HGG Advances

ARTICLE

Two-sample bi-directional causality between two traits with some invalid IVs in both directions using GWAS summary statistics

Siyi Chen[1,*]

Summary

Mendelian randomization (MR) is a widely used method for assessing causal relationships between risk factors and outcomes using genetic variants as instrumental variables (IVs). While traditional MR assumes uni-directional causality, bi-directional MR aims to identify the true causal direction. In uni-directional MR, invalid IVs due to pleiotropy can violate assumptions and introduce biases. In bi-directional MR, traditional MR can be performed separately for each direction, but the presence of invalid IVs poses even greater challenges. We introduce a new bi-directional MR method incorporating stepwise selection (Bidir-SW) designed to address these challenges. Our approach leverages public genome-wide association study (GWAS) datasets for two traits and uses model selection criteria to identify invalid IVs iteratively by stepwise selection. This method accounts for potential bi-directional causality in the presence of common invalid IVs for both directions, even if only GWAS summary statistics are provided. Through simulation studies, we demonstrate that our method outperforms traditional MR techniques, such as MR-Egger and inverse-variance weighted (IVW), with uncorrelated SNPs. We also provide simulations to compare our approach with existing transcriptome-wide association study (TWAS) to show its effectiveness. Finally, we apply the proposed method to genetic traits such as CRP levels and BMI to explore possible bi-directional relationships among these traits. We also used the proposed method to discover causal protein biomarkers. Our findings suggest that the Bidir-SW approach is a powerful tool for bi-directional MR or TWAS, which can provide a valuable framework for future genetic epidemiology studies.

Introduction

Mendelian randomization (MR) is a powerful method to distinguish correlation from causation between risk factors and observational outcome using genetic variants as instrumental variables (IVs).[1] In biomedical studies, randomized controlled trials (RCTs) are considered as the "gold standard" to demonstrate causality between a risk factor and an outcome,[2] but can be prohibitively expensive or even unfeasible sometimes. In contrast, MR offers a possible alternative for RCTs when RCTs are not available, by using genetic variation to infer causal relationships between the exposure and the outcome.[1] Traditionally, MR analyses have focused on a uni-directional causal pathway, where a single trait is considered as the exposure and another as the outcome. The basic framework for MR is IV regression, in which a valid IV must satisfy the following assumptions[3–5]:

(1) The IV is associated with the exposure;
(2) The IV does not affect the outcome directly;
(3) The IV is not associated with the unmeasured confounders.

Assumption 1 can be met by using only significant IVs, whereas assumption 2 is violated when some IVs have uncorrelated pleiotropic or direct effects on the outcome. The violation of assumption 3 leads to what is known as correlated pleiotropy. Kang et al.[6] proposed an invalid IV selection method called *sisVIVE* based on LASSO, which does not include statistical inference, while Windemeijer et al.[7] introduced a two-stage method that incorporates causal effect inference. Like traditional MR approaches, these methods typically only consider uni-directional causality, focusing on one trait as the exposure and another as the outcome. However, some biological relationships, especially in genetic studies, can be seen as bi-directional[8] because the true direction is unclear.

For bi-directional MR, generally we first test whether trait X has causal effects on trait Y, then we reverse the direction to examine if Y causally influences X. Just like traditional MR, bi-directional MR relies on the identification of (in)valid IVs for each direction. For bi-directional MR, invalid IVs can introduce greater complexities. These invalid IVs can distort causal estimates in both directions, making it more difficult to determine the true causal relationships compared with uni-directional MR. Practically, it is challenging to determine the (in)valid IVs, and it is also common for researchers to mistakenly use all significant SNPs as valid IVs for any of the directions. When an SNP directly influences X (and subsequently Y through X), it may appear to be

[1]School of Public Health, LSU Health Sciences Center New Orleans, New Orleans, LA 70112, USA
*Correspondence: sche11@lsuhsc.edu
https://doi.org/10.1016/j.xhgg.2025.100449.

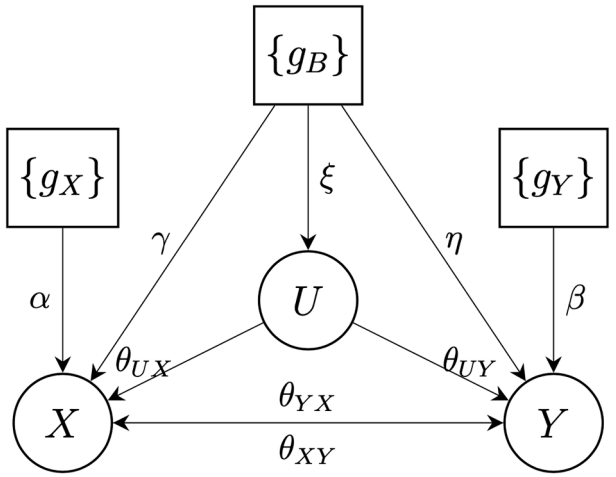

Figure 1. Causal diagram for the true bi-directional causal model of traits *X* and *Y*

associated with both traits, especially with larger sample sizes. This can lead to the false conclusion that a causal relationship exists in both directions.[8,9] Additionally, if an SNP affects X but not Y, yet the detection power for Y is higher, the SNP may incorrectly be used as a valid IV for Y,[9] resulting in a false conclusion that Y has causal effects on X.

Therefore, one of the key challenges in bi-directional MR is the identification of (in)valid IVs for each of the directions. For uni-directional MR, invalid IVs can arise from pleiotropy, where a genetic variant affects the outcome through pathways other than the exposure of interest, or due to linkage disequilibrium (LD) with other SNPs. Traditional MR methods such as inverse-variance weighted (IVW)[10] and MR-Egger[11] provide some robustness to pleiotropy but can still be biased when invalid IVs are present. Recent achievements[12,13] have focused on identifying and accounting for these invalid IVs to improve the reliability of causal estimates and statistical power. In bi-directional MR, an IV that is valid for X must be invalid for Y; in other words, validity for one direction cannot be true for the other, therefore adding more complexity to this problem.

In this research, we present a new model selection approach to address the presence of invalid IVs in bi-directional MR settings. We first select (in)valid IVs, and once valid IVs are identified for each direction, we implement a simple screening rule based on Steiger's approach[8] to ensure that no single SNP is used as a valid IV in both directions. This proposed method is able to control type-I errors and to provide more reliable inferences for complex genetic scenarios. We also conduct simulation studies to demonstrate the robustness and reliability of our proposed method. We compare our approach to various MR or transcriptome-wide association study (TWAS) techniques, such as IVW, MR-Egger, TWAS,[14] and imaging-wide association study (IWAS).[15] Following these simulations, we apply the method to publicly available GWAS summary statistics to infer causal directions among body mass index (BMI), sleep duration (SLP), and C-reactive protein (CRP) levels. Additionally, we perform the analysis on proteomics GWAS summary data, exploring potential bi-directional relationships between various cardiovascular disease (CVD)-related proteins and these traits. Our proposed approaches allow us to identify some interesting bi-directional causal connections, providing potential new insights into the underlying biological processes.

Methods

Model

Suppose there are n individuals, and we observe (X,Y,G), where X and Y are the two traits of interest, and G denotes the genetic variants. We assume that the set of G could be partitioned into three pairwise disjoint sets: $\{g_X\}$, $\{g_Y\}$ and $\{g_B\}$. Similar to Xue et al.[16] and Li et al.,[17] the true causal model is given by the following structural equations:

$$\begin{aligned} U &= \xi g_B + \varepsilon_U, \\ X &= \theta_{YX}Y + \alpha g_X + \gamma g_B + \theta_{UX}U + \varepsilon_X, \\ Y &= \theta_{XY}X + \beta g_Y + \eta g_B + \theta_{UY}U + \varepsilon_Y \end{aligned} \qquad \text{(Equation 1)}$$

In the proposed true model (Equation 1), U represents the unobserved confounder, $\{g_X\}$ and $\{g_Y\}$ stand for the sets of valid IVs for X and Y, respectively. $\{g_X\}$ and $\{g_Y\}$ are considered valid IVs for either direction. IVs in $\{g_B\}$, on the other hand, are considered common invalid IVs for both directions. The details of the three sets are discussed later. Arrows are used to indicate the directions of causal effects, with the corresponding effect sizes labeled, as presented in Figure 1. Specifically, θ_{XY} and θ_{YX} are used to denote the causal effects of interest for the directions $X\to Y$ and $Y\to X$, respectively.

