Human Genetics
https://doi.org/10.1007/s00439-024-02640-x

ORIGINAL INVESTIGATION

Bayesian network-based Mendelian randomization for variant prioritization and phenotypic causal inference

Jianle Sun[1] · Jie Zhou[1] · Yuqiao Gong[1] · Chongchen Pang[1] · Yanran Ma[1] · Jian Zhao[2,3] · Zhangsheng Yu[1] · Yue Zhang[1]

Received: 14 November 2023 / Accepted: 5 January 2024

Abstract

Mendelian randomization is a powerful method for inferring causal relationships. However, obtaining suitable genetic instrumental variables is often challenging due to gene interaction, linkage, and pleiotropy. We propose Bayesian network-based Mendelian randomization (BNMR), a Bayesian causal learning and inference framework using individual-level data. BNMR employs the random graph forest, an ensemble Bayesian network structural learning process, to prioritize candidate genetic variants and select appropriate instrumental variables, and then obtains a pleiotropy-robust estimate by incorporating a shrinkage prior in the Bayesian framework. Simulations demonstrate BNMR can efficiently reduce the false-positive discoveries in variant selection, and outperforms existing MR methods in terms of accuracy and statistical power in effect estimation. With application to the UK Biobank, BNMR exhibits its capacity in handling modern genomic data, and reveals the causal relationships from hematological traits to blood pressures and psychiatric disorders. Its effectiveness in handling complex genetic structures and modern genomic data highlights the potential to facilitate real-world evidence studies, making it a promising tool for advancing our understanding of causal mechanisms.

Introduction

Identifying genuine causality is crucial to understanding physiological processes and discovering therapeutic targets, but it is also a tricky issue. Randomized controlled trials (RCTs) are usually regarded as the golden standard for causal inference but are restricted due to methodological, ethical, and economic concerns. Mendelian randomization (MR) is a promising approach to estimating causal effects using genetic variants as instrumental variables (IVs) (Sanderson et al. 2022). In general, MR analysis relies on three core assumptions (Fig. 1a): (i) relevance: a reliable correlation exists between the instrument and exposure; (ii) exogeneity or exchangeability: the instrument is independent with any confounders between the exposure and outcome ($Z \perp\!\!\!\perp U$); and (iii) exclusion restriction: the instrument should affect the outcome only through the exposure ($Z \perp\!\!\!\perp Y|X, U$).

Unfortunately, the rigorous assumptions are often violated (Fig. 1b), making it challenging to identify appropriate genetic instruments. First, although genome-wide association studies (GWAS) have identified numerous risk loci, in particular single nucleotide polymorphisms (SNPs), the effect on a polygenic complex trait is usually small, leading to weak-instrument bias (Davies et al. 2015). The multiple

✉ Zhangsheng Yu
yuzhangsheng@sjtu.edu.cn

✉ Yue Zhang
yue.zhang@sjtu.edu.cn

Jianle Sun
sjl-2017@sjtu.edu.cn

Jie Zhou
jie.zhou@sjtu.edu.cn

Yuqiao Gong
gyq123@sjtu.edu.cn

Chongchen Pang
pangchongchen@sjtu.edu.cn

Yanran Ma
qiyumyr@sjtu.edu.cn

Jian Zhao
zhaoj@sustech.edu.cn

[1] Department of Bioinformatics and Biostatistics, Shanghai Jiao Tong University, Shanghai, China

[2] School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, China

[3] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Published online: 21 February 2024

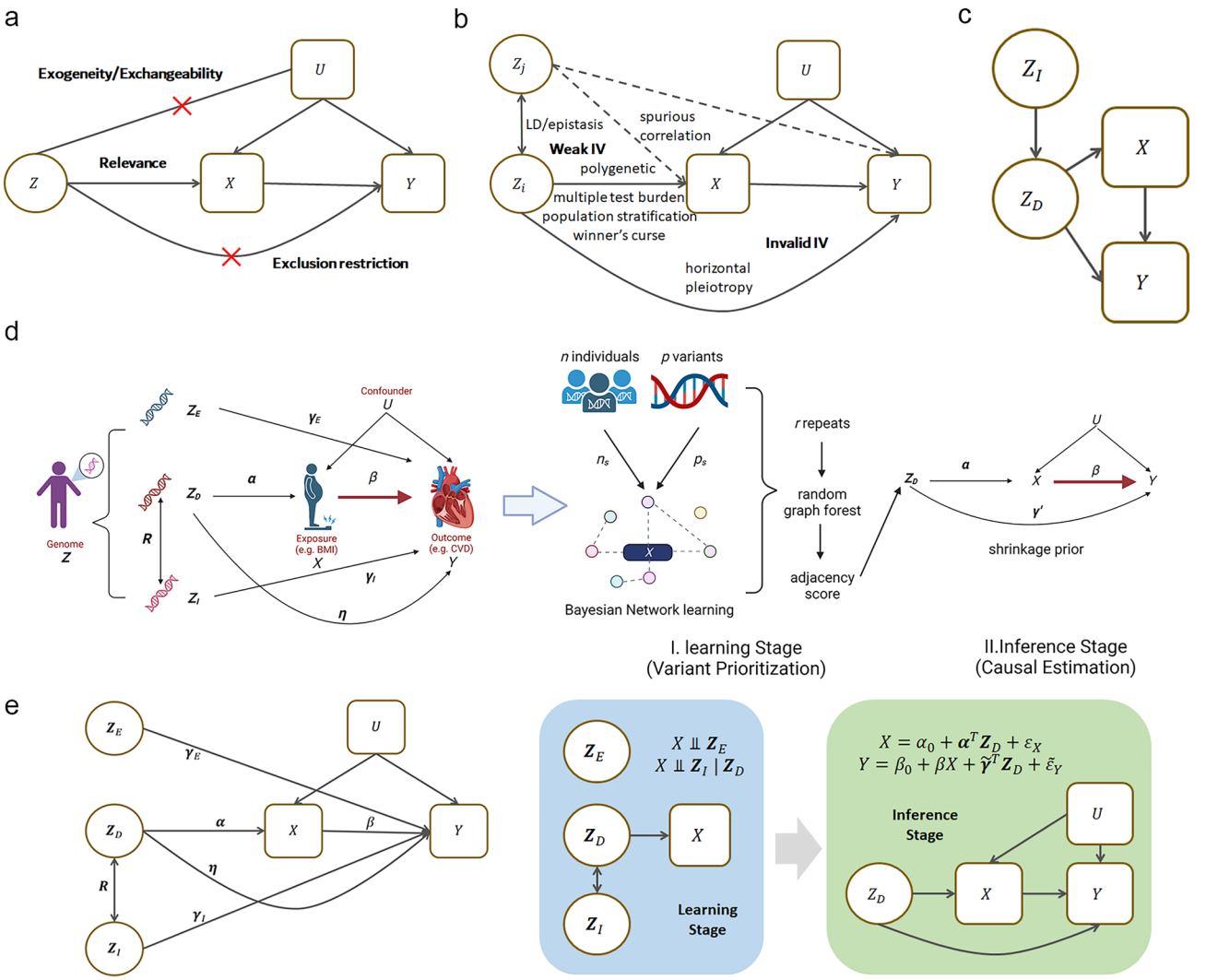


Fig. 1 The overview of BNMR. **a** The three core assumptions of IV. **b** The problems in current MR. Weak IVs are primarily due to the small individual contribution of a single locus to the trait and the low statistical power of GWAS, as well as the presence of linkage and interaction effects, leading to numerous false-positive discoveries. Invalid IVs are mainly caused by horizontal pleiotropy and linkage disequilibrium, which break the exclusion restriction assumption. **c** Correlated horizontal pleiotropy induced by gene interactions. Z_D is a pleiotropic variants with independent effect on X and Y, and Z_I is associated with Z_D. The causal pathway IV$\rightarrow X \rightarrow Y$ and horizontal pleiotropic pathway IV$\rightarrow Y$ will be correlated when Z_I is selected as IV. **d** The BNMR model. In the learning stage, we leverage the random graph forest to prioritize variants from a large interacting set and select variants with a true effect on the exposure as instruments. In the inference stage, we impose shrinkage prior on the Bayesian MR model to obtain a pleiotropy-robust estimate. **e** Notations used in the BNMR model

testing burdens, 'winner's curse', linkage disequilibrium (LD), and population stratification increase the risk of false-positive signals in GWAS (Tam et al. 2019). It can be improved by applying multiple instruments (Dudbridge 2021), whereas correlated instruments will also lead to unstable estimates and introduce additional genetic confoundings when including non-causal variants (Gkatzionis et al. 2023). Proposed strategies such as LD stepwise pruning (Yang et al. 2012), principal components analysis (PCA) (Burgess et al. 2017), and penalization aim to extract a suitable number of independent instruments from a large set of correlated weak variants, but confront criticism on robustness (Gkatzionis et al. 2022).

Another problem is that many IVs are actually invalid due to horizontal pleiotropy (a variant affects the outcome via alternative pathways other than the exposure of interest). Gene interactions, such as LD and epistasis, can also violate exclusion restrictions analogous to pleiotropy. For individual-level data, lasso-type methods like sisVIVE (Kang et al. 2016) and post-adaptive Lasso (Windmeijer et al. 2019) help to control the influence of the pleiotropic effect. Recent approaches such as TSHT (Guo et al. 2018)

and CIIV (Windmeijer et al. 2021) mitigate pleiotropy by identifying valid instruments from candidate sets.

