is available at `https://github.com/noahlorinczcomi/simmr`, where you can also find all R code to reproduce the results in Figure 3.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grant HG011052 (to XZ) from the National Human Genome Research Institute (NHGRI)

Authors' contributions

NLC conceived of the idea, wrote the manuscript, and created the software. YY and XZ edited the manuscript and software.

Acknowledgements

Not applicable.

References

1. Frederick J Boehm and Xiang Zhou. Statistical methods for mendelian randomization in genome-wide association studies: A review. *Computational and Structural Biotechnology Journal*, 20:2338–2351, 2022.
2. Jack Bowden, George Davey Smith, and Stephen Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *International journal of epidemiology*, 44(2):512–525, 2015.
3. Jack Bowden, George Davey Smith, Philip C Haycock, and Stephen Burgess. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.*, 40(4):304–314, 2016.
4. Ioan Gabriel Bucur, Tom Claassen, and Tom Heskes. Inferring the direction of a causal link and estimating its effect via a bayesian mendelian randomization approach. *Statistical Methods in Medical Research*, 29(4):1081–1111, 2020.
5. Stephen Burgess, Adam Butterworth, and Simon G Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology*, 37(7):658–665, 2013.
6. Stephen Burgess, Neil M Davies, and Simon G Thompson. Bias due to participant overlap in two-sample mendelian randomization. *Genetic epidemiology*, 40(7):597–608, 2016.
7. Stephen Burgess, Frank Dudbridge, and Simon G Thompson. Combining information on multiple instrumental variables in mendelian randomization: comparison of allele score and summarized data methods. *Statistics in medicine*, 35(11):1880–1906, 2016.
8. Stephen Burgess, Christopher N Foley, Elias Allara, James R Staley, and Joanna MM Howson. A robust and efficient method for mendelian randomization with hundreds of genetic variants. *Nat. Commun.*, 11(1):1–11, 2020.
9. Stephen Burgess, Simon G Thompson, and Crp Chd Genetics Collaboration. Avoiding bias from weak instruments in mendelian randomization studies. *International journal of epidemiology*, 40(3):755–764, 2011.
10. Stephen Burgess, Verena Zuber, Elsa Valdes-Marquez, Benjamin B Sun, and Jemma C Hopewell. Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. *Genetic epidemiology*, 41(8):714–725, 2017.
11. Qing Cheng, Tingting Qiu, Xiaoran Chai, Baoluo Sun, Yingcun Xia, Xingjie Shi, and Jin Liu. Mr-corr2: a two-sample mendelian randomization method that accounts for correlated horizontal pleiotropy using correlated instrumental variants. *Bioinformatics*, 38(2):303–310, 2022.
12. Qing Cheng, Yi Yang, Xingjie Shi, Kar-Fu Yeung, Can Yang, Heng Peng, and Jin Liu. Mr-ldp: a two-sample mendelian randomization for gwas summary statistics accounting for linkage disequilibrium and horizontal pleiotropy. *NAR genomics*

and bioinformatics, 2(2):lqaa028, 2020.