Since all variables can be centered to have zero means, thus the last two equations of the model are rewritten as follows:

$$\begin{aligned} X &= \theta_{YX}Y + \alpha g_X + (\gamma + \theta_{UX}\xi)g_B + e_X, \\ Y &= \theta_{XY}X + \beta g_Y + (\eta + \theta_{UY}\xi)g_B + e_Y, \end{aligned} \qquad \text{(Equation 2)}$$

where $e_X = \theta_{UX}\varepsilon_U + \varepsilon_X$ and $e_Y = \theta_{UY}\varepsilon_U + \varepsilon_Y$. Based on previous assumptions, there are three sets of IVs: (1) $\{g_X\}$ is the set for valid IVs of Y; (2) $\{g_Y\}$ is the set for valid IVs of X; (3) $\{g_B\}$ is the set for invalid IVs of both directions. We can further conclude that

$$\begin{aligned} X = \frac{1}{1-\theta_{XY}\theta_{YX}}\Big(\alpha g_X + \big[(\gamma + \theta_{UX}\xi) + \theta_{YX}(\eta + \theta_{UY}\xi\big]g_B \\ + \theta_{YX}\beta g_Y + \theta_{YX}e_Y + e_X), \\ Y = \frac{1}{1-\theta_{XY}\theta_{YX}}(\beta g_Y + [(\eta + \theta_{UY}\xi) + \theta_{XY}(\gamma + \theta_{UX}\xi)]g_B \\ + \theta_{XY}\alpha g_X + \theta_{XY}e_X + e_Y), \end{aligned} \qquad \text{(Equation 3)}$$

where $1 - \theta_{XY}\theta_{YX} \neq 0$.

Notice when the causal direction is from $X\to Y$, from the true causal model, we can see that both g_X and g_B have direct effects on outcome Y; therefore, $\{g_B\}\cup\{g_X\}$ can be considered as the set of invalid IVs for Y. Similarly, when we consider the direction of $Y\to X$, $\{g_B\}\cup\{g_Y\}$ becomes the set of invalid IVs for X. We continue to have the following assumptions and the theorem:

Assumption 1. Suppose for causal direction $X\to Y$, we say an IV g is valid for Y if $g \in \{g_X\}$, and g is invalid if $g \in \{g_B\}\cup\{g_Y\}$.

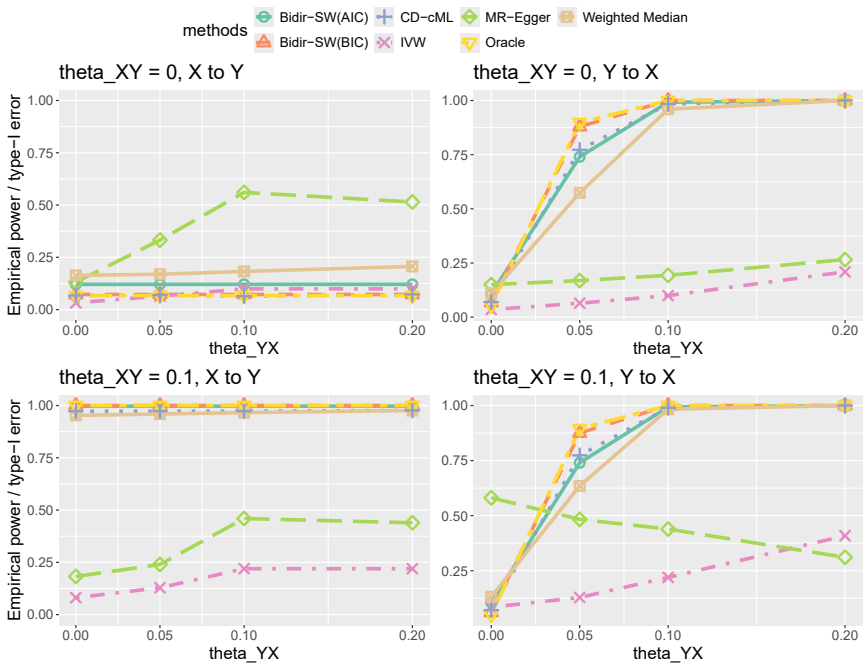

Figure 2. Simulation studies for independent SNPs comparing empirical type-I error (when $\theta_{XY} = 0$) and power (when $\theta_{XY} = 0.1$) across different values of θ_{YX} (0, 0.05, 0.1, 0.2)

Results are shown for different methods, including the oracle condition. Sample sizes are $n_1 = n_2 = 50,000$, and ξ is sampled from $U(-0.2, 0.2)$. The upper panel corresponds to $\theta_{XY} = 0$ (type-I error), and the lower panel corresponds to $\theta_{XY} = 0.1$ (power).

Similarly, for direction of $Y \to X$, an IV g is valid for X if $g \in \{g_Y\}$, and g is invalid if $g \in \{g_B\} \cup \{g_X\}$. We say that an IV $g \in \{g_X\} \cup \{g_Y\}$ is valid for either direction for a bi-directional causal model.

Assumption 2. For a bi-directional causal model, consider causal direction $X \to Y$, we have that $\theta_{XY}^2 \mathrm{Var}(X) < \mathrm{Var}(Y)$. Naturally, for causal direction $Y \to X$, we also have $\theta_{YX}^2 \mathrm{Var}(Y) < \mathrm{Var}(X)$. This assumption is adopted from Xue and Pan.[9]

Theorem 1. Suppose that we only know that IV g is valid for either causal direction, but we are not sure exactly for which direction g is a valid IV. If g satisfies $|\mathrm{corr}(g, Y)| < |\mathrm{corr}(g, X)|$ then we can say that g is a valid IV for causal direction $X \to Y$; otherwise, g is considered valid for causal direction $Y \to X$.

Proof of Theorem 1 can be found in section 3.1 of the supplemental information.

Table 1. Simulation studies for independent SNPs comparing empirical type-I error of various methods across different values of θ_{YX} (0, 0.05, 0.10, 0.20) with sample sizes $n_1 = n_2 = 50,000$ and $\xi \sim U(-0.2, 0.2)$, when $\theta_{XY} = 0$

	θ_{YX}			
Method	**0.00**	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.120	0.120	0.120	0.120
Bidir-SW (BIC)	0.065	0.065	0.065	0.063
CD-cML	0.062	0.063	0.058	0.059
IVW	0.025	0.054	0.099	0.099
MR-Egger	0.121	0.296	0.548	0.522
Oracle	0.057	0.057	0.057	0.055
Weighted Median	0.150	0.149	0.169	0.215
Method	θ_{YX}			
	0.00	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.126	0.754	0.985	1.000
Bidir-SW (BIC)	0.071	0.860	1.000	1.000
CD-cML	0.068	0.764	0.985	1.000
IVW	0.027	0.054	0.099	0.211
MR-Egger	0.139	0.162	0.190	0.266
Oracle	0.052	0.880	1.000	1.000
Weighted Median	0.113	0.574	0.950	1.000

The top table corresponds to the direction $X \to Y$, and the bottom table corresponds to $Y \to X$.

Table 2. Simulation studies for independent SNPs comparing empirical power of various methods across different values of θ_{YX} (0, 0.05, 0.10, 0.20) with sample sizes $n_1 = n_2 = 50,000$ and $\xi \sim U(-0.2, 0.2)$, when $\theta_{XY} = 0.1$

	θ_{YX}			
Method	**0.00**	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.998	0.998	0.998	0.998
Bidir-SW (BIC)	1.000	1.000	1.000	1.000
CD-cML	0.981	0.981	0.981	0.984
IVW	0.097	0.142	0.217	0.217
MR-Egger	0.167	0.218	0.430	0.428
Oracle	1.000	1.000	1.000	1.000
Weighted Median	0.954	0.960	0.966	0.985
Method	θ_{YX}			
	0.00	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.128	0.761	0.986	1.000
Bidir-SW (BIC)	0.065	0.870	1.000	1.000
CD-cML	0.071	0.777	0.986	1.000
IVW	0.099	0.139	0.219	0.391
MR-Egger	0.560	0.495	0.440	0.294
Oracle	0.048	0.887	1.000	1.000
Weighted Median	0.139	0.652	0.971	1.000

The top table corresponds to the direction $X \to Y$, and the bottom table corresponds to $Y \to X$.

The algorithm

The main concept of the proposed Bidir-SW approach is as follows: (1) We first identify invalid IVs for X and Y, denoted as $\widehat{\mathcal{A}}_X$ and $\widehat{\mathcal{A}}_Y$, respectively. (2) Since $\widehat{\mathcal{A}}_X$ and $\widehat{\mathcal{A}}_Y$ contain IVs that are potentially invalid for both directions, we explicitly identify these common invalid IVs $\{g_B\}$ by taking the intersection: $\{g_B\} = \widehat{\mathcal{A}}_X \cap \widehat{\mathcal{A}}_Y$. (3) After removing common invalid IVs, the remaining IVs form a candidate set of potentially valid IVs, denoted by $\{g_C\} = \widehat{\{g_B\}}^C$; therefore IVs in g_C are not considered as invalid for both directions simultaneously. Each IV in this complement set is thus invalid for at most one direction and remains a valid candidate IV for the other direction. Further validation will be performed in the next step. (4) In the final step, we will determine whether each IV in $\{g_C\}$ belongs to to the valid IV set $\{g_X\}$ or the valid IV set $\{g_Y\}$ by directly applying Theorem 1.

