

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

Noah Lorincz-Comi¹ | Yihe Yang¹ | Xiaofeng Zhu¹

¹Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH 44106, USA

Correspondence

Noah Lorincz-Comi, Department of Population and Quantitative Health Sciences, Case Western Reserve University
Email: njl96@case.edu

Present address

[†]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH 44106, USA

Funding information

National Human Genome Research Institute (NHGRI), Grant/Award Number: HG011052 and HG011052-03S1 (to XZ)

Mendelian Randomization (MR) has become a popular tool for inferring causality of risk factors on disease. There are currently over 45 different methods available to perform MR, reflecting this 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 MR methods in a range of scenarios encountered in practice. Researchers can directly modify the `simmr` source code so that the research community may arrive at a widely accepted framework for researchers to evaluate the performance of different MR methods.

KEYWORDS

Mendelian Randomization, bias, causal inference, instrumental variables, simulation, statistical genetics

1 | 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 (Sanderson et al., 2022a; Richmond and Smith, 2022). The validity of inferences made using MR relies heavily on the satisfaction of three primary assumptions (de Leeuw et al., 2022):

Abbreviations: MR, Mendelian Randomization; LD, linkage disequilibrium; UVMR, univariable Mendelian Randomization; MVMR, multivariable Mendelian Randomization; SNP, single nucleotide polymorphism; GWAS, genome-wide association study; DAG, directed acyclic graph

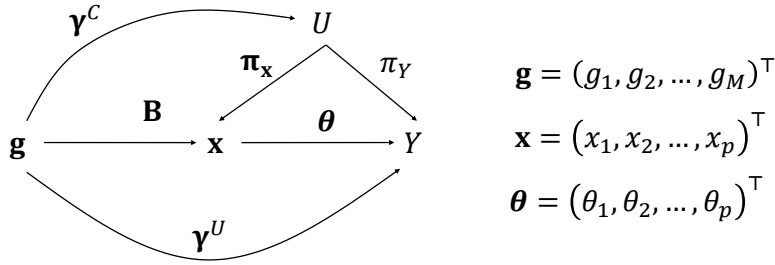


FIGURE 1 This DAG is used to produce the simulation models described in the text. Users of `simmr` can modify all parameters that are present in this DAG.

the genetic instruments are (i) strongly associated with the exposure(s) (Davies et al., 2015; Sanderson et al., 2021), (i) not associated with the outcome conditional on the exposures (Hemani et al., 2018), and (iii) not associated with any confounders of the exposure(s)-outcome relationship(s) (Morrison et al., 2020). Violation of any of these assumptions can lead to bias in causal estimation, therefore inflating false positive and/or false negative rates (Bowden et al., 2015). Three additional challenges are present in MR analyses: (a) some individuals may have been present in both the exposure(s) and outcome GWAS which leads to sample overlap bias (Burgess et al., 2016a), (b) applying strict IV selection criteria to satisfy assumption (i) above can introduce a ‘winner’s curse’ bias (Jiang et al., 2022b), and (c) strong LD between IVs (Gkatzionis et al., 2023), and its imprecise estimation using reference panels, can drastically inflate false positive rates (Gleason et al., 2021).

At least 45 statistical methods for performing MR 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 (Boehm and Zhou, 2022). Clearly, the literature is saturated with a variety of MR estimators. In each of the initial studies introducing these 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 phenotypic (Verbanck et al., 2018; Zhu et al., 2021) or genetic (Lin et al., 2023) variance while others have used hundreds (Lorincz-Comi et al., 2023) or even thousands (Mounier and Kutalik, 2023) of SNPs. Additionally, some have used GWAS sample sizes no larger than 500 (Deng et al., 2022) while others have used GWAS sample sizes no less than 100k (Xu et al., 2021). Nearly all multivariable MR simulation sets do not accurately reflect reality. For example, some have 50 or less IVs in simulation (Lin et al., 2023; Sanderson et al., 2019, 2022b), but experience with real data analyses with multiple exposures suggest that hundreds or thousands of SNPs meeting the valid IV conditions may be identified (Davies et al., 2019; Lorincz-Comi et al., 2023). Since weak IV bias generally becomes more severe as more IVs are used (Davies et al., 2015; Lorincz-Comi et al., 2023), 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 (Lin et al., 2023; Grant and Burgess, 2021), a scenario unlikely to be encountered in practice and one that does not require MVMR methods.