13. Qing Cheng, Xiao Zhang, Lin S Chen, and Jin Liu. Mendelian randomization accounting for complex correlated horizontal pleiotropy while elucidating shared genetic etiology. *Nat. Commun.*, 13(1):1–13, 2022.
14. Neil M Davies, Stephanie von Hinke Kessler Scholder, Helmut Farbmacher, Stephen Burgess, Frank Windmeijer, and George Davey Smith. The many weak instruments problem and mendelian randomization. *Statistics in medicine*, 34(3):454–468, 2015.
15. Neil Martin Davies, W David Hill, Emma L Anderson, Eleanor Sanderson, Ian J Deary, and George Davey Smith. Multivariable two-sample mendelian randomization estimates of the effects of intelligence and education on health. *Elife*, 8:e43990, 2019.
16. Christiaan de Leeuw, Jeanne Savage, Ioan Gabriel Bucur, Tom Heskes, and Danielle Posthuma. Understanding the assumptions underlying mendelian randomization. *European Journal of Human Genetics*, 30(6):653–660, 2022.
17. Yangqing Deng, Dongsheng Tu, Chris J O'Callaghan, Derek J Jonker, Christos S Karapetis, Jeremy Shapiro, Geoffrey Liu, and Wei Xu. A bayesian approach for two-stage multivariate mendelian randomization with mixed outcomes. *Statistics in Medicine*, 2023.
18. Yangqing Deng, Dongsheng Tu, Chris J O'Callaghan, Geoffrey Liu, and Wei Xu. Two-stage multivariate mendelian randomization on multiple outcomes with mixed distributions. *Statistical Methods in Medical Research*, page 09622802231181220, 2022.
19. Frank Dudbridge and Paul J Newcombe. Accuracy of gene scores when pruning markers by linkage disequilibrium. *Human heredity*, 80(4):178–186, 2016.
20. Christopher N Foley, Amy M Mason, Paul DW Kirk, and Stephen Burgess. Mr-clust: clustering of genetic variants in mendelian randomization with similar causal estimates. *Bioinformatics*, 37(4):531–541, 2021.
21. Apostolos Gkatzionis, Stephen Burgess, and Paul J Newcombe. Statistical methods for cis-mendelian randomization with two-sample summary-level data. *Genetic epidemiology*, 47(1):3–25, 2023.
22. Kevin J Gleason, Fan Yang, and Lin S Chen. A robust two-sample mendelian randomization method integrating gwas with multi-tissue eqtl summary statistics. *bioRxiv*, pages 2020–06, 2020.
23. Kevin J Gleason, Fan Yang, and Lin S Chen. A robust two-sample transcriptome-wide mendelian randomization method integrating gwas with multi-tissue eqtl summary statistics. *Genetic epidemiology*, 45(4):353–371, 2021.
24. Andrew J Grant and Stephen Burgess. Pleiotropy robust methods for multivariable mendelian randomization. *Statistics in medicine*, 40(26):5813–5830, 2021.
25. Andrew J Grant and Stephen Burgess. A bayesian approach to mendelian randomization using summary statistics in the univariable and multivariable settings with correlated pleiotropy. *bioRxiv*, pages 2023–05, 2023.
26. Fernando Pires Hartwig, George Davey Smith, and Jack Bowden. Robust inference in summary data mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology*, 46(6):1985–1998, 2017.