The above approaches relying on many implausible assumptions are tricky to model sophisticated real genetic patterns. In particular, due to complex gene interactions (GxG), incorporating non-causal variants as IVs not only leads to unstable estimates and impacts statistical power but also may introduce GxG-induced correlated pleiotropy, violating the Instrument strength independent of direct effect (InSIDE) assumption that many methods for correcting pleiotropy rely on (Fig. 1c). Causal diagram model provides an alternative way to represent the underlying causal relationships (Nogueira et al. 2022). With causal diagrams, machine learning techniques like the causal Bayesian network (BN) are currently applied to identify genetic interactions and causal variants (Lyu et al. 2021). They will also be a profitable complement to conventional MR (Howey et al. 2020; Amar et al. 2021).

In this paper, we propose a two-stage Bayesian network-based Mendelian randomization (BNMR) approach by integrating causal discovery and inference (Fig. 1d). We aim to tackle correlated weak instruments in learning stage and cope with pleiotropy in inference stage. Using the random graph forest (RGF), an ensemble approach comprised of a series of BN structure learning processes, we prioritize variants with effects that are small and interacting and identify variants with direct effect on exposure as instruments. Then we estimate the causal effects via the Bayesian MR framework with a shrinkage prior to cope with potential horizontal pleiotropy (Berzuini et al. 2020). We demonstrate that BNMR is superior to conventional approaches in both instrument selection and effect estimation via simulations. With application to the UK BioBank, we examine causal pathways from hematological parameters to blood pressures and psychiatric disorders, bringing new biological insights.

Methods

Overview of the BNMR model

BNMR is a two-stage MR framework using individual-level data. In the learning stage, we propose RGF to select variants with reliable relevance from a large number of correlated weak instruments. We utilize BNs to characterize the complex conditional probability relationships and partition the variant set $\mathcal{Z}$ into three subsets according to their relationships with the exposure of interest X (DIE partition),

$$\mathcal{Z} = \mathcal{Z}_D \cup \mathcal{Z}_I \cup \mathcal{Z}_E. \tag{1}$$

We use notations in calligraphic font to represent the variant set, bold font to represent the vector of genotypes, and Italic capital letter to represent single genotype. Variants in $\mathcal{Z}_D$ directly affect the exposure, variants in $\mathcal{Z}_I$ indirectly affect the exposure via gene interaction or linkage, i.e., variants in $\mathcal{Z}_I$ and X are d-separated by variants in $\mathcal{Z}_D$ ($\boldsymbol{Z}_I \perp\!\!\!\perp X|\boldsymbol{Z}_D$), while variants in $\mathcal{Z}_E$ do not affect the exposure ($\boldsymbol{Z}_E \perp\!\!\!\perp X$). The three subsets are distinguished via BN under the causal Markov, faithfulness, and sufficiency assumptions (Nogueira et al. 2022), and only $\mathcal{Z}_D$ can be parents of X in the causal graph.

In the inference stage, we model the potential horizontal pleiotropy explicitly. Since quantitative traits are determined by both genetic and environmental factors, assuming linearity and no interaction, we have

$$X = \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{Z}_D + \varepsilon_X, \tag{2}$$

and

$$Y = \gamma_0 + \boldsymbol{\gamma}_D^T \boldsymbol{Z}_D + \boldsymbol{\gamma}_I^T \boldsymbol{Z}_I + \boldsymbol{\gamma}_E^T \boldsymbol{Z}_E + \varepsilon_Y, \tag{3}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ represent corresponding effect size on X and Y. Variants $Z_D \in \mathcal{Z}_D$ affect the outcome Y through two different pathways: with the mediation of exposure X ($\boldsymbol{Z}_D \xrightarrow{\boldsymbol{\alpha}} X \xrightarrow{\beta} Y$), the causal pathway of interest, and via direct pathway or through other mediators other than X ($\boldsymbol{Z}_D \xrightarrow{\boldsymbol{\eta}} Y$), known as (horizontal) pleiotropy (Pingault et al. 2018). Under the assumption that both pathways are independent (the InSIDE assumption) (Pingault et al. 2018), we have

$$\boldsymbol{\gamma}_D = \beta\boldsymbol{\alpha} + \boldsymbol{\eta}, \tag{4}$$

where β is the causal effect of X on Y, while $\boldsymbol{\eta}$ represents the pleiotropic effects. By introducing the correlation matrix $\boldsymbol{R}$ between $\boldsymbol{Z}_D$ and $\boldsymbol{Z}_I$,

$$\boldsymbol{Z}_I = \boldsymbol{R}_0 + \boldsymbol{R}\boldsymbol{Z}_D + \boldsymbol{\varepsilon}_R, \tag{5}$$

we can rewrite Eqs. (2) and (3) as

$$\begin{aligned} X &= \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{Z}_D + \varepsilon_X, \\ Y &= \beta_0 + \beta X + \widetilde{\boldsymbol{\gamma}}^T \boldsymbol{Z}_D + \widetilde{\varepsilon}_Y, \end{aligned} \tag{6}$$

where $\beta_0 = \gamma_0 + \boldsymbol{\gamma}_I^T \boldsymbol{R}_0 + \boldsymbol{\gamma}_E^T \boldsymbol{Z}_E - \beta\alpha_0$, $\widetilde{\boldsymbol{\gamma}} = \boldsymbol{\eta} + \boldsymbol{R}\boldsymbol{\gamma}_I$, and $\widetilde{\varepsilon}_Y = \boldsymbol{\gamma}_I^T \boldsymbol{\varepsilon}_R + \varepsilon_Y - \beta\varepsilon_X$. ε_X and $\widetilde{\varepsilon}_Y$ are correlated, but are both independent with $\boldsymbol{Z}_D$. Actually, only a subset of variants in $\mathcal{Z}_D$ needs to be included as IVs, and consequently, variants in $\mathcal{Z}_D$ are required to be identified with high precision. We then impose a shrinkage prior on nuisance parameters $\widetilde{\boldsymbol{\gamma}}$ to make β identifiable (Berzuini et al. 2020). Details on the derivation can be found in Supplementary Note SN1.

BN structure learning in the random graph forest

To reduce the computational complexity of structure learning and assess confidence of each edge, we propose RGF, inspired by the random forest. In RGF, r sub-graphs are created using

bootstrapping or subsampling, in which n_s of n individuals and p_s of p variants are sampled in each sub-graph. Consequently, we boil down the process of DIE partitioning to the structure learning of a series of causal BNs.

Since variants are reasons of traits naturally, we can simplify structure learning to graph skeleton determination. We identify $\mathcal{Z}_D$ by scanning the variants directly adjacent to the exposure in each graph and calculating the adjacency score (the frequency of the presence of $Z - X$ edge in all sub-graphs) for each variant, which is the confidence of the variant-exposure relevance in the average causal graph. Variants with higher adjacency scores are at a higher confidence level to be identified as $Z_D \in \mathcal{Z}_D$. We can select a specified number of lead variants or variants with scores higher than a given threshold $\frac{\alpha^* p_s}{p}$ as IVs.

Various algorithms are proposed for BN structure learning. Scored-based approaches ascertain the optimal network by exhaustively or heuristically exploring candidate graphs and maximizing the network score, while constraint-based approaches leverage a sequence of conditional independence tests to establish the edge constraints between nodes and subsequently refine the directions (Nogueira et al. 2022). We implement score-based approaches including Hill-Climbing (hc) and Tabu Search (tabu), constrained-based approaches including stable PC (pc.stable), Incremental Association (iamb), and Grow-Shrink (gs), as well as hybrid learning methods including Max-Min Hill-Climbing (mmhc) and Restricted Maximization (rsmax2). All these methods are implemented using the R package bnlearn.

Bayesian MR estimation with a shrinkage prior

We specify model (6) under the Bayesian framework. The total error term ε_X and $\widetilde{\varepsilon}_Y$ can be decomposed into a confounding-related term δ and a completely random term σ, i.e., $\mathbb{E}(\varepsilon_X) = \mathbb{E}(\widetilde{\varepsilon}_Y) = 0$, $\mathrm{Var}(\varepsilon_X) = \delta_1^2 + \sigma_1^2$, and $\mathrm{Var}(\widetilde{\varepsilon}_Y) = \delta_2^2 + \sigma_2^2$. Assuming that the two completely random terms are uncorrelated, we have $\mathrm{Cov}(\varepsilon_X, \widetilde{\varepsilon}_Y) = \delta_1\delta_2$.