Some of these issues arise from researchers reproducing the unrealistic simulation settings of others as described above (e.g., (Zhu et al., 2021) reproduced (Verbanck et al., 2018); (Lin et al., 2023) reproduced (Grant and Burgess,

2021)). 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 ((Mounier and Kutalik, 2023) and (Ye et al., 2021)); the power of MR-Robust in the presence of 50% invalid IVs, being either less than 20% or greater than 80% ((Lin et al., 2023) and (Slob and Burgess, 2020)); or MR-RAPS having Type I error greater than 80% or controlled at 5% in the presence of horizontal pleiotropy ((Qi and Chatterjee, 2021) and (Xue et al., 2021)).

We introduce `simmr`, an open-source and flexible tool for generating data to use in MR simulations. The last time a tool such as this was introduced was `MR_predictor` 10 years ago (Voight, 2014), however `MR_predictor` cannot accommodate horizontal pleiotropy, sample overlap, multiple exposures, weak instruments, correlated IVs, imprecise LD estimation, or winner's curse. `simmr` can accommodate each of these and even more scenarios simultaneously in either the UVMR or MVMR settings. With no modern standard set in the literature, the purpose of the open-source `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.

2 | SIMULATION MODELS

The simulation data-generating process is based on the directed acyclic graph (DAG) in Figure 1 using the input parameters displayed in Figure 2. That is, simulation data are generated under the following models:

$$U = \mathbf{g}^T \boldsymbol{\gamma}^C + \epsilon_U, \quad (1)$$

$$\mathbf{x} = \mathbf{B}^T \mathbf{g} + \pi_X U + \epsilon_X, \quad (2)$$

$$Y = \mathbf{x}^T \boldsymbol{\theta} + \pi_Y U + \mathbf{g}^T \boldsymbol{\gamma}^U + \epsilon_Y \quad (3)$$

$$= \boldsymbol{\alpha}^T \mathbf{g} + \tilde{\epsilon}_Y \quad (4)$$

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 (Dudbridge and Newcombe, 2016). Users also have the option to perform IV pruning such that a specific F-statistic for instrument strength (Burgess et al., 2011) 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., (Lin et al., 2023), (Lorincz-Comi et al., 2023), (Cheng et al., 2020), (Mounier and Kutalik, 2023)). These methods calculate $\boldsymbol{\Omega}$ using non-significant GWAS summary statistics (Lorincz-Comi et al., 2023; Mounier and Kutalik, 2023; Zhu et al., 2015; Zhao et al., 2020b) or LD score regression (Lin et al., 2023; Cheng et al., 2020). `simmr` will directly calculate $\boldsymbol{\Omega}$ using the methods in (LeBlanc et al., 2018) without any estimation error and return it to the user.

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

Require: $\text{Var}(Y|x_k, U)$ for $k = 1, \dots, p$, $\text{sign}(\text{Cov}[Y, x_k])$, $\text{Cov}(x|U)$, $\text{Cov}(x|g)$, phenotypic and genetic correlation between exposures S_P and 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} , the number of individuals in the LD reference panel n_{ref} , 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 σ_{YU}^2 and σ_{YC}^2 , 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 $E(g_j) = 0$, $\text{Var}(g_j) = 1$; $j = 1, \dots, M$
2. Draw m^C CHP effects from $\gamma_j^C \sim \text{Uniform}(-1, 1)$ and re-scale to explain σ_{YC}^2 variance in U
3. Draw m^U UHP effects from $\gamma_j^U \sim \text{Uniform}(-1, 1)$ and re-scale to explain σ_{YU}^2 variance in Y
4. Draw un-transformed true SNP associations with the exposures ($\tilde{\mathbf{B}}$) 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 $[\text{diag}(\tilde{\mathbf{B}})]_k = 1 - [\text{diag}(\text{Cov}\{x|g\})]_k$
5. Transform $\tilde{\mathbf{B}}$ using the copula method to convert to a multivariate uniform distribution: $[\mathbf{B}]_{jk} = \Phi\left([\tilde{\mathbf{B}}\mathbf{S}_G^{-1/2}]_{jk}\right)$
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 $k = 1, \dots, p$
7. Define sets \mathcal{P}_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 $x_{\mathcal{P}_X}$ 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 S
11. Draw LD matrix for set of IVs \mathbf{R}_S from $\text{Wishart}(n_{\text{ref}}, \mathbf{R})$ and standardize to be a matrix of LD correlations

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

3 | SOFTWARE

`simmr` uses the popular R software v4.1.1 or later (R Core Team, 2021) 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.