Algorithm 1Bidir-SW algorithm with Bayesian Information Criteria (BIC) or Akaike Information Criteria (AIC):

Input: Two independent GWAS datasets: GWAS1 for trait X and GWAS2 for trait Y.

(1) Consider causal direction $X \to Y$, determine the invalid IV set $\widehat{\mathcal{A}}_X$ by stepwise model selection. Then similarly we can identify invalid IV set $\widehat{\mathcal{A}}_Y$ for causal direction $Y \to X$. The estimated invalid IV set for both direction is given by Y. $\widehat{\{g_B\}} = \widehat{\mathcal{A}}_X \cap \widehat{\mathcal{A}}$
(2) $\{g_C\} = \widehat{\{g_B\}}^C$ is the complement of set $\widehat{\{g_B\}}$; we use $\{g_C\}$ to denote the set for valid IVs of either causal direction.
(3) Consider each of the IVs in $\{g_C\}$, calculate $|\mathrm{corr}(g, X)|$ and $|\mathrm{corr}(g, Y)|$ to determine the estimated valid IV sets $\widehat{\{g_X\}}$ and $\widehat{\{g_Y\}}$ for both directions, according to Theorem 1.
(4) For each direction, include IVs from $\widehat{\{g_X\}}$ for $X \to Y$ and IVs from $\widehat{\{g_Y\}}$ for $Y \to X$, respectively, to perform two-stage least squares (TSLS) regression.

Output: The estimated causal effects $\widehat{\theta}_{XY}$ and $\widehat{\theta}_{YX}$ (with corresponding standard errors) for directions $X \to Y$ and $Y \to X$, obtained by two-sample TSLS regression.

More details of the proposed algorithm will be given in the following sections.

Selection of invalid IVs for both directions

First, consider the direction of $X \to Y$ as an example. In this case, we treat X as exposure and Y as the outcome. We adopt the TSLS-stepwise framework in Chen et al.[18]

Consider two independent GWAS samples GWAS1 and GWAS2, where GWAS1 has n_1 individuals and GWAS2 has n_2 individuals. For each GWAS dataset, there are p SNPs as candidate IVs. GWAS1 $\{X_1, G_1\}$ only contains exposure X_1 and SNPs G_1, while GWAS2 $\{Y_2, G_2\}$ only has the information for outcome Y_2 and genetic variants G_2. Ideally, when invalid IVs are treated as covariates and X_2 is observed, the TSLS framework leads to the following two-stage linear model:

$$\begin{aligned} X_1 &= \omega G_1 + \varepsilon_{XY}, \\ Y_2 &= \theta_{XY} X_2 + \alpha G_{2\mathcal{A}_X} + \nu_{XY}, \end{aligned} \qquad \text{(Equation 4)}$$

where ω is the vector of causal effects of IVs on the exposure and ε_{XY} and ν_{XY} are error terms. However, in practice, X_2 is typically unobserved. In such cases, we estimate X_2 using the fitted values from the first-stage regression: $\widehat{X_2} = \widehat{\omega} G_2$, where $\widehat{\omega}$ is obtained via ordinary least squares (OLS) regression of X_1 on all p

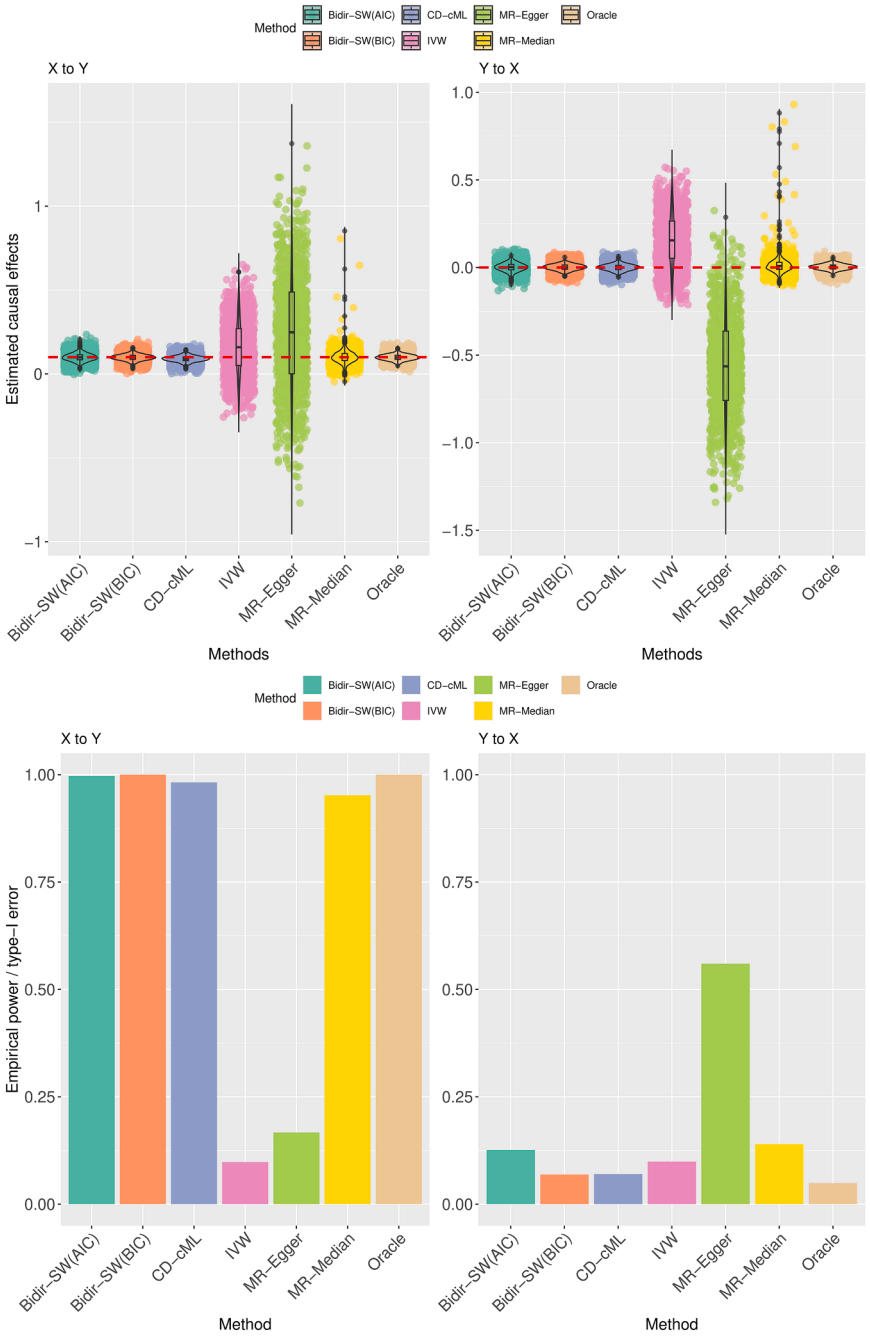

Figure 3. Simulation studies for independent SNPs showing empirical distributions of estimates, empirical type-I error (when $\theta_{XY} = 0$), and power (when $\theta_{XY} = 0.1$) for parameters θ_{XY} and θ_{YX}, with $\xi \sim \mathcal{U}(-0.2, 0.2)$
The true parameter values are set as $\theta_{XY} = 0.1$ and $\theta_{YX} = 0$, respectively.

instruments in G_1. Substituting this estimate into the second stage yields the working model:

$$\begin{aligned} X_1 &= \omega G_1 + \varepsilon_{XY}, \\ Y_2 &= \theta_{XY}\widehat{X_2} + \alpha G_{2\mathcal{A}_X} + \eta_{XY}, \end{aligned} \qquad \text{(Equation 5)}$$

where η_{XY} denotes the modified error term reflecting the use of estimated $\widehat{X_2}$. Previously we proposed a stepwise model selection MR technique[18] to identify invalid IVs using only BIC[19] as the model selection criterion. In this research, we not only consider BIC but also consider AIC,[20] since BIC generally penalizes the number of invalid IVs more than AIC and we would like to see if there will be any significant difference. For each candidate set of invalid IVs C, the corresponding AIC and BIC are given as follows:

$$\begin{aligned} \mathrm{AIC}(C) &= 2|C| + n_2 \log\frac{SSE_{XY(C)}}{n_2}, \\ \mathrm{BIC}(C) &= \log(n_2)|C| + n_2 \log\frac{SSE_{XY(C)}}{n_2}, \end{aligned}$$

where $SSE_{XY(C)}$ is the *sum of squared errors* of the model under the invalid IV set C. The invalid IVs selection procedure can be summarized as follows: we start with an empty invalid IV set and obtain the corresponding BIC or AIC; then we add one invalid IV each time to the model to see if any combination would result in a lower BIC or AIC. Next, we continue previous steps until no more invalid IVs could be added; therefore, we can obtain the estimated $\widehat{\mathcal{A}}_X$, which is the estimated set of invalid IVs for X.