We only need to select a subset of $\mathcal{Z}_D$ as instruments, and have the Bayesian MR model (Berzuini et al. 2020)

$$\begin{aligned} X|Z,U &\sim \mathcal{N}\left(\alpha_0 + \sum_{j=1}^{J}\alpha_j Z_j + \delta_1 U, \sigma_1^2\right) \\ Y|X,Z,U &\sim \mathcal{N}\left(\beta_0 + \beta X + \sum_{j=1}^{J}\gamma_j Z_j + \delta_2 U, \sigma_2^2\right), \\ U &\sim \mathcal{N}(0,1). \end{aligned} \tag{7}$$

where $Z_j \in \mathcal{Z}_D$. To make causal parameter β identifiable, we assume that not all IVs selected take pleiotropic effects (i.e., some components of $\boldsymbol{\gamma}$ are zero) and impose a shrinkage prior on $\boldsymbol{\gamma}$ under the Bayesian framework (Berzuini et al. 2020). The Bayesian estimation is conducted using Markov Chain Monte-Carlo (MCMC) with Rstan and PyMC. The first half of the iteration is used for burn-in, and the second half is used for sampling.

BNMR can be extended to binary outcomes by modifying the Eq. 7 to probit or logistic regressions, i.e.,

$$Y|X,Z,U \sim \text{Bernoulli}\left(h(\beta_0 + \beta X + \sum_{j=1}^{J}\gamma_j Z_j + \delta_2 U)\right), \tag{8}$$

where the link function $h(\cdot)$ can be inverse-probit or inverse-logit.

We compare estimates of BNMR with other IV selection and MR estimation approaches. We implement PCA with R package stats and penalized regressions with R package glmnet, where 10-fold cross validation is used to determine the best value of λ. Compared methods are implemented with the R packages AER, ivmodel, MendelianRandomization, cause, R2BGLiMS, and CIIV. We implement BNMR as an R package, with source codes available at https://github.com/sjl-sjtu/bnmr2.

Simulations

We use both simulated and real genomics from UK Biobank in simulations. For simulated genomics, k independent loci sampled from multinomial distributions, whose genotype frequencies satisfy the Hardy–Weinberg equilibrium (HWE), with the effect allele frequency π from $U(0.05, 0.95)$. m correlated loci for each locus are simulated according to LD squared correlation coefficient (r^2) (Pritchard and Przeworski 2001) that sampled from $U(0.01, 0.99)$, and genomics with $p = k(m+1)$ loci are synthesized. Real genomic data used to simulate phenotypes are derived from variants on chromosomes 10, 17 and 22 in the European ancestry population of UK Biobank.

Phenotypes are generated from linear model

$$X = \alpha_0 + \sum_j \alpha_j G_j + \delta_x U + \varepsilon_x,\ \varepsilon_x \sim \mathcal{N}(0, \sigma_x^2), \tag{9}$$

and

$$Y = \beta_0 + \beta X + \sum_{\substack{k \\ \text{unilateral}}} \gamma_k G_k + \sum_{\substack{m \\ \text{pleiotropic}}} \gamma_m G_m + \delta_y U + \varepsilon_y,\ \varepsilon_y \sim \mathcal{N}(0, \sigma_y^2), \tag{10}$$

with causal effect $\beta = 0.5$, G_m from a subset of G_j. Confounder U is generated from the standard Gaussian distribution $\mathcal{N}(0,1)$, with coefficients $\delta_x = \delta_y = 1$. Variants affecting X are randomly selected from the simulated genome, with effect size $\alpha_j \sim 0.1 + |\mathcal{N}(0, 0.05^2)|$. Variants affecting Y are either unilateral (those only affect Y) or pleiotropic (those that also affect X). For unilateral variants $\gamma_k \sim \mathcal{N}(0.1, 0.05^2)$, and for pleiotropic variants $\gamma_j \sim \mathcal{N}(\mu_\gamma, 0.05^2)$, with $\mu_\gamma = 0$

(balanced pleiotropy) or $\mu_\gamma = 0.05$ (directional pleiotropy). Although unilateral loci do not directly affect X, they can perform as a background noise to interfere IV selection through gene interactions. We utilize 100 replicates for each scenario and report the average.

For scenarios with non-additive genetic effects, we simulate phenotypes using more complicated polygenetic models with simple multiplicative effects, interactive multiplicative effects, and interactive threshold effects (Marchini et al. 2005). Details can be found at Supplementary Note SN3.

Results

BNMR can efficiently identify effect variants from numerous weak, interacting variants with high precision in learning stage

We first compare the fine-mapping performance of RGF with different hyperparameters and structure learning algorithms in simulated datasets (Fig. 2). RGF exhibits a lower false discovery rate (FDR) and a higher AUC, with increasing subsample sizes and numbers of subsamples, though this improvement is accompanied by an increase in time consumption. Constrained-based approaches yield lower FDR, while the score-based approaches are superior in speed. As the selection threshold increases, the number of identified variants diminishes with increased precision.

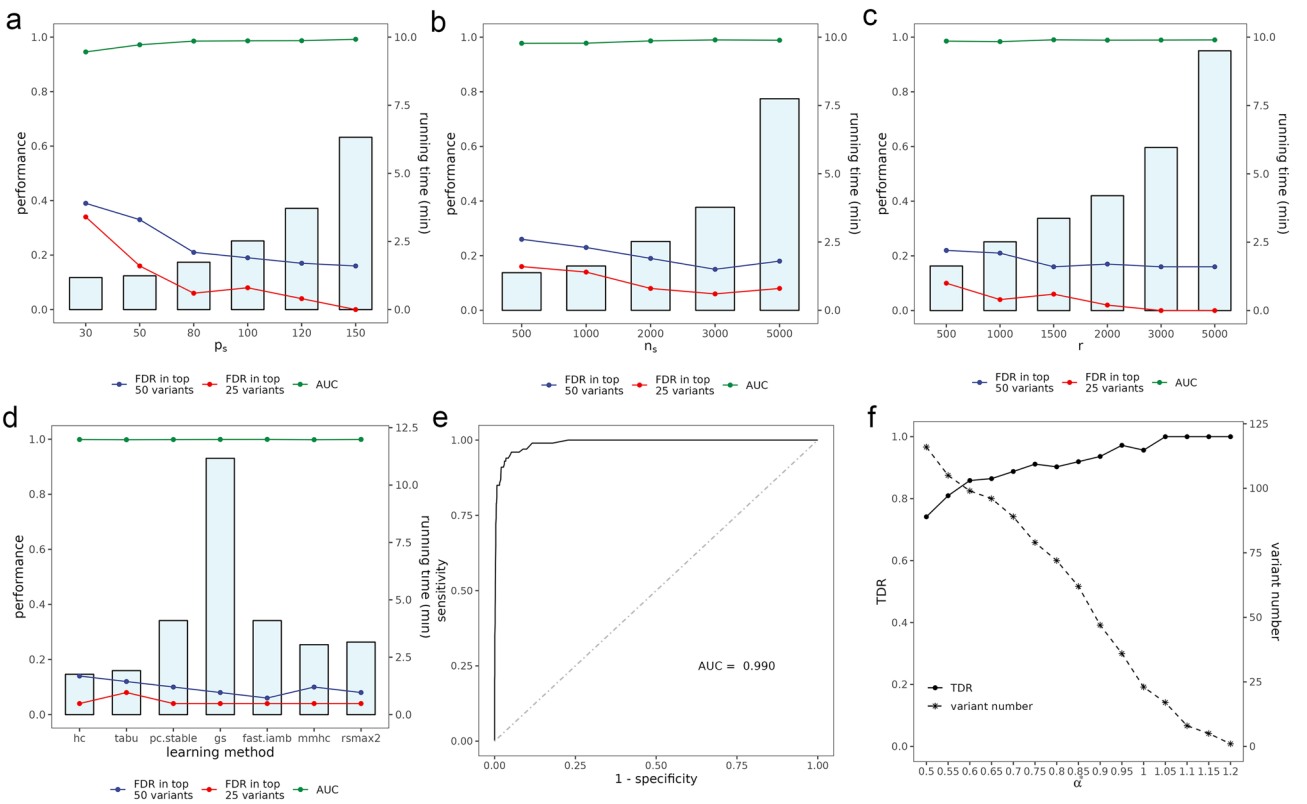


Fig. 2 The performance of the RGF with different hyperparameters and BN structure learning methods in simulations. **a** The performance and time consumption of RGF with different numbers of subsampling variants using the Hill-Climbing (hc) algorithm. **b** The performance and time consumption of RGF with different numbers of subsampling individuals using the hc algorithm. **c** The performance and time consumption of RGF with different numbers of subsamples. **d** The performance and time consumption of RGF with different BN structure learning methods. We evaluate score-based approaches including hc and tabu, constrained-based approaches including stable PC (pc.stable), Incremental Association (iamb), and Grow-Shrink (gs), as well as hybrid learning methods including Max-Min Hill-Climbing (mmhc) and Restricted Maximization (rsmax2). The lines show the corresponding FDR and AUC, while the gray bars display the changes in consumed time (min). **e** The ROC curve for RGF ($n_s = 2000$, $p_s = 120$, $r = 1000$). **f** The relationships among selection threshold (α^*), number of selected variants, and FDR. Simulated data size: $n = 5000$, $p = 2000$, with 100 true effect variants for the exposure. FDR: false discovery rate. AUC: area under the receiver operating characteristic (ROC) curve

Compared to the conventional association test (linear regression), RGF, LD stepwise pruning, and penalized regressions (especially lasso and elastic net) can all reduce FDR, while the RGF achieves the highest precision, performing as a effective tool in prioritizing candidate effect variants and identifying true effect variants ($\mathcal{Z}_D$) (Fig. 3). Before employing these variable selection strategies, we conduct pre-filtering to reduce the number of candidate variants. A more strict P threshold before RGF increases the precision in top variants but may reduce the recall. A threshold of around $\frac{1}{p}$ to $\frac{0.01}{p}$ may be proper. Besides, adjacency score can be regarded as the confidence that the variant belongs to $\mathcal{Z}_D$, and thus, it is also an assessment of instrument strength. The correlation between the adjacency score and the commonly used F statistics (Fig. S2) indicate that RGF is capable of choosing instruments reliable relevance to reduce weak-instrument bias.