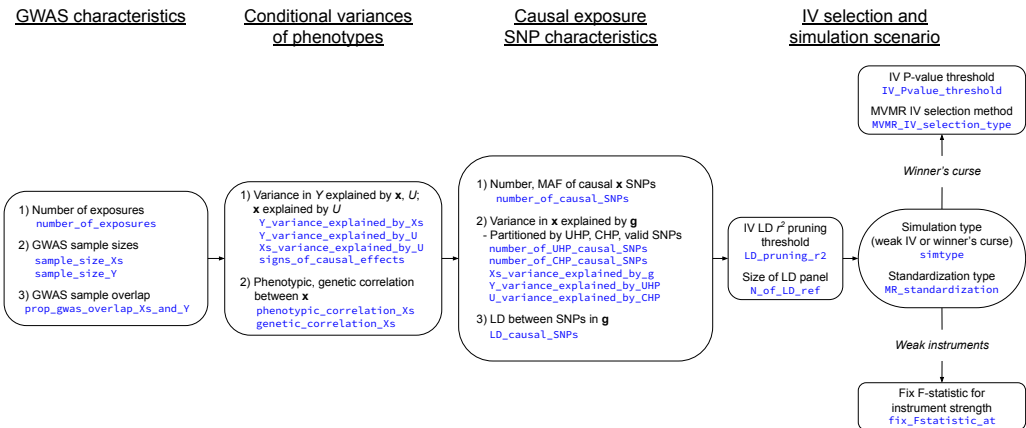


FIGURE 2 This figure provides a conceptual overview of the `simmr` data-generating process. First, users specify the sizes of exposure and outcome GWAS sample sizes and their degree of overlap. Next, users specify the true causal model by changing variances in each phenotype explained by each other and 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. Users can view a summary of the outputted data, in the format of the right panel in Figure 3, simply by executing `plot_simdata()` after `generate_data()`.

4 | 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: (i) standardized SNP association estimates between all IVs and each exposure in the R matrix 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 in `mIVs`, (vii) an indication for each IV if it was generated under a non-UHP/CHP, UHP, or CHP model in the vector `IVType`, (viii) the matrix of estimated LD correlations between all IVs in the matrix `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 the matrix `RhoME`. Users can view a summary visualization of their simulated data by executing `plot_simdata()` after `generate_data()`.

5 | CONCLUSION

There is currently no standard simulation framework for performing simulations in Mendelian Randomization research. Different researchers have independently performed simulations designed to reflect 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

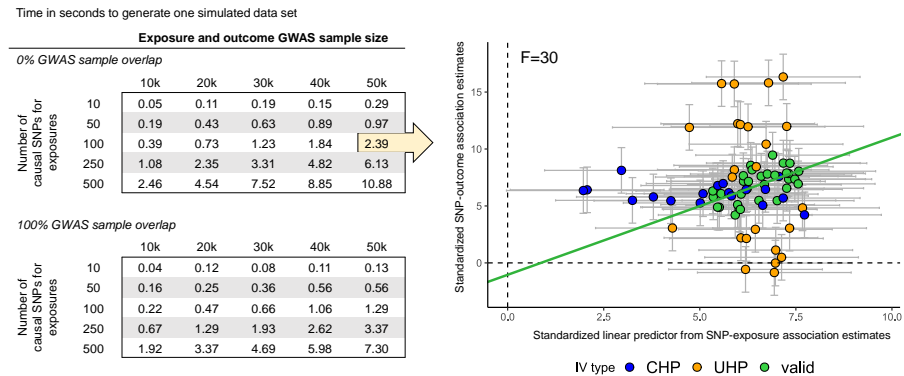


FIGURE 3 This figure shows computation times to generate one simulated data set using `simmr` (left) and an example of the generated summary data for evaluation of MR methods (right). The data in the right panel is only of the final selected and pruned set of instruments and was created simply by executing `plot_simdata()` in R. 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 ($\hat{B}\theta$) 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.

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` source code. 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 facilitate future MR method development and evaluation.