Table 3. Summary of simulation results for independent SNPs, presenting estimated causal effects (θ_{XY} and θ_{YX}), empirical power, type-I error rates, standard deviation (SD), and standard error (SE) for both causal directions

***X* To *Y*, $\theta_{XY} = 0.1$**					***Y* To *X*, $\theta_{YX} = 0$**			
Method	Power	Mean(θ_{XY})	SD	SE	Type-I error	Mean(θ_{YX})	SD	SE
Bidir-SW (AIC)	0.997	9.9793e−02	0.0238	0.0186	0.126	1.5337e−03	0.0241	0.0180
Bidir-SW (BIC)	1.000	9.8242e−02	0.0179	0.0170	0.069	1.6027e−03	0.0171	0.0160
Naive stepwise (AIC)	0.819	1.0503e−01	0.1385	0.0157	0.872	9.2352e−02	0.1611	0.0150
Naive stepwise (BIC)	0.816	1.0456e−01	0.1218	0.0149	0.872	9.1845e−02	0.1603	0.0149
CD-cML	0.982	8.9754e−02	0.0176	0.0186	0.070	7.8123e−05	0.0180	0.0173
IVW	0.098	1.6311e−01	0.1605	0.2030	0.099	1.5721e−01	0.1512	0.1984
MR-Egger	0.167	2.4245e−01	0.3438	0.3100	0.560	−5.6090e−01	0.2860	0.2579
MR-Median	0.952	1.0219e−01	0.0476	0.0238	0.140	1.6161e−02	0.0689	0.0244
Oracle	1.000	9.8942e−02	0.0172	0.0171	0.051	2.3653e−03	0.0162	0.0161

The left side of the table corresponds to the direction $X \to Y$, and the right side corresponds to $Y \to X$.

The details of the TSLS-stepwise approach are provided in Chen et al.[18] In line with the original TSLS-stepwise framework—which is designed for a uni-directional setting, requires at least two valid IVs—we adopt the same assumption in this study, requiring at least two valid IVs for each causal direction.

For the causal direction of $Y \to X$, where we treat Y as exposure and X as the outcome. Similarly, the working model is thereby given by

$$\begin{aligned} Y_1 &= \mu G_1 + \varepsilon_{YX}, \\ X_2 &= \theta_{YX}\widehat{Y_2} + \beta G_{2\mathcal{A}_Y} + \eta_{YX}. \end{aligned} \quad \text{(Equation 6)}$$

In the first stage, we estimate $\widehat{Y_2}$ by $\widehat{Y_2} = \widehat{\mu} G_2$, where $\widehat{\mu}$ is obtained by regressing Y_1 on all p IVs in G_1. For every possible combination of invalid IVs C', the AIC and BIC are given by

$$\begin{aligned} \mathrm{AIC}(C') &= 2|C'| + n_1 \log\frac{SSE_{YX(C')}}{n_1}, \\ \mathrm{BIC}(C') &= \log(n_1)|C'| + n_1 \log\frac{SSE_{YX(C')}}{n_1}, \end{aligned} \quad \text{(Equation 7)}$$

where $SSE_{YX}(C')$ is the sum of squared errors of the model for set C'. Thus we can obtain the estimated $\widehat{\mathcal{A}}_Y$ accordingly.

Common invalid IVs, causal directions, and post-IV-selection inference

The previous section described how to estimate the sets of invalid IVs in both directions, $\widehat{\mathcal{A}}_X$ and $\widehat{\mathcal{A}}_Y$. To find the IVs that are commonly invalid for both directions, we take the intersection of these sets:

$$\widehat{\{g_B\}} = \widehat{\mathcal{A}}_X \cap \widehat{\mathcal{A}}_Y.$$

This intersection identifies the IVs that are invalid in both directions. Consequently, we estimate the set of IVs that remain valid in either direction by taking the complement of $\widehat{\{g_B\}}$, namely $\{g_C\} = \widehat{\{g_B\}}^C$.

By Theorem 1, any instrument $g \in \{g_C\}$ that satisfies $|\mathrm{corr}(g, Y)| < |\mathrm{corr}(g, X)|$ is valid for the direction $X \to Y$; otherwise, g is valid for $Y \to X$. We test each $g \in \{g_C\}$ against this condition and collect all g satisfying $|\mathrm{corr}(g, Y)| < |\mathrm{corr}(g, X)|$ into $\widehat{\{g_X\}}$, which is the estimated set of valid IVs for $X \to Y$. Consequently, the set of valid IVs for $Y \to X$ is given by $\widehat{\{g_Y\}} = \{g_C\} \setminus \widehat{\{g_X\}}$. Having identified valid IVs for each direction, we then apply two-stage least squares (TSLS) separately, using $\widehat{\{g_X\}}$ for $X \to Y$ and $\widehat{\{g_Y\}}$ for $Y \to X$. This procedure can yield the estimated causal effects and the corresponding standard errors for both directions.

Once we have identified the estimated $\widehat{\{g_X\}}$ (X to Y) and $\widehat{\{g_Y\}}$ (Y to X), we can move on to statistical inference. Without loss of generality, we will illustrate the inference for the direction of X to Y as an example; the procedure for Y to X is analogous. Consider valid IVs only, we use G_{1X} and G_{2X} to denote the matrices for valid IVs in the first sample and the second sample, respectively. In particular, Pacini and Windmeijer et al.[21] showed that the estimator $\widehat{\theta}_{XY}$ for the causal effect θ_{XY} satisfies

$$\sqrt{n_2}(\widehat{\theta}_{XY} - \theta_{XY}) \xrightarrow{d} N(0, \sigma^2_{XY}),$$

Table 4. Summary of simulation results for independent SNPs, presenting the true discovery rate, false discovery rate, number of true valid IVs identified, and total number of valid IVs discovered for the direction *X*→*Y*

***X* To *Y*, $\theta_{XY} = 0.1$**				
Method	TDR	FDR	Mean of true valid IVs (SD)	Mean of total valid IVs (SD)
Naive stepwise (AIC)	0.8534	0.2157	8.53 (1.08)	10.97 (1.59)
Bidir-SW (AIC)	0.8571	0.0034	8.57 (1.11)	8.60 (1.12)
Naive stepwise (BIC)	0.9948	0.1898	9.95 (0.22)	12.38 (1.18)
Bidir-SW (BIC)	0.9984	0.0103	9.98 (0.13)	10.10 (0.35)

Table 5. Summary of simulation results for independent SNPs, presenting the true discovery rate, false discovery rate, number of true valid IVs identified, and total number of valid IVs discovered for the direction *Y* to *X*

***Y* To *X*, $\theta_{YX} = 0$**

Method	TDR	FDR	Mean of true valid IVs (SD)	Mean of total valid IVs (SD)
Naive stepwise (AIC)	0.7467	0.2557	7.47 (1.09)	10.03 (0.18)
Bidir-SW (AIC)	0.8032	0.0036	8.03 (1.18)	8.06 (1.18)
Naive stepwise (BIC)	0.7549	0.2534	7.55 (1.09)	10.12 (0.34)
Bidir-SW (BIC)	0.9964	0.0098	9.96 (0.19)	10.07 (0.38)

where the variance σ^2_{XY} is estimated by

$$\widehat{\sigma}^2_{XY} = \widehat{Var}(\widehat{\theta}_{XY}) = \widehat{\sigma}^2_2 \left(\widehat{X_2}^T \widehat{X_2}\right)^{-1} + \widehat{\theta}^2_{XY} \widehat{\sigma}^2_1 \widehat{C} \left(G^T_{1X} G_{1X}\right)^{-1} \widehat{C}^T,$$

where

$$\widehat{\sigma}^2_2 = \frac{||Y_2 - \widehat{X_2}\widehat{\theta}_{XY}||^2}{n_2}$$

is the estimated variance given by the second sample, and

$$\widehat{\sigma}^2_1 = \frac{||X_1 - \widehat{X_1}||^2}{n_1}$$

is the variance estimated in the first stage, with $\widehat{X_1}$ being the fitted value of X_1. Additionally, the matrix

$$\widehat{C} = \left(\widehat{X_2}^T \widehat{X_2}\right)^{-1} \widehat{X_2}^T G_{2X}$$

can be obtained from the second sample.