GWAS becomes tricky when dealing with non-additive genetic effects (Tam et al. 2019). We simulate phenotypes from the combination of three interacting polygenetic models, including simple multiplicative effects, interactive multiplicative effects, and interactive threshold effects (Supplementary Note SN3) (Marchini et al. 2005). The results show that RGF performs well in settling the puzzle of gene interaction and epistasis (Fig. S4).

Since genotypes are notoriously difficult to simulate, we also generate phenotypes using the same procedures (Eq. 10)

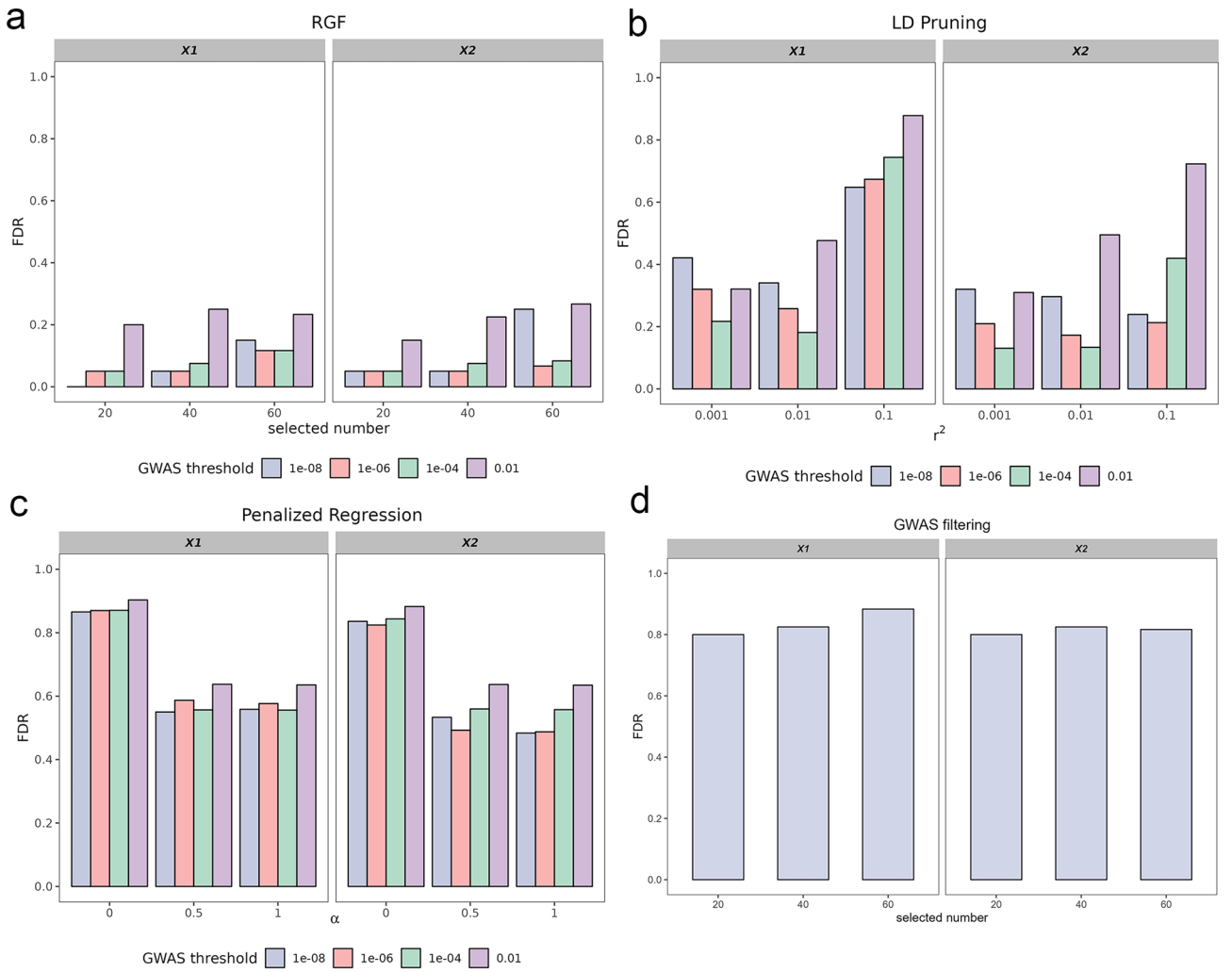


Fig. 3 False discovery rate (FDR) of different IV selection strategies in simulations. **a** FDR of RGF. **b** FDR of LD stepwise pruning. r^2: correlation thresholds for LD pruning. **c** FDR of penalized regressions. α is set to be 1.0 (Lasso), 0.5 (elastic net), and 0 (ridge regression), with penalty factor λ determined by 10-fold cross validation. **d** FDR of GWAS lead SNPs. Variants are pre-filtered according to specific GWAS P-value thresholds before LD pruning, penalized regressions, or RGF. Simulated data size: $n = 10{,}000$, $p = 10{,}000$, with 300 true effect variants for each trait. The environmental variance (σ_x^2) for X_2 is imposed to be twice as much as that of X_1 to represent traits with lower heritability. RGF is conducted with $n_s = 2000$, $p_s = 150$, $r = 5000$ using hc algorithm

but based on real genotype data from the UK BioBank. Using the synthetic datasets, we demonstrate the adaptation of RGF to different scales of genomes (Fig. 4). RGF is capable of handling the complex structure of real genetic data and exhibits good adaptability to datasets with different scales.

BNMR can effectively reduce mean square error of estimates, enhance statistical power, and is robust to horizontal pleiotropy in inference stage

We first compare the performance of BNMR to other existing MR approaches (Supplementary Table S5), including two-stage least square (TSLS), limited information maximum likelihood ratio (LIML) (Boehm and Zhou 2022), inverse-variance weighted (IVW) (Burgess et al. 2013), MR-Egger (Bowden et al. 2015), weighted median (Bowden et al. 2016), weighted mode (Hartwig et al. 2017), JAM-MR (Gkatzionis et al. 2021), CAUSE (Morrison et al. 2020), and CIIV (Windmeijer et al. 2021) (Fig. 5a). We include two types of pleiotropic loci in our simulation: pleiotropy loci that independently affect exposure and outcome, or the effects on the exposure and outcome correlated resulting from gene interaction like linkage. For each scenario, we examine the performance of various methods under settings where the expected average pleiotropic effect of all loci was either 0 (balanced pleiotropy) or non-zero (directional pleiotropy). Most prevailing approaches perform relatively well in balanced pleiotropy, but fail to cope with scenarios with complex directional and correlated pleiotropy. Due to its sensitivity to the InSIDE assumption, MR-Egger performs noticeably worse when the number of correlated pleiotropic variants increases. Some two-sample methods, like CAUSE, confront an inflated estimation variance, and bias from sample overlap when applying to one-sample studies (Burgess et al. 2016). Methods based on plurality rule in a broad sense such as weighted median and mode estimators exhibit commendable stability. In general, BNMR is superior to the existing approaches in terms of mean squared error (MSE), particularly due to its smaller variance of estimates, yielding augmented statistical power. Despite relying on the InSIDE assumption, the process of using RGF for IV selection enhances the robustness of the InSIDE assumption, making it more resilient to correlated pleiotropy arising from gene–gene interactions.

To show the bonus BN brings to the conventional Bayesian MR, we then evaluate the improvement in Bayesian MR estimation by using IVs obtained from BN (BNMR) compared to using GWAS lead SNPs directly as IVs (BMR). A noticeable reduction in MSE can be observed when there is presence of correlated pleiotropy by gene interaction (Fig. 5b). This enhancement primarily manifests in the attenuation of estimated variance in balanced scenarios, while both bias and variance deflate when the pleiotropic effect is directional. To better understand the role of BN, we examine the performance of RGF-selected IVs on other traditional MR methods (Fig. 5c). Due to overlapping samples in single-sample designs and the requirement for IVW and MR-Egger to use independent IVs, while RGF aims to identify IVs that have a direct impact on the exposure (which may not