Acknowledgements

This work was supported by grant HG011052 and HG011052-03S1 (to X.Z.) from the National Human Genome Research Institute (NHGRI)

Conflict of interest

The authors declare that they have no competing interests.

6 | APPENDIX

We mentioned in the Introduction section that at least 45 statistical methods for performing Mendelian Randomization (MR) exist. These methods include IVW (Burgess et al., 2013), dIVW (Ye et al., 2021), pIVW (Xu et al., 2022), MR-Egger (Bowden et al., 2016), MR-RAPS (Zhao et al., 2020b), MRAID (Yuan et al., 2022), MRMix (Qi and Chatterjee, 2019), MR-cML (Xue et al., 2021), MVMR-cML (Lin et al., 2023), MR-PRESSO (Verbanck et al., 2018), IMRP (Zhu et al., 2021), MR-Median (Bowden et al., 2016), MR-MaxLike (Burgess et al., 2016b), MR-Corr (Cheng et al., 2022a), MR-Robust (Rees et al., 2019), MR-Lasso (Kang et al., 2016), MR-Conmix (Burgess et al., 2020), CAUSE (Morrison et al., 2020), MR-CUE (Cheng et al., 2022b), MR-Horse (Grant and Burgess, 2023), MR-BMA (Zuber et al., 2020), MR-Robin (Gleason et al., 2020), EMIC (Jiang et al., 2022a), MR-Mode (Hartwig et al., 2017), MRBEE (Lorincz-Comi et al., 2023), MR-Lap (Mounier and Kutalik, 2023), the Wald test (Palmer et al., 2008), JAM (Newcombe et al., 2016), MR using factor analysis (Patel et al., 2023), mixIE (Lin et al., 2021), MRMO (Deng et al., 2022), BMRMO (Deng et al., 2023), BWMR (Zhao et al., 2020a), moPMR-Egger (Liu et al., 2021), sisVIVE (Kang et al., 2016), MR-LDP (Cheng et al., 2020), MR-CIP (Xu et al., 2021), MR-PATH (long et al., 2020), MR-Clust (Foley et al., 2021), BayesMR (Bucur et al., 2020), BMRE (Schmidt and Dudbridge, 2018), MR-link (van Der Graaf et al., 2020), OMR (Wang et al., 2021b), CoJo (Yang et al., 2012), MR using PCA (Burgess et al., 2017), and GRAPPLE (Wang et al., 2021a).

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` (required) and `ggplot2` (optional) R packages
License: MIT

references

Boehm, F. J. and Zhou, X. (2022) Statistical methods for mendelian randomization in genome-wide association studies: A review. *Computational and Structural Biotechnology Journal*, **20**, 2338–2351.

Bowden, J., Davey Smith, G. and Burgess, S. (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *International journal of epidemiology*, **44**, 512–525.

Bowden, J., Davey Smith, G., Haycock, P. C. and Burgess, S. (2016) Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.*, **40**, 304–314.

Bucur, I. G., Claassen, T. and Heskes, T. (2020) Inferring the direction of a causal link and estimating its effect via a bayesian mendelian randomization approach. *Statistical Methods in Medical Research*, **29**, 1081–1111.

Burgess, S., Butterworth, A. and Thompson, S. G. (2013) Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology*, **37**, 658–665.

Burgess, S., Davies, N. M. and Thompson, S. G. (2016a) Bias due to participant overlap in two-sample mendelian randomization. *Genetic epidemiology*, **40**, 597–608.