27. Gibran Hemani, Jack Bowden, and George Davey Smith. Evaluating the potential role of pleiotropy in mendelian randomization studies. *Human molecular genetics*, 27(R2):R195–R208, 2018.
28. Daniel Iong, Qingyuan Zhao, and Yang Chen. A latent mixture model for heterogeneous causal mechanisms in mendelian randomization. *arXiv preprint arXiv:2007.06476*, 2020.
29. Lin Jiang, Lin Miao, Guorong Yi, Xiangyi Li, Chao Xue, Mulin Jun Li, Hailiang Huang, and Miaoxin Li. Powerful and robust inference of complex phenotypes' causal genes with dependent expression quantitative loci by a median-based mendelian randomization. *The American Journal of Human Genetics*, 109(5):838–856, 2022.
30. Tao Jiang, Dipender Gill, Adam S Butterworth, and Stephen Burgess. An empirical investigation into the impact of winner's curse on estimates from mendelian randomization. *medRxiv*, pages 2022–08, 2022.
31. Hyunseung Kang, Anru Zhang, T Tony Cai, and Dylan S Small. Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. *J. Am. Stat. Assoc.*, 111(513):132–144, 2016.
32. Zhaotong Lin, Yangqing Deng, and Wei Pan. Combining the strengths of inverse-variance weighting and egger regression in mendelian randomization using a mixture of regressions model. *PLoS genetics*, 17(11):e1009922, 2021.
33. Zhaotong Lin, Haoran Xue, and Wei Pan. Robust multivariable mendelian randomization based on constrained maximum likelihood. *The American Journal of Human Genetics*, 110(4):592–605, 2023.
34. Lu Liu, Ping Zeng, Fuzhong Xue, Zhongshang Yuan, and Xiang Zhou. Multi-trait transcriptome-wide association studies with probabilistic mendelian randomization. *The American Journal of Human Genetics*, 108(2):240–256, 2021.
35. Noah Lorincz-Comi, Yihe Yang, Gen Li, and Xiaofeng Zhu. Mrbee: A novel bias-corrected multivariable mendelian randomization method. *bioRxiv*, pages 2023–01, 2023.
36. Jean Morrison, Nicholas Knoblauch, Joseph H Marcus, Matthew Stephens, and Xin He. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat. Genet.*, 52(7):740–747, 2020.
37. Ninon Mounier and Zoltán Kutalik. Bias correction for inverse variance weighting mendelian randomization. *Genetic Epidemiology*, 2023.
38. Paul J Newcombe, David V Conti, and Sylvia Richardson. Jam: a scalable bayesian framework for joint analysis of marginal snp effects. *Genetic epidemiology*, 40(3):188–201, 2016.
39. Tom M Palmer, John R Thompson, Martin D Tobin, Nuala A Sheehan, and Paul R Burton. Adjusting for bias and unmeasured confounding in mendelian randomization studies with binary responses. *International journal of epidemiology*, 37(5):1161–1168, 2008.
40. A Patel, D Gill, P Newcombe, and Burgess S. Conditional inference in cis-mendelian randomization using weak genetic factors. *Biometrics*, 2023.
41. Guanghao Qi and Nilanjan Chatterjee. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nature*

communications, 10(1):1941, 2019.

42. Guanghao Qi and Nilanjan Chatterjee. A comprehensive evaluation of methods for mendelian randomization using realistic simulations and an analysis of 38 biomarkers for risk of type 2 diabetes. *International journal of epidemiology*, 50(4):1335–1349, 2021.
43. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.
44. Jessica MB Rees, Angela M Wood, Frank Dudbridge, and Stephen Burgess. Robust methods in mendelian randomization via penalization of heterogeneous causal estimates. *PloS One*, 14(9):e0222362, 2019.
45. Rebecca C Richmond and George Davey Smith. Mendelian randomization: concepts and scope. *Cold Spring Harbor perspectives in medicine*, 12(1):a040501, 2022.
46. Eleanor Sanderson, George Davey Smith, Frank Windmeijer, and Jack Bowden. An examination of multivariable mendelian randomization in the single-sample and two-sample summary data settings. *International journal of epidemiology*, 48(3):713–727, 2019.
47. Eleanor Sanderson, M Maria Glymour, Michael V Holmes, Hyunseung Kang, Jean Morrison, Marcus R Munafò, Tom Palmer, C Mary Schooling, Chris Wallace, Qingyuan Zhao, et al. Mendelian randomization. *Nature Reviews Methods Primers*, 2(1):6, 2022.
48. Eleanor Sanderson, Tom G Richardson, Tim T Morris, Kate Tilling, and George Davey Smith. Estimation of causal effects of a time-varying exposure at multiple time points through multivariable mendelian randomization. *PLoS Genetics*, 18(7):e1010290, 2022.
49. Eleanor Sanderson, Wes Spiller, and Jack Bowden. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable mendelian randomization. *Statistics in medicine*, 40(25):5434–5452, 2021.
50. AF Schmidt and F Dudbridge. Mendelian randomization with egger pleiotropy correction and weakly informative bayesian priors. *International journal of epidemiology*, 47(4):1217–1228, 2018.
51. Eric AW Slob and Stephen Burgess. A comparison of robust mendelian randomization methods using summary data. *Genetic epidemiology*, 44(4):313–329, 2020.
52. Adriaan van Der Graaf, Annique Claringbould, Antoine Rimbert, BIOS Consortium Heijmans Bastiaan T. 8 Hoen Peter AC't 9 van Meurs Joyce BJ 10 Jansen Rick 11 Franke Lude 1 2, Harm-Jan Westra, Yang Li, Cisca Wijmenga, and Serena Sanna. Mendelian randomization while jointly modeling cis genetics identifies causal relationships between gene expression and lipids. *Nature communications*, 11(1):4930, 2020.
53. Marie Verbanck, Chia-Yen Chen, Benjamin Neale, and Ron Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nat. Genet.*, 50(5):693–698, 2018.
54. Jingshu Wang, Qingyuan Zhao, Jack Bowden, Gibran Hemani, George Davey Smith, Dylan S Small, and Nancy R Zhang. Causal inference for heritable