Adaption for GWAS summary statistics

Consider independent GWAS summary-level datasets for traits X and Y, the marginal association effect sizes ($\widehat{\beta_X}$ and $\widehat{\beta_Y}$) and standard errors (SE($\widehat{\beta_X}$) and SE($\widehat{\beta_Y}$)) are provided. When only GWAS summary statistics, instead of individual GWAS datasets, are given due to possible privacy concerns, we can therefore access the marginal SNP-trait effect estimates and the variances from GWAS summary statistics. In this section, the extension of our proposed approach to GWAS summary statistics is discussed.

Invalid IV selection under GWAS summary statistics

For causal direction $X \to Y$, we adopt results from previous studies, such as Knutson et al.,[15] Chen et al.,[18] and Pattee and Pan,[22] to estimate $\widehat{G_2^T G_2}$, $\widehat{G_2^T Y_2}$, and $\widehat{Y_2{}^T Y_2}$ by using GWAS summary statistics and a reference panel provided by the 1000 Genomes Project.[23] Without the loss of generality, we propose the following matrix $\widehat{\mathcal{W}}$:

$$\widehat{\mathcal{W}} = \begin{bmatrix} 1(\textbf{invalid})_1 & 0 & \cdots & 0 & \widehat{\omega_1} \\ 0 & 1(\textbf{invalid})_2 & \cdots & 0 & \widehat{\omega_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1(\textbf{invalid})_p & \widehat{\omega_p} \end{bmatrix}.$$

The matrix $\widehat{\mathcal{W}}$ contains the causal effects of each of the IVs on the exposures, which can be obtained in the first stage using OLS, are $(\widehat{\omega_1}, \widehat{\omega_2}, \cdots, \widehat{\omega_p})^T$, and indicators $1(\textbf{invalid})$ for each of the SNPs. If the i-th SNP is an invalid IV, then the value of this indictor is 1, otherwise the value returns 0. The estimated coefficient $\widehat{\phi}_{XY} = [\widehat{\theta}_{XY}, \widehat{\alpha}]^T$ is thereby given as

$$\widehat{\phi}_{XY} = \begin{pmatrix} \widehat{\theta}_{XY} \\ \widehat{\alpha} \end{pmatrix} = \left(\widehat{\mathcal{W}}^T G_2^T G_2 \widehat{\mathcal{W}}\right)^{-1} \widehat{\mathcal{W}}^T G_2^T Y_2.$$

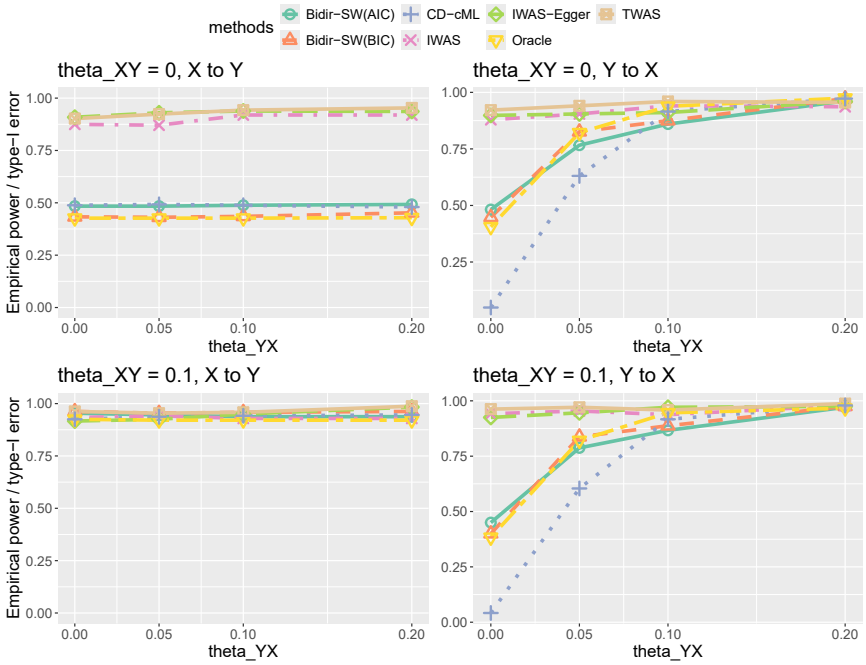

Figure 4. Simulation studies for correlated SNPs comparing empirical type-I error (when $\theta_{XY} = 0$) and power (when $\theta_{XY} = 0.1$) across different values of θ_{YX} (0, 0.05, 0.1, 0.2)

Results are shown for different methods, including the oracle condition. Sample sizes are $n_1 = n_2 = 50,000$, and ξ are sampled from $U(-0.2, 0.2)$. The upper panel corresponds to $\theta_{XY} = 0$ (type-I error), and the lower panel corresponds to $\theta_{XY} = 0.1$ (power).

Table 6. Simulation studies for correlated SNPs comparing empirical type-I error of various methods across different values of θ_{YX} (0, 0.05, 0.10, 0.20) with sample sizes $n_1 = n_2 = 50,000$ and $\xi \sim U(-0.2, 0.2)$, when $\theta_{XY} = 0$

	θ_{YX}			
Method	**0.00**	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.498	0.500	0.502	0.511
Bidir-SW (BIC)	0.455	0.452	0.455	0.471
CD-cML	0.494	0.498	0.499	0.492
IWAS	0.895	0.885	0.921	0.921
IWAS-Egger	0.916	0.924	0.930	0.946
Oracle	0.445	0.445	0.445	0.445
TWAS	0.918	0.939	0.949	0.958
Method	θ_{YX}			
	0.00	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.503	0.770	0.850	0.961
Bidir-SW (BIC)	0.465	0.833	0.858	0.980
CD-cML	0.045	0.629	0.921	0.975
IWAS	0.904	0.912	0.936	0.932
IWAS-Egger	0.897	0.902	0.918	0.953
Oracle	0.432	0.838	0.919	0.981
TWAS	0.940	0.948	0.958	0.957

The top table corresponds to the direction $X \to Y$, and the bottom table corresponds to $Y \to X$.

Assume that the set of k invalid IVs is C, then the SSE_{XY} under C becomes $Y_2^T Y_2 - \widehat{\phi}_{XY} \widehat{\mathcal{W}}^T G_2^T Y_2$. The corresponding AIC and BIC are given as follows:

$$\mathrm{AIC}(C) = n_2 \log \frac{SSE_{XY}}{n_2} + 2k,$$

and

$$\mathrm{BIC}(C) = n_2 \log \frac{SSE_{XY}}{n_2} + \log(n_2)k.$$

Likewise, for causal direction $Y \to X$, it is also possible for us to derive AIC and BIC without accessing individual-level GWAS. Therefore we can perform IV selection using only summary-level GWAS data.

Sample correlation calculation under GWAS summary statistics

The determination of causal directions depends on sample correlations. For two independent GWAS summary statistics datasets, according to Xue et al.,[9] the sample correlations of traits X and Y with each SNP g are given by

$$\begin{aligned} \mathrm{corr}(g, X) &= \frac{\widehat{\beta_X}}{\sqrt{\widehat{\beta_X}^2 + (n_1 - 2)\mathrm{Var}(\widehat{\beta_X})}}, \\ \mathrm{corr}(g, Y) &= \frac{\widehat{\beta_Y}}{\sqrt{\widehat{\beta_Y}^2 + (n_2 - 2)\mathrm{Var}(\widehat{\beta_Y})}}, \end{aligned} \qquad \text{(Equation 8)}$$

where $\mathrm{Var}(\widehat{\beta_X})$ and $\mathrm{Var}(\widehat{\beta_Y})$ can be estimated by $\mathrm{SE}^2(\widehat{\beta_X})$ and $\mathrm{SE}^2(\widehat{\beta_Y})$; both are provided in GWAS summary statistics. In this way, we can determine the valid IVs for each direction as long as GWAS summary statistics are provided.