- 141 Burgess, S., Dudbridge, F. and Thompson, S. G. (2016b) Combining information on multiple instrumental variables in mendelian
142 randomization: comparison of allele score and summarized data methods. *Statistics in medicine*, **35**, 1880–1906.
- 143 Burgess, S., Foley, C. N., Allara, E., Staley, J. R. and Howson, J. M. (2020) A robust and efficient method for mendelian random-
144 ization with hundreds of genetic variants. *Nat. Commun.*, **11**, 1–11.
- 145 Burgess, S., Thompson, S. G. and Collaboration, C. C. G. (2011) Avoiding bias from weak instruments in mendelian randomiza-
146 tion studies. *International journal of epidemiology*, **40**, 755–764.
- 147 Burgess, S., Zuber, V., Valdes-Marquez, E., Sun, B. B. and Hopewell, J. C. (2017) Mendelian randomization with fine-mapped
148 genetic data: choosing from large numbers of correlated instrumental variables. *Genetic epidemiology*, **41**, 714–725.
- 149 Cheng, Q., Qiu, T., Chai, X., Sun, B., Xia, Y., Shi, X. and Liu, J. (2022a) Mr-corr2: a two-sample mendelian randomization method
150 that accounts for correlated horizontal pleiotropy using correlated instrumental variants. *Bioinformatics*, **38**, 303–310.
- 151 Cheng, Q., Yang, Y., Shi, X., Yeung, K.-F., Yang, C., Peng, H. and Liu, J. (2020) Mr-ldp: a two-sample mendelian randomization for
152 gwas summary statistics accounting for linkage disequilibrium and horizontal pleiotropy. *NAR genomics and bioinformatics*,
153 **2**, lqaa028.
- 154 Cheng, Q., Zhang, X., Chen, L. S. and Liu, J. (2022b) Mendelian randomization accounting for complex correlated horizontal
155 pleiotropy while elucidating shared genetic etiology. *Nat. Commun.*, **13**, 1–13.
- 156 Davies, N. M., Hill, W. D., Anderson, E. L., Sanderson, E., Deary, I. J. and Davey Smith, G. (2019) Multivariable two-sample
157 mendelian randomization estimates of the effects of intelligence and education on health. *Elife*, **8**, e43990.
- 158 Davies, N. M., von Hinke Kessler Scholder, S., Farbmacher, H., Burgess, S., Windmeijer, F. and Smith, G. D. (2015) The many
159 weak instruments problem and mendelian randomization. *Statistics in medicine*, **34**, 454–468.
- 160 Deng, Y., Tu, D., O'Callaghan, C. J., Jonker, D. J., Karapetis, C. S., Shapiro, J., Liu, G. and Xu, W. (2023) A bayesian approach for
161 two-stage multivariate mendelian randomization with mixed outcomes. *Statistics in Medicine*.
- 162 Deng, Y., Tu, D., O'Callaghan, C. J., Liu, G. and Xu, W. (2022) Two-stage multivariate mendelian randomization on multiple
163 outcomes with mixed distributions. *Statistical Methods in Medical Research*, 09622802231181220.
- 164 van Der Graaf, A., Claringbould, A., Rimbart, A., van Meurs Joyce BJ 10 Jansen Rick 11 Franke Lude 1 2, B. C. H. B. T. . H. P.
165 A. ., Westra, H.-J., Li, Y., Wijmenga, C. and Sanna, S. (2020) Mendelian randomization while jointly modeling cis genetics
166 identifies causal relationships between gene expression and lipids. *Nature communications*, **11**, 4930.
- 167 Dudbridge, F. and Newcombe, P. J. (2016) Accuracy of gene scores when pruning markers by linkage disequilibrium. *Human
168 heredity*, **80**, 178–186.
- 169 Foley, C. N., Mason, A. M., Kirk, P. D. and Burgess, S. (2021) Mr-clust: clustering of genetic variants in mendelian randomization
170 with similar causal estimates. *Bioinformatics*, **37**, 531–541.
- 171 Gkatzionis, A., Burgess, S. and Newcombe, P. J. (2023) Statistical methods for cis-mendelian randomization with two-sample
172 summary-level data. *Genetic epidemiology*, **47**, 3–25.
- 173 Gleason, K. J., Yang, F. and Chen, L. S. (2020) A robust two-sample mendelian randomization method integrating gwas with
174 multi-tissue eqtl summary statistics. *bioRxiv*, 2020–06.
- 175 – (2021) A robust two-sample transcriptome-wide mendelian randomization method integrating gwas with multi-tissue eqtl
176 summary statistics. *Genetic epidemiology*, **45**, 353–371.
- 177 Grant, A. J. and Burgess, S. (2021) Pleiotropy robust methods for multivariable mendelian randomization. *Statistics in medicine*,
178 **40**, 5813–5830.