phenotypic risk factors using heterogeneous genetic instruments. *PLoS genetics*, 17(6):e1009575, 2021.

55. Lu Wang, Boran Gao, Yue Fan, Fuzhong Xue, and Xiang Zhou. Mendelian randomization under the omnigenic architecture. *Briefings in Bioinformatics*, 22(6):bbab322, 2021.
56. Siqi Xu, Wing Kam Fung, and Zhonghua Liu. Mrcip: a robust mendelian randomization method accounting for correlated and idiosyncratic pleiotropy. *Briefings in Bioinformatics*, 22(5):bbab019, 2021.
57. Siqi Xu, Peng Wang, Wing Kam Fung, and Zhonghua Liu. A novel penalized inverse-variance weighted estimator for mendelian randomization with applications to covid-19 outcomes. *Biometrics*, 2022.
58. Haoran Xue, Xiaotong Shen, and Wei Pan. Constrained maximum likelihood-based mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. *The American Journal of Human Genetics*, 108(7):1251–1269, 2021.
59. Jian Yang, Teresa Ferreira, Andrew P Morris, Sarah E Medland, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication, Meta analysis (DIAGRAM) Consortium, Pamela AF Madden, Andrew C Heath, Nicholas G Martin, Grant W Montgomery, et al. Conditional and joint multiple-snp analysis of gwas summary statistics identifies additional variants influencing complex traits. *Nature genetics*, 44(4):369–375, 2012.
60. Ting Ye, Jun Shao, and Hyunseung Kang. Debiased inverse-variance weighted estimator in two-sample summary-data mendelian randomization. *The Annals of statistics*, 49(4):2079–2100, 2021.
61. Zhongshang Yuan, Lu Liu, Ping Guo, Ran Yan, Fuzhong Xue, and Xiang Zhou. Likelihood-based mendelian randomization analysis with automated instrument selection and horizontal pleiotropic modeling. *Science Advances*, 8(9):eabl5744, 2022.
62. Jia Zhao, Jingsi Ming, Xianghong Hu, Gang Chen, Jin Liu, and Can Yang. Bayesian weighted mendelian randomization for causal inference based on summary statistics. *Bioinformatics*, 36(5):1501–1508, 2020.
63. Qingyuan Zhao, Jingshu Wang, Gibran Hemani, Jack Bowden, and Dylan S Small. Statistical inference in two-sample summary-data mendelian randomization using robust adjusted profile score. 2020.
64. Xiaofeng Zhu, Xiaoyin Li, Rong Xu, and Tao Wang. An iterative approach to detect pleiotropy and perform mendelian randomization analysis using gwas summary statistics. *Bioinformatics*, 37(10):1390–1400, 2021.
65. Verena Zuber, Johanna Maria Colijn, Caroline Klaver, and Stephen Burgess. Selecting likely causal risk factors from high-throughput experiments using multivariable mendelian randomization. *Nature communications*, 11(1):29, 2020.