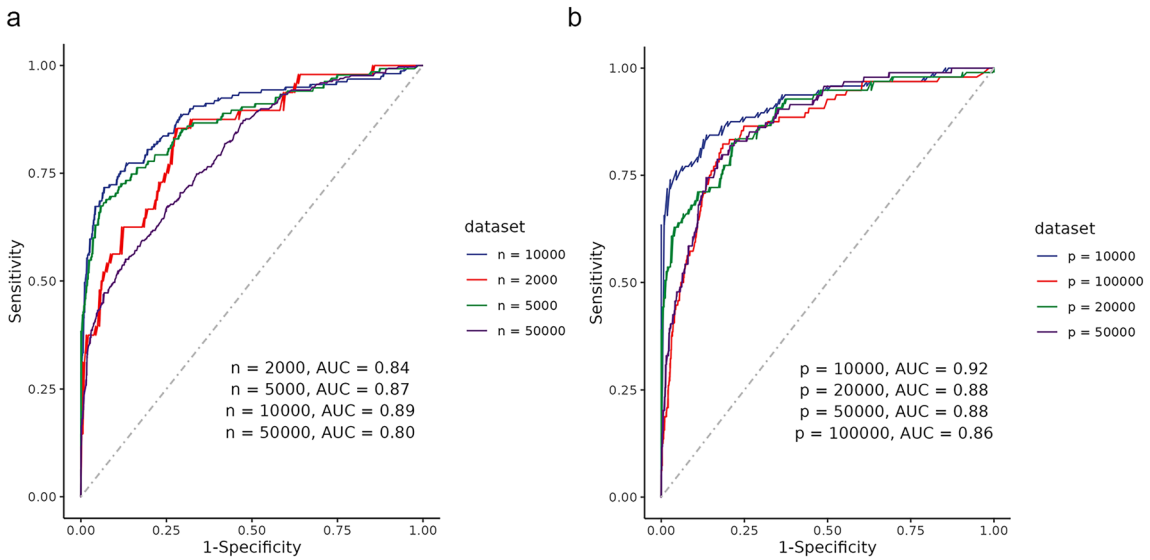
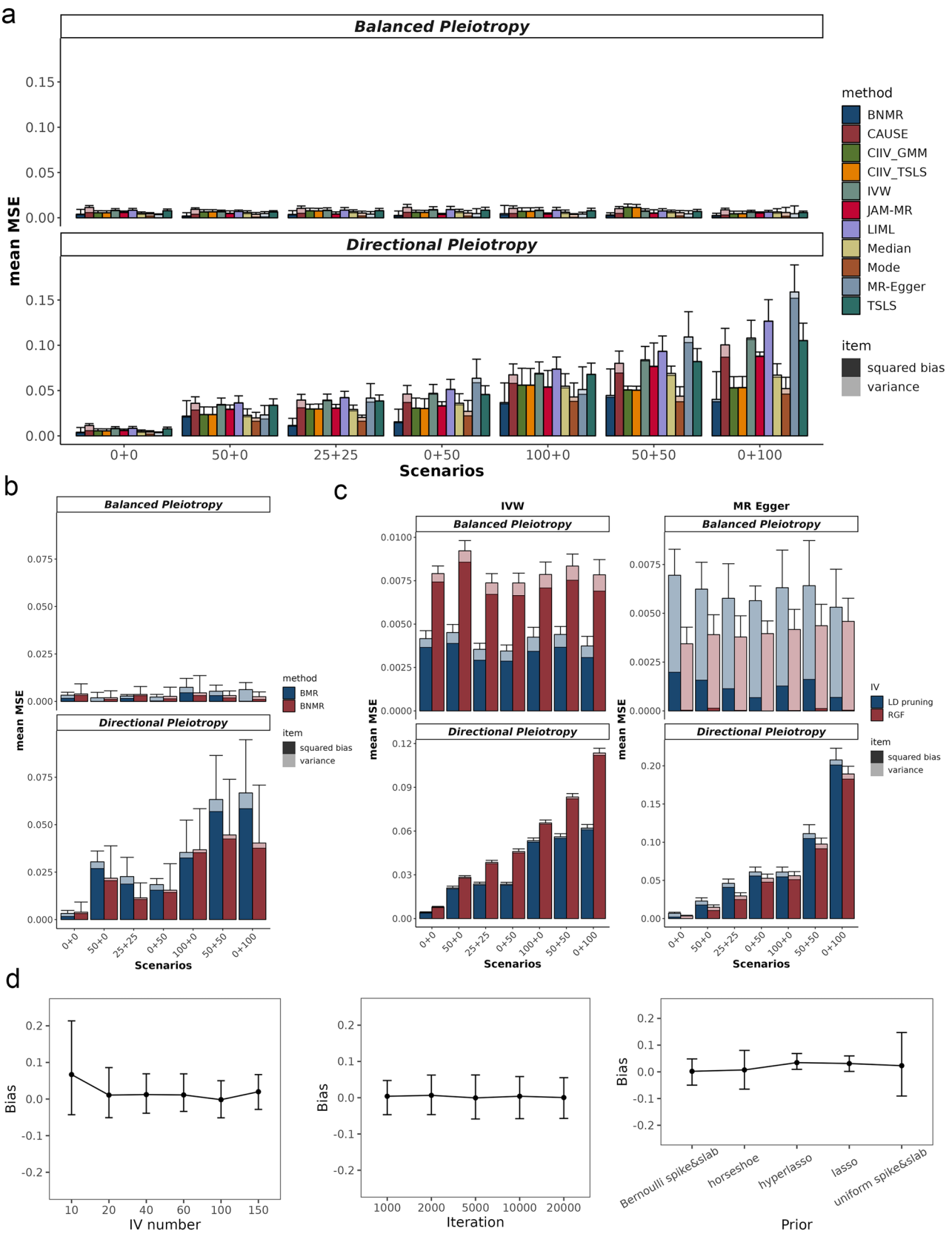


Fig. 4 Performance of BNMR on simulated phenotypes based on real genotypes from the UK Biobank. **a** The ROC curves for real genotype data with different sample size (n). Here p is fixed at 20,000. **b** The ROC curves for real genotype data with different genome size (p). Here n is fixed at 5000


a
Balanced Pleiotropy
Directional Pleiotropy
mean MSE
0.15
0.10
0.05
0.00
0+0
50+0
25+25
0+50
100+0
50+50
0+100
Scenarios
method
BNMR
CAUSE
CIIV_GMM
CIIV_TSLS
IVW
JAM-MR
LIML
Median
Mode
MR-Egger
TSLS
item
squared bias
variance
b
Balanced Pleiotropy
Directional Pleiotropy
mean MSE
0.075
0.050
0.025
0.000
Scenarios
method
BMR
BNMR
item
squared bias
variance
c
IVW
MR Egger
Balanced Pleiotropy
Directional Pleiotropy
mean MSE
0.0100
0.0075
0.0050
0.0025
0.0000
0.12
0.09
0.06
0.03
0.00
0.20
0.15
0.10
0.05
Scenarios
IV
LD pruning
RGF
item
squared bias
variance
d
Bias
0.2
0.1
0.0
-0.1
-0.2
10
20
40
60
100
150
IV number
1000
2000
5000
10000
20000
Iteration
Bernoulli spike&slab
horseshoe
hyperlasso
lasso
uniform spike&slab
Prior

◂**Fig. 5** Performance on causal effect estimation in simulations. **a** Averaged mean square errors (MSE) of different MR approaches. Robust regression and penalized weights are adopted in IVW and MR-Egger. IVs used in TSLS, LIML, IVW, MR-Egger, weighted Median and Mode, and CIIV are obtained through LD stepwise pruning. IVs used in JAM-MR, CAUSE, and BNMR are selected using their own filtering procedures from GWAS lead SNPs. Standard errors (SEs) of CAUSE are transformed from confidence intervals, of BNMR are calculated from MCMC posterior sampling. The scenarios are represented as "the number of independent pleiotropic variants (variants with independent direct effects on both X and Y) + the number of correlated pleiotropic variants (the effects on X and Y are correlated)". For each scenario, the pleiotropic effects are simulated to be either balanced (mean pleiotropic effect $\mu_Y = 0$) or directional (mean pleiotropic effect $\mu_Y = 0.05$). Each bar represents the average MSE of the estimation for the respective method in that scenario with error bar showing the empirical standard error of mean MSE. The darker section at the bottom represents the squared bias, while the lighter section at the top represents the variance. **b** MSE of Bayesian MR using IVs from BN (BNMR) and using the same number of IVs from GWAS lead SNPs (BMR). **c** MSE of IVW and MR-Egger using IVs from LD pruning and RGF. Both estimators were used their penalized robust versions. **d** The average bias and error bars under different IV numbers (left), iterations (middle), and shrinkage priors (right)

be independent of each other), the use of RGF-selected IVs does not perform well and even increases bias when using IVW estimator. Even so, on the other hand, when using MR-Egger estimator, the use of RGF-selected IVs reduced bias and variance in estimation compared to IVs obtained through LD pruning. We believe this is because MR-Egger and Bayesian MR are based on similar assumptions, the InSIDE assumption, requiring that the effect of Z on X and the pleiotropic pathways from Z to Y are independent. It is violated when Z affects the confounder U that affects both X and Y (i.e., correlated horizontal pleiotropy). If this correlated horizontal pleiotropy is caused by gene–gene interactions, where U is another genetic locus Z' (Fig. 1c), RGF will tend to select Z' as IV ensuring that the InSIDE assumption is still satisfied. Therefore, RGF enhances the robustness of the InSIDE assumption.