- (2023) A bayesian approach to mendelian randomization using summary statistics in the univariable and multivariable settings with correlated pleiotropy. *bioRxiv*, 2023–05.
- Hartwig, F. P., Davey Smith, G. and Bowden, J. (2017) Robust inference in summary data mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology*, **46**, 1985–1998.
- Hemani, G., Bowden, J. and Davey Smith, G. (2018) Evaluating the potential role of pleiotropy in mendelian randomization studies. *Human molecular genetics*, **27**, R195–R208.
- long, D., Zhao, Q. and Chen, Y. (2020) A latent mixture model for heterogeneous causal mechanisms in mendelian randomization. *arXiv preprint arXiv:2007.06476*.
- Jiang, L., Miao, L., Yi, G., Li, X., Xue, C., Li, M. J., Huang, H. and Li, M. (2022a) 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**, 838–856.
- Jiang, T., Gill, D., Butterworth, A. S. and Burgess, S. (2022b) An empirical investigation into the impact of winner's curse on estimates from mendelian randomization. *medRxiv*, 2022–08.
- Kang, H., Zhang, A., Cai, T. T. and Small, D. S. (2016) Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. *J. Am. Stat. Assoc.*, **111**, 132–144.
- LeBlanc, M., Zuber, V., Thompson, W. K., Andreassen, O. A., Schizophrenia, of the Psychiatric Genomics Consortium, B. D. W. G., Frigessi, A. and Andreassen, B. K. (2018) A correction for sample overlap in genome-wide association studies in a polygenic pleiotropy-informed framework. *BMC genomics*, **19**, 1–15.
- de Leeuw, C., Savage, J., Bucur, I. G., Heskes, T. and Posthuma, D. (2022) Understanding the assumptions underlying mendelian randomization. *European Journal of Human Genetics*, **30**, 653–660.
- Lin, Z., Deng, Y. and Pan, W. (2021) Combining the strengths of inverse-variance weighting and egger regression in mendelian randomization using a mixture of regressions model. *PLoS genetics*, **17**, e1009922.
- Lin, Z., Xue, H. and Pan, W. (2023) Robust multivariable mendelian randomization based on constrained maximum likelihood. *The American Journal of Human Genetics*, **110**, 592–605.
- Liu, L., Zeng, P., Xue, F., Yuan, Z. and Zhou, X. (2021) Multi-trait transcriptome-wide association studies with probabilistic mendelian randomization. *The American Journal of Human Genetics*, **108**, 240–256.
- Lorincz-Comi, N., Yang, Y., Li, G. and Zhu, X. (2023) Mrbee: A novel bias-corrected multivariable mendelian randomization method. *bioRxiv*, 2023–01.
- Morrison, J., Knoblach, N., Marcus, J. H., Stephens, M. and He, X. (2020) Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat. Genet.*, **52**, 740–747.
- Mounier, N. and Kutalik, Z. (2023) Bias correction for inverse variance weighting mendelian randomization. *Genetic Epidemiology*.
- Newcombe, P. J., Conti, D. V. and Richardson, S. (2016) Jam: a scalable bayesian framework for joint analysis of marginal snp effects. *Genetic epidemiology*, **40**, 188–201.
- Palmer, T. M., Thompson, J. R., Tobin, M. D., Sheehan, N. A. and Burton, P. R. (2008) Adjusting for bias and unmeasured confounding in mendelian randomization studies with binary responses. *International journal of epidemiology*, **37**, 1161–1168.
- Patel, A., Gill, D., Newcombe, P. and S, B. (2023) Conditional inference in cis-mendelian randomization using weak genetic factors. *Biometrics*.