Simulation studies

Simulation studies for independent SNPs

We follow the main simulation settings as mentioned in Xue and Pan.[16] We first generated two independent, non-overlapping GWAS datasets with sample sizes $n_1 = n_2 = 50,000$ at the individual level, based on the true causal model illustrated in Figure 1 with two continuous traits X and Y. We designated 10, 10, and five independent SNPs in sets $\{g_X\}, \{g_Y\}$ and $\{g_B\}$, respectively. α and β are the causal effects from $\{g_X\}$ and $\{g_X\}$ to traits X and Y. We sampled α and β from $\mathcal{U}((-0.3, -0.2) \cup (0.2, 0.3))$ independently. The causal effects of $\{g_B\}$ to traits X and Y, say γ and η, were also sampled independently from $\mathcal{U}((-0.3, -0.2) \cup (0.2, 0.3))$. The causal effects of hidden confounder U on traits X and Y were set to be $\theta_{UX} = \theta_{UY} = 1$. The minor allele frequencies (MAFs) for all SNPs were set to be 0.3. The random errors ε_X and ε_Y were sampled from $\mathcal{N}(0, 1)$ distributions, while ε_U were sampled from $\mathcal{N}(0, 2)$. For the true causal effects θ_{XY} and θ_{YX}, the results from the following combinations are presented: true causal effect sizes for $\theta_{XY} = 0$ and 0.1, with θ_{YX} taking values from 0, 0.05, 0.1, and 0.2. In order to create more correlated pleiotropic effects to the simulated data, we sampled ξ values from uniform distribution $\mathcal{U}(-0.2, 0.2)$, therefore the IVs in $\{g_B\}$ are correlated with the hidden confounder U. Notice that this is also the case when the InSIDE (INstrument Strength Independent of Direct Effect) assumption[11] is

Table 7. Simulation studies for correlated SNPs comparing empirical power of various methods across different values of θ_{YX} (0, 0.05, 0.10, 0.20) with sample sizes $n_1 = n_2 = 50{,}000$ and $\xi \sim U(-0.2, 0.2)$, when $\theta_{XY} = 0.1$

	θ_{YX}			
Method	**0.00**	**0.05**	**0.10**	**0.20**
Bidir-SW(AIC)	0.953	0.949	0.945	0.948
Bidir-SW(BIC)	0.969	0.967	0.966	0.962
CD-cML	0.940	0.942	0.941	0.946
IWAS	0.922	0.934	0.929	0.929
IWAS-Egger	0.914	0.923	0.944	0.974
Oracle	0.901	0.900	0.901	0.900
TWAS	0.947	0.956	0.957	0.981
Method	θ_{YX}			
	0.00	**0.05**	**0.10**	**0.20**
Bidir-SW (AIC)	0.486	0.760	0.867	0.970
Bidir-SW (BIC)	0.448	0.808	0.883	0.984
CD-cML	0.058	0.623	0.920	0.981
IWAS	0.935	0.943	0.936	0.976
IWAS-Egger	0.921	0.945	0.954	0.978
Oracle	0.433	0.832	0.921	0.981
TWAS	0.958	0.968	0.962	0.982

The top table corresponds to the direction $X \to Y$, and the bottom table corresponds to $Y \to X$.

violated. We also provided the simulation results when ξ is 0; and this is when the IVs in $\{g_B\}$ are not correlated with the hidden confounder. For each simulation set-up, we conducted 1,000 simulations.

We present the simulation results of empirical type-I errors and power in Figure S1 (when $\xi = 0$) and Figure 2 (when ξ is non-zero) for all simulation settings. At the meantime, we also provide the estimates of the oracle set-ups where the valid IVs for each direction are known, for comparison purposes. For $\xi \sim \mathcal{U}(-0.2, 0.2)$, the numerical results of the simulations are provided in Tables 1 and 2. Our proposed Bidir-SW approaches, especially the BIC option, control the type-I error while maintaining relatively high statistical power when compared with other established MR methods, such as MR-Egger, IVW, and Weighted Median,[24] as well as the recently developed bi-directional MR approach CD-cML by Xue et al.[16] In these simulation results, we could also see that when ξ is 0, both CD-cML and our opposed BIC approach generate the closest empirical power (or type-I error) comparing to the oracle settings. When ξ is non-zero, our BIC approach has the best performance. We also notice that our AIC approach has similar performance to CD-cML under this set-up.

We continue to consider the following parameter combinations: (1) true values are $\theta_{XY} = 0.1$ and $\theta_{YX} = 0$ and (2) true values are $\theta_{XY} = 0$ and $\theta_{YX} = 0$, are also provided. For setting (1) $\theta_{XY} = 0.1$ and $\theta_{YX} = 0$, the simulation results of empirical distributions of both estimated θ_{XY} and θ_{YX}, plus estimated empirical power/type-I errors, can be found in Figure 3 and Table 3; for setting (2) $\theta_{XY} = 0$ and $\theta_{YX} = 0$, we provide graphical presentations in Figure S2. We also include the oracle results as a benchmark for comparison. We also observe that while CD-cML performs well, our proposed BIC approach has the best overall performance when compared with the oracle results.

We also present the mean and the standard deviation of the number of true and total estimated valid IVs, using our proposed Bidir-SW methods and the naive stepwise approach as presented in Chen et al.,[18] for $X \to Y$ and $Y \to X$, as shown in Tables 4 and 5, respectively. For the naive stepwise approach, we apply Steiger's filtering before performing stepwise selection, as recommended by the reviewers. We assume the total number of correctly identified valid IVs in 1,000 simulations is N_T and the number of total incorrectly identified valid IVs is N_F. We also define *true selection rate* as N_T/N_0 and *false selection rate* is given by $N_F/(N_F + N_T)$. In Tables 4 and 5 we also provide *true selection rate* (TDR) and *false selection rate* (FDR), besides the mean and the standard deviation (SD) of the number of true estimated valid IVs and total valid IVs. The simulation results for the naive stepwise approach are also presented in Table 3. In comparison, our method yields more accurate and reliable inference for both θ_{XY} and θ_{YX}. Specifically, it achieves higher empirical power when a true causal effect is present and better control of type-I error under the null. Overall, the results highlight that our approach not only

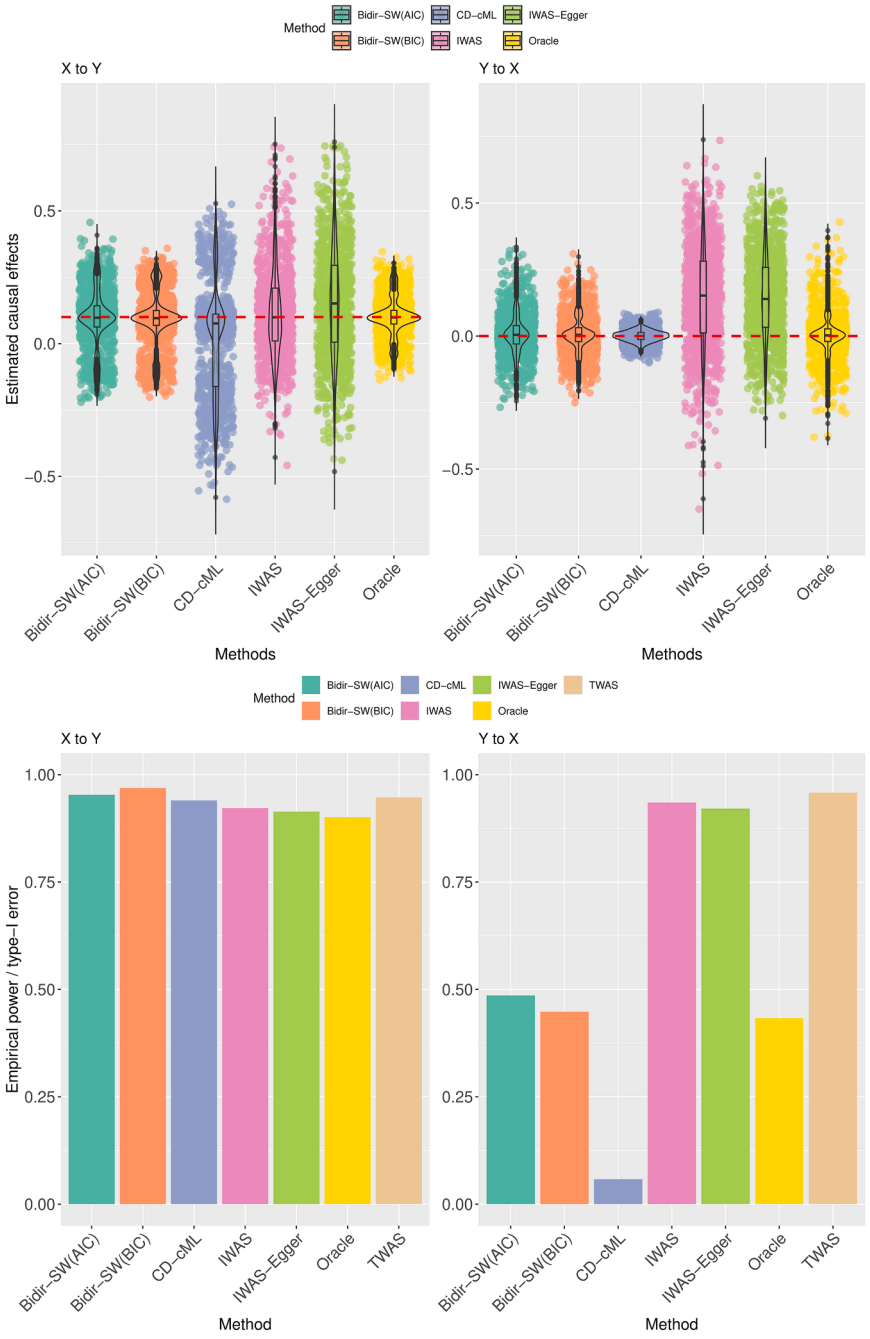

Figure 5. Simulation studies for correlated SNPs showing empirical distributions of estimates, empirical type-I error (when $\theta_{XY} = 0$), and power (when $\theta_{XY} = 0.1$) for parameters θ_{XY} and θ_{YX}, with $\xi \sim \mathcal{U}(-0.2, 0.2)$
The true parameter values are set as $\theta_{XY} = 0.1$ and $\theta_{YX} = 0$, respectively.

enhances the selection of (in)valid IVs but also improves the robustness of statistical inference compared with the naive approach.