We also conduct sensitivity analysis on the instrument numbers and iterations, and different shrinkage priors (Van Erp et al. 2019) (Fig. 5, Supplementary Note SN4–SN5). Bias increases when there are too many or too few instruments. BNMR estimates are not sensitive to priors in general, despite the fact that uniform spike and slab prior is a bit more inefficient than the others based on the error bars (Fig. 5d). The Bayesian Lasso prior shows the fastest sampling speed and a deflated standard error, but has a slightly higher bias. The horseshoe prior, although slightly less efficient, is superior in the performance of convergence due to the lowest Rhat (Table S4 and Fig. S5).

BNMR with large-scale BioBank-level data vindicates causality from erythrocyte-related traits to blood pressures

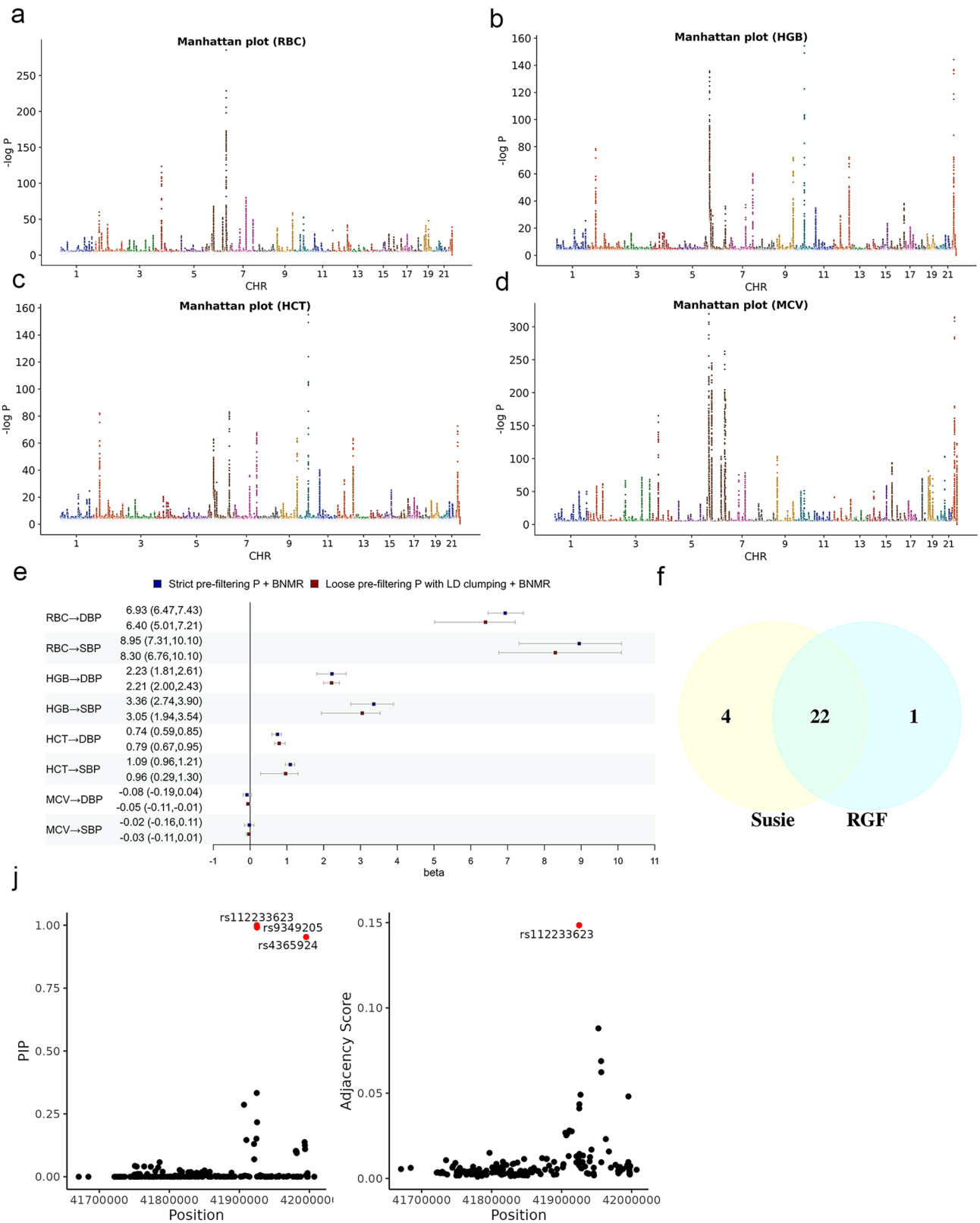
To highlight the practical significance and applicability of our method on extensive modern genetic datasets, we provide illustrative examples of two real-world studies featuring both continuous and binary outcomes, utilizing data sourced from the UK BioBank.

Hematological indices usually vary in a variety of physiological processes and are potential indicators for related disorders. Correlational studies have proposed that erythrocyte-related characteristics, including red blood cell count (RBC), hemoglobin concentration (HGB), hematocrit (HCT), the proportion of RBCs to the plasma, and mean red cell volume (MCV) (Enawgaw et al. 2017), are in strong correlation with systolic and diastolic blood pressures (SBP and DBP), and abnormalities of erythrocytes might be indicators of some cardiovascular and cerebrovascular diseases such as hypertension (Tsuda 2020).

To examine the causal effects of erythrocyte parameters on blood pressures, we involve 246,659 participants of Caucasian ancestry, self-reported as free from hypertension or other cardiovascular diseases (UK Biobank Non-cancer Illness Coding 1065-1094), and with available blood routine measurements at the time of enrollment. Genome quality control is conducted using PLINK 2.0, with corresponding thresholds for the SNP missing rate, minor allele frequency (MAF), and HWE test are 0.05, 0.01, and 1e−6. Fast posterior sampling with the large dataset is conducted with the Python packages PyMC and JAX. To increase power, we conducted preliminary GWAS filtering using summary statistics from a different dataset of the same ethnic group but with distinct samples (Astle et al. 2016).

We exert two pre-filtering strategies to reduce the amount of candidate variants and then conduct BNMR analysis. The first strategy utilizes a more stringent GWAS P threshold of 1e−20, while the second strategy employs a looser P threshold of 5e−8, followed by LD clumping. The results (Fig. 6e) consistently show that RBC, HGB, and HCT show significant positive effects on both DBP and SBP, and the effect magnitude is larger on SBP than on DBP. Whereas MCV shows a non-significant negative effect on blood pressure instead. Alternative approaches use top GWAS significant SNPs after LD clumping as instruments, and the varied and even conflicting results (Table S5) remind us of the importance of MR methodology. MR-Egger test shows that all causal relationships are not significant. On the other hand, TSLS indicates that RBC, HGB, and HCT have a significant effect on DBP but not on SBP, while CIIV estimates a positive effect on DBP and a negative effect on SBP.

Comparing the causal variants identified by RGF and probabilistic fine-mapping methods such as Susie (Wang



et al. 2020) is quite interesting. In general, RGF focuses on the genomic global landscape, while fine-mapping methods focuse more on local features. Taking RBC as an example, if we coarse fine-mapping using both methods on all candidate loci after preliminary screening, we would find that the majority of signals are shared between the two approaches

◂**Fig. 6** Causal relationships from erythrocyte-related traits to blood pressures. **a**–**d** Manhattan plots for RBC (**a**), HGB (**b**), HCT (**c**), and MCV (**d**), where the vertical coordinate shows the negative logarithm of the GWAS association test *P*-values for each locus. **e** Forest plot of the causal estimations. The units of RBC, HGB, HCT, MCV, and blood pressure are million/mm^3, g/dl, %, fL, and mmHg. We adopt two different pre-filtering strategies: one uses a more strict GWAS *P* threshold (1e−20), the other uses a looser *P* threshold (5e−8) and then conducts LD clumping (threshold: window = 10,000 kb, $r^2 = 0.01$, MAF=0.01). The variant loci obtained from pre-filtering are then further selected via RGF ($n_s = 4000$, $p_s = 150$, $r = 5000$) to identify 20 instruments for each exposure, shown in Supplementary Note SN6, and sample 4 chains with 5000 iterations per chain in MCMC. **f** Fine-mapping results of Susie and RGF using all variants after pre-filtering using the second strategy. **g** Fine-mapping results of Susie and RGF in the GWAS peak region chr6 41,600,000–42,200,000

(Fig. 6f). When we specifically examine the local structure near a GWAS peak (e.g., the region from 41,600,000 to 42,200,000 on chromosome 6), although Susie tends to identify more causal loci, the most significant signal (rs112233623) is the same for both methods (Fig. 6g). Considering that in IV selection, we are more concerned about false-positive signals in GWAS caused by genetic correlation and winner's curse, the relative conservatism of RGF is not a disadvantage.

The underlying mechanisms may relate to blood viscosity. Higher RBC, HGB, and HCT mean an increase in blood viscosity and peripheral resistance to blood flow, resulting in hypertension (Enawgaw et al. 2017). Besides, erythrocytes and hemoglobin also influence nitric oxide bioavailability, a crucial signal in the regulation of vessel psychology such as vasodilatation, thrombosis inhibition, and vessel formation (Helms et al. 2018). Although the molecular mechanisms still remain to be uncovered, the findings indicate that those hematological indices may be not only indicators but potential therapeutic targets for hypertension.