- Qi, G. and Chatterjee, N. (2019) Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nature communications*, **10**, 1941.
- (2021) 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**, 1335–1349.
- R Core Team (2021) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rees, J. M., Wood, A. M., Dudbridge, F. and Burgess, S. (2019) Robust methods in mendelian randomization via penalization of heterogeneous causal estimates. *PloS One*, **14**, e0222362.
- Richmond, R. C. and Smith, G. D. (2022) Mendelian randomization: concepts and scope. *Cold Spring Harbor perspectives in medicine*, **12**, a040501.
- Sanderson, E., Davey Smith, G., Windmeijer, F. and Bowden, J. (2019) An examination of multivariable mendelian randomization in the single-sample and two-sample summary data settings. *International journal of epidemiology*, **48**, 713–727.
- Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., Palmer, T., Schooling, C. M., Wallace, C., Zhao, Q. et al. (2022a) Mendelian randomization. *Nature Reviews Methods Primers*, **2**, 6.
- Sanderson, E., Richardson, T. G., Morris, T. T., Tilling, K. and Davey Smith, G. (2022b) Estimation of causal effects of a time-varying exposure at multiple time points through multivariable mendelian randomization. *PLoS Genetics*, **18**, e1010290.
- Sanderson, E., Spiller, W. and Bowden, J. (2021) Testing and correcting for weak and pleiotropic instruments in two-sample multivariable mendelian randomization. *Statistics in medicine*, **40**, 5434–5452.
- Schmidt, A. and Dudbridge, F. (2018) Mendelian randomization with egger pleiotropy correction and weakly informative bayesian priors. *International journal of epidemiology*, **47**, 1217–1228.
- Slob, E. A. and Burgess, S. (2020) A comparison of robust mendelian randomization methods using summary data. *Genetic epidemiology*, **44**, 313–329.
- Verbanck, M., Chen, C.-Y., Neale, B. and Do, R. (2018) Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nat. Genet.*, **50**, 693–698.
- Voight, B. F. (2014) Mr_predictor: a simulation engine for mendelian randomization studies. *Bioinformatics*, **30**, 3432–3434.
- Wang, J., Zhao, Q., Bowden, J., Hemani, G., Davey Smith, G., Small, D. S. and Zhang, N. R. (2021a) Causal inference for heritable phenotypic risk factors using heterogeneous genetic instruments. *PLoS genetics*, **17**, e1009575.
- Wang, L., Gao, B., Fan, Y., Xue, F. and Zhou, X. (2021b) Mendelian randomization under the omnigenic architecture. *Briefings in Bioinformatics*, **22**, bbab322.
- Xu, S., Fung, W. K. and Liu, Z. (2021) Mrcip: a robust mendelian randomization method accounting for correlated and idiosyncratic pleiotropy. *Briefings in Bioinformatics*, **22**, bbab019.
- Xu, S., Wang, P., Fung, W. K. and Liu, Z. (2022) A novel penalized inverse-variance weighted estimator for mendelian randomization with applications to covid-19 outcomes. *Biometrics*.
- Xue, H., Shen, X. and Pan, W. (2021) Constrained maximum likelihood-based mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. *The American Journal of Human Genetics*, **108**, 1251–1269.
- Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., of ANthropometric Traits (GIANT) Consortium, G. I., Replication, D. G., analysis (DIAGRAM) Consortium, M., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W. et al. (2012) Conditional and joint multiple-snp analysis of gwas summary statistics identifies additional variants influencing complex traits. *Nature genetics*, **44**, 369–375.

- 257 Ye, T., Shao, J. and Kang, H. (2021) Debiased inverse-variance weighted estimator in two-sample summary-data mendelian
258 randomization. *The Annals of statistics*, **49**, 2079–2100.
- 259 Yuan, Z., Liu, L., Guo, P., Yan, R., Xue, F. and Zhou, X. (2022) Likelihood-based mendelian randomization analysis with auto-
260 mated instrument selection and horizontal pleiotropic modeling. *Science Advances*, **8**, eabl5744.
- 261 Zhao, J., Ming, J., Hu, X., Chen, G., Liu, J. and Yang, C. (2020a) Bayesian weighted mendelian randomization for causal inference
262 based on summary statistics. *Bioinformatics*, **36**, 1501–1508.
- 263 Zhao, Q., Wang, J., Hemani, G., Bowden, J. and Small, D. S. (2020b) Statistical inference in two-sample summary-data
264 mendelian randomization using robust adjusted profile score.
- 265 Zhu, X., Feng, T., Tayo, B. O., Liang, J., Young, J. H., Franceschini, N., Smith, J. A., Yanek, L. R., Sun, Y. V., Edwards, T. L. et al.
266 (2015) Meta-analysis of correlated traits via summary statistics from gwas with an application in hypertension. *The*
267 *American Journal of Human Genetics*, **96**, 21–36.
- 268 Zhu, X., Li, X., Xu, R. and Wang, T. (2021) An iterative approach to detect pleiotropy and perform mendelian randomization
269 analysis using gwas summary statistics. *Bioinformatics*, **37**, 1390–1400.
- 270 Zuber, V., Colijn, J. M., Klaver, C. and Burgess, S. (2020) Selecting likely causal risk factors from high-throughput experiments
271 using multivariable mendelian randomization. *Nature communications*, **11**, 29.