Simulation studies for correlated SNPs

To evaluate the performance of our methods with LD correlations in real TWAS applications, we repeated the simulation in section 3.1 of the supplemental information except that we considered correlated SNPs instead of independent SNPs. In order to do so, we first created an uncorrelated matrix using binomial random values based on the MAF of 0.3; next, we introduced a correlation structure by constructing a correlation matrix where half of the IVs have weak correlations, with an off-diagonal value of 0.2. We thus generated a mix of correlated and uncorrelated SNPs for analysis. We also conducted 1,000 simulations for each scenario similar to what we illustrated in the previous section, with GWAS sample sizes of $n_1 = n_2 = 50,000$. Similarly, we account for correlated pleiotropic effects by adjusting the values of ξ.

When $\xi = 0$, we present the simulation results in Figure S3. When ξ is drawn from $\mathcal{U}(-0.2, 0.2)$, we provide the simulation results in Figure 4 and Tables 6 and 7, to illustrate the empirical type-I error and power. These results show that our proposed methods effectively control type-I error under correlated SNPs and perform better compared with TWAS,[14] IWAS,[15] and IWAS-Egger.[15] Although CD-cML is designed for uncorrelated IVs, we still provide the corresponding results of CD-cML when IVs are correlated. The simulation results indicate that our approaches, especially the BIC option, have the best performance among all methods compared with oracle settings. Not surprisingly, the CD-cML approach does not perform well under SNP correlations.

We continue to provide the simulation results of empirical distributions of both estimated θ_{XY} and θ_{YX}. The results are shown for the following parameter combinations: (1) true values $\theta_{XY} = 0.1$ and $\theta_{YX} = 0$, simulation results are presented in Figure 5 and Table 8; (2) true values $\theta_{XY} = 0$ and $\theta_{YX} = 0$, shown in Figure S4. Our proposed approaches also have better performance among all existing approaches such as IWAS, comparing to the oracle scenario.

We continue to present the mean, standard deviation, TDR, and FDR of the number of true estimated valid IVs and total valid IVs for both directions in Tables 9 and 10. These simulations also indicate that our proposed method performs better than the naive stepwise approach.

Table 8. Summary of simulation results for correlated SNPs, presenting estimated causal effects (θ_{XY} and θ_{YX}), empirical power, type-I error rates, standard deviation (SD), and standard error (SE) for both causal directions

***X* To *Y*, $\theta_{XY} = 0.1$**					***Y* To *X*, $\theta_{YX} = 0$**			
Method	Power	Mean	SD	SE	Type-I error	Mean	SD	SE
Bidir-SW (AIC)	0.953	0.0931	0.1108	0.0187	0.486	0.0086	0.0806	0.0184
Bidir-SW (BIC)	0.969	0.0908	0.1007	0.0168	0.448	0.0086	0.0721	0.0163
CD-cML	0.940	0.0162	0.2128	0.0275	0.058	0.00004	0.0200	0.0195
IWAS	0.922	0.1164	0.1664	0.0109	0.935	0.1449	0.1948	0.0114
IWAS-Egger	0.914	0.1539	0.2083	0.0137	0.921	0.1449	0.1640	0.0126
TWAS	0.947	N/A	N/A	N/A	0.958	N/A	N/A	N/A
Oracle	0.901	0.1001	0.0753	0.0175	0.433	0.0048	0.0890	0.0162

The left side of the table corresponds to the direction $X \to Y$, and the right side corresponds to $Y \to X$.

Additional simulations

We conduct additional simulation studies to evaluate the performance of our proposed methods in the presence of weak instrumental variables (IVs), both for independent and correlated SNPs. We also compare our approaches against existing MR and TWAS methods. Details of these simulations are provided in section 3.2 of the supplemental materials. Our results demonstrate that our proposed methods provide reliable power estimates and robust IV selection, even in the presence of weak IVs.

We also adjusted our previous simulation settings by increasing the proportion of invalid IVs, as suggested by the reviewers. Simulations are repeated to evaluate the TDR and the FDR; the simulation results are also presented in section 3.2 of the supplemental materials.

Real data applications

Olink CVD proteins GWAS summary data

The GWAS summary statistics for Olink CVD-I proteins are derived from 14 studies from individuals of European ancestry. Folkersen et al.[25] conducted a genome-wide meta-analysis of 90 cardiovascular-related proteins in more than 30,000 individuals from these 14 studies. The summary statistics for all 90 proteins are publicly available at https://zenodo.org/records/2615265. This study includes many proteins that are well-established as prognostic biomarkers or drug targets, analyzed using the Olink Proximity Extension Assay CVD-I panel across the combined cohort.

CVD proteins and SLP

We first examine the bi-directional relationships between each of the 90 CVD proteins and sleep duration (SLP), focusing on potential causal connections between these two traits. The GWAS summary statistics for SLP is provided by Jones et al.,[26] with more than 120,000 White British individuals from the UK Biobank study. For each protein as the exposure, we filtered a set of SNPs that are either associated with SLP or the protein with p value less than 5×10^{-8}, to reduce weak instrument bias. We also performed LD clumping of the genetic variants applying a criterion of low LD of $r^2 \leq 0.2$ and clumping window being kb $= 2000$, using 1000 Genomes European population reference panel. We only keep proteins that share at least three SNPs with SLP, resulting 75 of them included in the analysis. We compared our results with established TWAS methods such as traditional TWAS, IWAS, and IWAS-Egger.

For causal direction from proteins to sleep duration, as presented in Figure 6, we present the significant proteins identified by each method, after Bonferroni-corrected significance level $\alpha = 0.05/75$. We notice that TWAS does not find any protein that is significantly associated with SLP, and for all other methods except for traditional TWAS, SPON1 is commonly detected as significantly causal to SLP. We provide our findings along with literature support information in Table 11. For direction from SLP to each of the proteins, we did not identify any significant causal relationships using our proposed approach.

Table 9. Summary of simulation results for correlated SNPs, presenting the true discovery rate, false discovery rate, number of true valid IVs identified, and total number of valid IVs discovered for the direction $X \to Y$

***X* To *Y*, $\theta_{XY} = 0.1$**				
Method	TDR	FDR	Mean of true valid IVs (SD)	Mean of total valid IVs (SD)
Naive stepwise (AIC)	0.6384	0.2848	6.38 (2.19)	8.83 (2.37)
Bidir-SW(AIC)	0.7888	0.0337	7.89 (1.45)	8.21 (1.62)
Naive stepwise (BIC)	0.7619	0.2479	7.62 (2.25)	10.07 (2.43)
Bidir-SW(BIC)	0.9187	0.0429	9.19 (1.21)	9.66 (1.53)

Table 10. Summary of simulation results for correlated SNPs, presenting the true discovery rate, false discovery rate, number of true valid IVs identified, and total number of valid IVs discovered for the direction $Y \to X$

Y **To** X, $\theta_{YX} = 0$

Method	TDR	FDR	Mean of true valid IVs (SD)	Mean of total valid IVs (SD)
Naive stepwise (AIC)	0.7423	0.2849	7.42 (1.10)	10.43 (0.77)
Bidir-SW(AIC)	0.7748	0.0725	7.74 (1.35)	8.47 (1.77)
Naive stepwise (BIC)	0.7614	0.2843	7.61 (1.06)	10.71 (0.96)
Bidir-SW(BIC)	0.9610	0.0744	9.61 (0.74)	10.52 (1.53)

CVD proteins and BMI

Next we study the bi-directional relationships between the CVD proteins and BMI. The GWAS summary statistics of BMI[36] can be found at https://gwas.mrcieu.ac.uk/datasets/ieu-a-835/. With similar filtering and clumping procedures, we include 88 proteins in the final analysis. For causal direction from proteins to BMI, we present the numbers of significant proteins identified by each method after Bonferroni correction in Figure 7, with more detailed information as well as literature findings in Table 12. For direction from BMI to each of the proteins, there are no significant causal relationships identified by the proposed methods either.