BNMR indicates that increased leukocytes contribute to the risk of depression

The neuro-immune interaction has been an appealing topic in recent years. Widespread bidirectional circuits exist between the two systems. The nervous system regulates immune activity and cytokine balance via the direct connection of sympathetic and parasympathetic nerves, and some neurotransmitters and neuroendocrine hormones can also serve as immunomodulators. Meanwhile, the immune system participates in the elimination and plasticity of synapses during development and modulates brain activity as well (Dantzer 2018). Immune-related hematological biomarkers provide a new insight into the pathological mechanisms of many psychiatric disorders. For instance, immune dysregulation has long been regarded associated with psychological disorders including depression (Drevets et al. 2022). Recent studies report the correlation between leukocyte count and depression (Reay et al. 2022; Sørensen et al. 2023).

We leverage disease records from UK Biobank and construct a case–control study by randomly selecting the same number of healthy individuals of the same ethnicity to assess whether leukocyte count, as well as its two subtypes, lymphocyte and monocyte counts, will causally affect depression. Subjects with extreme values exceeding 3 σ are excluded, and 22,324 cases and 22,861 controls are included in the analysis. All participants are of Caucasian ancestry.

Significant differences in leukocyte, lymphocyte, and monocyte counts are manifested between the case and control groups (Fig. 7a). Results from BNMR and weighted median indicate that an elevated leukocyte count will increase the risk of depression (Fig. 7b). The reciprocal MR analysis supports the causal direction from leukocyte count to depression. However, when we examine the two subtypes of leukocytes—lymphocytes and monocytes—this significant positive causal relationship disappears. This suggests that the causal mechanism from immune cells to mental disorders is more complex than anticipated, and warrants a careful examination of the influence of various cell type counts and compositional ratios (Sørensen et al. 2023).

We further conduct gene mapping and functional annotation by FUMA (Watanabe et al. 2017) using the top 1500 variants identified in RGF, which maps 273 depression-related protein coding genes. Functional analysis shows that these genes are enriched in the KEGG systemic lupus erythematosus and glycosaminoglycan Degradation pathway (Fig. 7c), both related closely with immune system (Handel and Dyer 2021). Depression-related genes also indicates enrichment in many cytokine and immune response pathways, including reactomes related to signaling of interleukin 9, Wnt, biocarta, and butyrophilin family (Supplementary Fig. S6), consistent with previous study (Wray et al. 2018). Depressive symptoms often share resemblance with inflammation-induced syndrome such as lethargy and inactivity, and the findings support role of immunity in the development of depression.

Immune targets for therapeutic development in depression has become a promising area in recent years (Drevets et al. 2022). Our analysis supports the idea of modulating immune cell composition as an intervention for psychological depression. However, the results should still be interpreted with caution due to recipient inclusion and sample size, population heterogeneity, and other potential confounding factors. Noncollapsibility of the logistic model may damage the estimation of binary responses (Schuster et al. 2021). Collaboration with evidence by means of triangulation (Lawlor et al. 2016) is vital to drawing a solid and reliable conclusion.

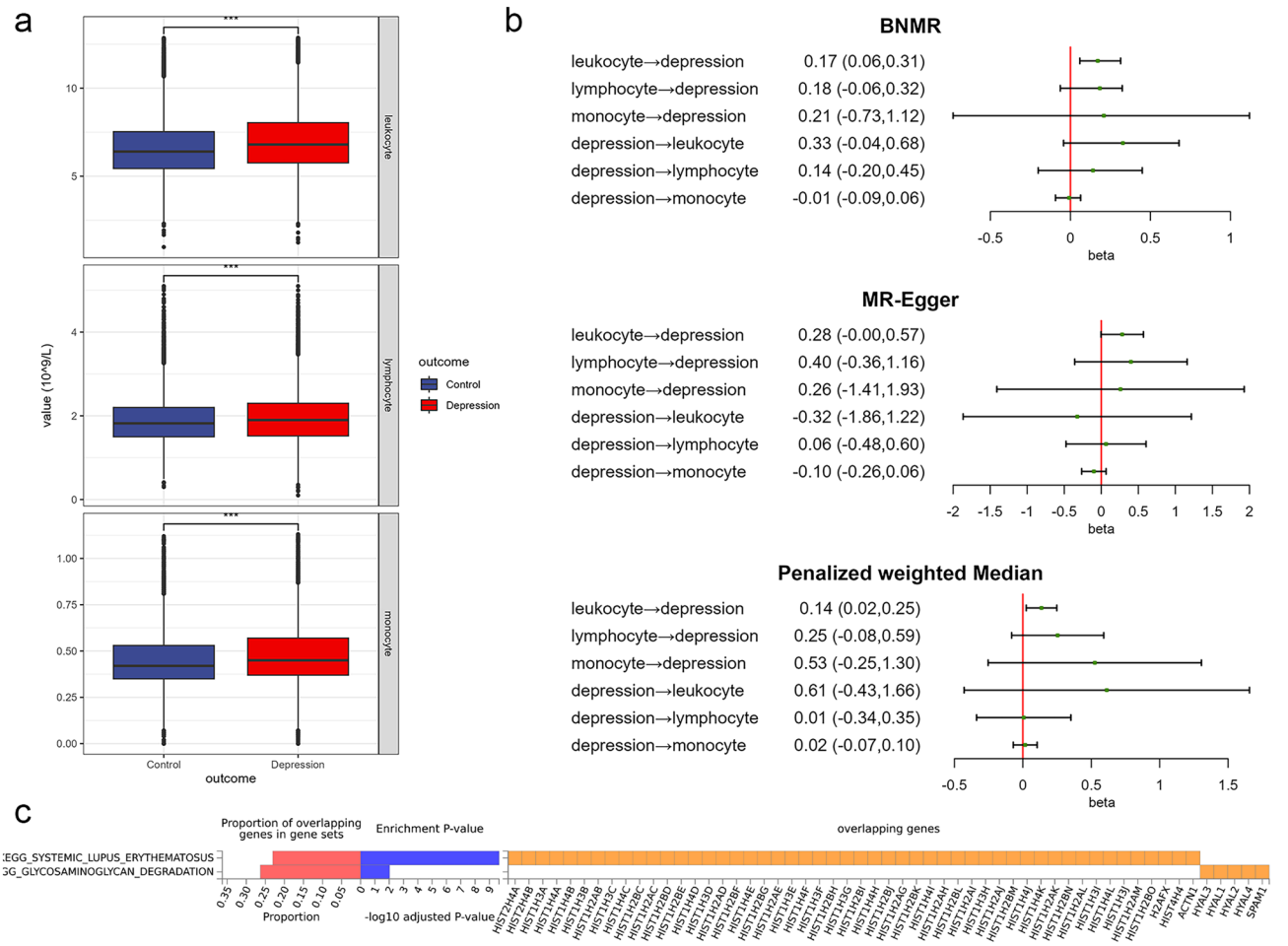


Fig. 7 The relationships between immune cell count and psychiatric disorders. **a** Differences in immune cell counts between the case and control groups. The significance level is calculated using the two-side T test. **b** Forest plot of the causal estimations of BNMR, penalized robust MR-Egger, and penalized weighted median. The unit of blood cell count is billion cells per liter. For BNMR, we select 20 instruments for each exposure via RGF ($n_s = 5000$, $p_s = 150$, $r = 5000$), shown in Supplementary Note SN6, and sample 4 chains with 5000 iterations per chain in MCMC. For MR-Egger and weighted median, GWAS lead variants after LD clumping are selected as instruments. **c** KEGG pathway enrichment of genes mapped by depression-related SNPs

Discussion

Causality is challenging to identify in observational studies due to unmeasured confounders. The introduction of genetic instruments in MR makes it possible to estimate causal effect in the presence of unobserved confounders, making MR increasingly appealing in real-world studies.

Tackling imperfect IVs has always been a tricky issue in MR. We propose BNMR to address the challenges by leveraging machine learning techniques and integrating causal discovery and inference. We use RGF to reduce FDR and improve statistical power when selecting instruments with true effects from numerous correlated weak variants due to polygenicity, epistasis, and LD. Then we control horizontal pleiotropy by imposing a shrinkage prior on the Bayesian MR. The selection of SNPs with direct effects on exposure enhances the robustness of InSIDE assumption and reduce correlated pleiotropy due to gene interaction, and the avoidance of false-positive signals in IV selection also contributes to reducing weak-instrument bias and enhancing statistical power.

To guarantee the faithfulness and sufficiency of causal diagrams, we impose constraints in RGF that limit the nodes in the graph to only include genetic loci and a single exposure. We tend not to involve multiple traits in a causal graph because the common causes of those traits may not be observed. Another advantage is that the criteria 'not d-separated from exposure by other variants' can be simply expressed as 'adjacent to exposure' in this scenario, which is convenient for DIE partition and IV selection.

Bayesian estimation with imposed shrinkage priors is conceptually similar to regularization in the traditional

model but with some obvious advantages, like simultaneously estimated penalty parameters, easily obtained credible intervals, and intuitive interpretation. In addition, domain-specific knowledge can be included as an informative prior. BNMR are not sensitive to priors, though we recommend horseshoe prior for better convergence performance if no additional information is accessible.

Although large-scale biobanks containing genotypes and phenotypes are now available, an increasing number of studies tend to report summary association statistics instead due to concerns on privacy and security. Bayesian meta-analysis is applied to assess pooled genetic relevance (Sun et al. 2022). Recent work has started to focus on learning causal diagrams with summary data (Zhang et al. 2017), while arduous task still remains.

BNMR is an example of post-selection inference and faces the issue that the inference stage does not account for uncertainty in the selection stage, causing more volatile results. The model also confronts computational challenges in BN learning and MCMC sampling, especially with increasing numbers of samples and variables. BN structural learning is an NP-hard problem. We leverage the bagging technique in ensemble learning and propose the RGF to split the whole genetic pattern into a series of subgraphs. The number of candidate variants is restricted via pre-filtering by GWAS association tests before RGF, since the removal of variants with low correlations will not influence the network structure severely due to the modularity of the causal diagram. To achieve a balance between sufficiently high precision and acceptable time consumption, we suggest conducting pre-filtering using a GWAS P threshold of approximately $\frac{1}{p}$ to $\frac{0.01}{p}$ and setting the value of $p_s r$ to be at least 100 times greater than the number of variants in RGF to ensure adequate sampling for each variant. We endorse the use of at least 4 chains and at least 2000 iterations in MCMC. For large-scale datasets, consolidation of posterior sampling in subsamples may be feasible.

In summary, BNMR is a practical model to prioritize and select proper instruments from massive, interacting, and weak variants and obtain pleiotropy-robust causal effect estimates. With accumulated genomic data available, BNMR will contribute to revealing more causal relationships and discovering potential therapeutic targets with real-world evidence.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00439-024-02640-x.

Author contributions JS and YZ designed the study. JS performed the research, analyzed the data, and wrote the original manuscript. YM participated in the real-world data management. ZY and YZ supervised the research. JS, JZ, YG, CP, YM, JZ, ZY, and YZ discussed and revised the manuscript.

Funding The research is supported by the National Natural Science Foundation of China (11901387 for YZ and 12171318 for ZY).

Data and code availability The GWAS summary statistics are accessed from GWAS Catalog (https://www.ebi.ac.uk/gwas/downloads/summary-statistics). The real-world data underlying this article are accessed from UK Biobank (https://www.ukbiobank.ac.uk/). We implement the BNMR framework as an R package, and all the codes are available at https://github.com/sjl-sjtu/bnmr2.

Declarations

Conflict of interest No competing interest exists.

Ethic approval UK Biobank has approval from the North West Multicenter Research Ethics Committee (21/NW/0157). The real-world studies have been conducted using the UK Biobank Resource under Application Number 100014.

Consent to participate Not applicable.

Consent to publish Not applicable.

References

Amar D, Sinnott-Armstrong N, Ashley EA et al (2021) Graphical analysis for phenome-wide causal discovery in genotyped population-scale biobanks. Nat Commun 12(1):1–11

Astle WJ, Elding H, Jiang T et al (2016) The allelic landscape of human blood cell trait variation and links to common complex disease. Cell 167(5):1415–1429

Berzuini C, Guo H, Burgess S et al (2020) A Bayesian approach to Mendelian randomization with multiple pleiotropic variants. Biostatistics 21(1):86–101

Boehm FJ, Zhou X (2022) Statistical methods for Mendelian randomization in genome-wide association studies: a review. Comput Struct Biotechnol J 20:2338–2351

Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 44(2):512–525

Bowden J, Davey Smith G, Haycock PC et al (2016) Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 40(4):304–314

Burgess S, Butterworth A, Thompson SG (2013) Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 37(7):658–665

Burgess S, Davies NM, Thompson SG (2016) Bias due to participant overlap in two-sample mendelian randomization. Genet Epidemiol 40(7):597–608

Burgess S, Zuber V, Valdes-Marquez E et al (2017) Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. Genet Epidemiol 41(8):714–725

Dantzer R (2018) Neuroimmune interactions: from the brain to the immune system and vice versa. Physiol Rev 98(1):477–504

Davies NM, von Hinke Kessler Scholder S, Farbmacher H et al (2015) The many weak instruments problem and Mendelian randomization. Stat Med 34(3):454–468

Drevets WC, Wittenberg GM, Bullmore ET et al (2022) Immune targets for therapeutic development in depression: towards precision medicine. Nat Rev Drug Discov 21(3):224–244

Dudbridge F (2021) Polygenic Mendelian randomization. Cold Spring Harb Perspect Med 11(2):a039586. https://doi.org/10.1101/cshperspect.a039586
Enawgaw B, Adane N, Terefe B et al (2017) A comparative cross-sectional study of some hematological parameters of hypertensive and normotensive individuals at the University of Gondar Hospital, Northwest Ethiopia. BMC Hematol 17(1):1–7
Gkatzionis A, Burgess S, Conti DV et al (2021) Bayesian variable selection with a pleiotropic loss function in Mendelian randomization. Stat Med 40(23):5025–5045
Gkatzionis A, Burgess S, Newcombe PJ (2023) Statistical methods for cis-Mendelian randomization with two-sample summary-level data. Genetic Epidemiol 47(1):3–25. https://doi.org/10.1002/gepi.22506
Guo Z, Kang H, Tony Cai T et al (2018) Confidence intervals for causal effects with invalid instruments by using two-stage hard thresholding with voting. J Roy Stat Soc Ser B (Stat Methodol) 80(4):793–815
Handel TM, Dyer DP (2021) Perspectives on the biological role of chemokine: glycosaminoglycan interactions. J Histochem Cytochem 69(2):87–91
Hartwig FP, Davey Smith G, Bowden J (2017) Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 46(6):1985–1998
Helms CC, Gladwin MT, Kim-Shapiro DB (2018) Erythrocytes and vascular function: oxygen and nitric oxide. Front Physiol 9:125
Howey R, Shin SY, Relton C et al (2020) Bayesian network analysis incorporating genetic anchors complements conventional Mendelian randomization approaches for exploratory analysis of causal relationships in complex data. PLoS Genet 16(3):e1008198
Kang H, Zhang A, Cai TT et al (2016) Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. J Am Stat Assoc 111(513):132–144
Lawlor DA, Tilling K, Davey Smith G (2016) Triangulation in aetiological epidemiology. Int J Epidemiol 45(6):1866–1886
Lyu R, Sun J, Xu D et al (2021) GESLM algorithm for detecting causal SNPs in GWAS with multiple phenotypes. Brief Bioinform 22(6):bbab276
Marchini J, Donnelly P, Cardon LR (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet 37(4):413–417
Morrison J, Knoblauch N, Marcus JH et al (2020) Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. Nat Genet 52(7):740–747
Nogueira AR, Pugnana A, Ruggieri S et al (2022) Methods and tools for causal discovery and causal inference. Wiley Interdiscipl Rev Data Min Knowl Discov 12(2):e1449
Pingault JB, O'reilly PF, Schoeler T et al (2018) Using genetic data to strengthen causal inference in observational research. Nat Rev Genet 19(9):566–580
Pritchard JK, Przeworski M (2001) Linkage disequilibrium in humans: models and data. Am J Hum Genet 69(1):1–14
Reay WR, Kiltschewskij DJ, Geaghan MP et al (2022) Genetic estimates of correlation and causality between blood-based biomarkers and psychiatric disorders. Sci Adv 8(14):eabj8969
Sanderson E, Glymour MM, Holmes MV et al (2022) Mendelian randomization. Nat Rev Methods Prim 2(1):1–21
Schuster NA, Twisk JWR, ter Riet G et al (2021) Noncollapsibility and its role in quantifying confounding bias in logistic regression. BMC Med Res Methodol. https://doi.org/10.1186/s12874-021-01316-8
Sørensen NV, Frandsen BH, Orlovska-Waast S et al (2023) Immune cell composition in unipolar depression: a comprehensive systematic review and meta-analysis. Mol Psychiatry 28(1):391–401
Sun J, Lyu R, Deng L et al (2022) SMetABF: a rapid algorithm for Bayesian GWAS meta-analysis with a large number of studies included. PLoS Comput Biol 18(3):e1009948
Tam V, Patel N, Turcotte M et al (2019) Benefits and limitations of genome-wide association studies. Nat Rev Genet 20(8):467–484
Tsuda K (2020) Red blood cell abnormalities and hypertension. Hypertens Res 43(1):72–73
Van Erp S, Oberski DL, Mulder J (2019) Shrinkage priors for Bayesian penalized regression. J Math Psychol 89:31–50
Wang G, Sarkar A, Carbonetto P et al (2020) A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Ser B Stat Methodol 82(5):1273–1300
Watanabe K, Taskesen E, Van Bochoven A et al (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8(1):1826
Windmeijer F, Farbmacher H, Davies N et al (2019) On the use of the lasso for instrumental variables estimation with some invalid instruments. J Am Stat Assoc 114(527):1339–1350
Windmeijer F, Liang X, Hartwig FP et al (2021) The confidence interval method for selecting valid instrumental variables. J Roy Stat Soc Ser B (Stat Methodol) 83(4):752–776
Wray NR, Ripke S, Mattheisen M et al (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 50(5):668–681
Yang J, Ferreira T, Morris AP et al (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 44(4):369–375
Zhang L, Pan Q, Wang Y et al (2017) Bayesian network construction and genotype-phenotype inference using GWAS statistics. IEEE/ACM Trans Comput Biol Bioinf 16(2):475–489