CVD proteins and CRP

The GWAS summary data for C-reactive protein (CRP) levels is provided by Ligthart et al.[50] and publicly available at https://gwas.mrcieu.ac.uk/datasets/ieu-b-35/. For the bi-directional relationships between the proteins and CRP, we ended up with 76 proteins to be included in further analysis. For causal direction from proteins to CRP, we present the results in Figure 8 and Table 13, with literature validation information. For direction from CRP to each of the proteins, we identified several significant causal relationships by the proposed methods, after Bonferroni correction. We present the results in Table 14, and none of the three proteins are identified as significant by TWAS, IWAS, or IWAS-Egger. We also notice that protein FS and CRP are bi-directionally significant to each other, according to our Bidir-SW(AIC) approach.

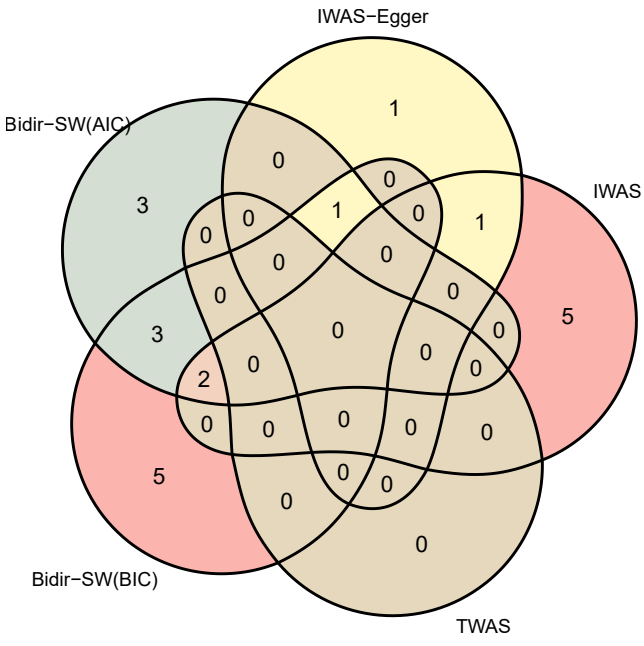

Figure 6. Venn diagram illustrating the significant proteins identified as causal for sleep duration (SLP) across five different methods, using a Bonferroni-corrected significance threshold of $\alpha = 0.05/75$

Causal relationships among SLP, BMI, and CRP

We also consider the combinations of the following three traits: SLP, BMI, and CRP, by choosing two of them as exposure-outcome pairs. We then conduct MR analysis to identify the causal relationship among these pairs. For the existing MR approaches, we applied the `TwoSampleMR` R package for LD clumping and the following harmonization procedures, and the `MendelianRandomization` R package to implement IVW, MR-Egger, and Weighted Median approaches. For the CD-cML approach to perform bi-directional causal inference mentioned in Xue and Pan,[16] we applied the `BiDirectCausal` R package available on GitHub. All genetic variants included in the analysis are pruned and clumped at $r^2 < 0.001$ to avoid LD correlations. We also applied the default clumping window for 2,000 kb. The data clumping is performed using the 1000 Genomes European population reference panel.

BMI and CRP

We present the results of causal estimates between CRP and BMI using different MR methods in Figure 9. In the outcome and exposure GWAS datasets, only SNPs with p values $\leq 5 \times 10^{-7}, 5 \times 10^{-6}, 5 \times 10^{-5}$ in both exposure and outcome data are retained in further analysis. The bi-directional MR analysis between CRP and BMI reveals a significant causal effect of CRP on BMI by Bidir-SW (AIC and BIC) and Weighted Median. Especially when p-value threshold is 5×10^{-7}, the strong significant positive causal estimates indicate that increased CRP levels may contribute to higher BMI, which agrees with previous findings that higher BMI is generally associated with increased CRP levels.[73,74] However, IVW, MR-Egger, and CD-cML do not find strong evidence for this relationship, as their results are non-significant (NS).

Table 11. Proteins identified as significantly causal for sleep duration (SLP) by different methods, using a Bonferroni-corrected significance threshold of $\alpha = 0.05/75$

Proteins	Bidir-SW (AIC)	Bidir-SW (BIC)	IWAS-Egger	IWAS	TWAS	Validation
EN-RAGE	Yes	Yes	No	No	No	
Gal-3	Yes	No	No	No	No	Smith et al.[27]
MPO	Yes	Yes	No	No	No	Zhu et al.[28]
PSGL-1	Yes	Yes	No	No	No	Díaz-García et al.[29]
RETN	Yes	No	No	No	No	
SPON1	Yes	Yes	Yes	Yes	No	Theorell-Haglöw et al.[30]
ST2	Yes	Yes	No	No	No	
TIE2	Yes	No	No	No	No	Gozal et al.[31]
TNFSF14	Yes	Yes	No	No	No	
IL-6RA	No	Yes	No	No	No	Wang et al.[32]
MMP-10	No	Yes	No	No	No	Franczak et al.[33]
RAGE	No	Yes	No	No	No	Olejarz et al.[34]
FAS	No	No	Yes	Yes	No	
IL-27	No	No	Yes	Yes	No	
FABP4	No	No	No	Yes	No	Català et al.[35]

Literature validation sources for the identified proteins are provided.

We continue to examine the causal effect of BMI on CRP. Both Bidir-SW methods (AIC and BIC) suggest significant effects of BMI on CRP under p-value thresholds of 5×10^{-7} and 5×10^{-6}. Methods such as Weighted Median and CD-cML also discover significant causal effects under the threshold of 5×10^{-6}, while IVW and MR-Egger do not find anything significant. Overall, these results suggest potential bi-directional causality between CRP and BMI.

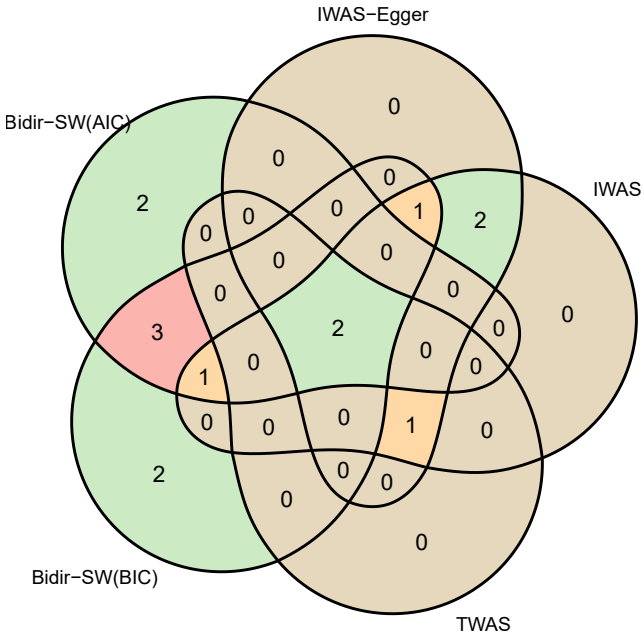

Figure 7. Venn diagram illustrating the significant proteins identified as causal for body mass index (BMI) across five different methods, using a Bonferroni-corrected significance threshold

SLP and CRP

We also explore the bi-directional relationship between CRP and SLP using different MR techniques; the results are presented in Figure 10. In order to include enough SNPs for analysis, we filtered SNPs with more relaxed p-value thresholds $5\times 10^{-5}, 5\times 10^{-4}$, and 1×10^{-3} for both exposure and outcome data. We first assess the causal effect of CRP on sleep duration. Under all p-value thresholds, Bidir-SW (AIC) and Bidir-SW (BIC) show significant causal relationships. Weighted Median also finds significant relationships when p-value threshold is 0.001, while other MR methods, including IVW, MR-Egger, and CD-cML, fail to show any strong evidence.

We then evaluate the effect of sleep duration on CRP. Bidir-SW (AIC and BIC) and Weighted Median identify significant causal relationships across all p-value thresholds, while IVW and MR-Egger do not find anything statistically significant. When the p-value threshold is 5×10^{-5}, CD-cML also detects a significant causal effect.

SLP and BMI

Finally, we examine the bi-directional relationship between SLP and BMI using different methods; the results are presented in Figure 11. Similarly, we filtered SNPs with more relaxed p thresholds 5×10^{-4} and 5×10^{-3} for both exposure and outcome GWAS summary data. When examining the effect of BMI on sleep duration, Bidir-SW (AIC and BIC), Weighted Median, and CD-cML under p-value threshold of 5×10^{-4} reveal significant causal effects. Other methods such as IVW and MR-Egger do not identify anything significant.

Table 12. Proteins identified as significantly causal for body mass index (BMI) by different methods, using a Bonferroni-corrected significance threshold

Proteins	Bidir-SW (AIC)	Bidir-SW (BIC)	IWAS-Egger	IWAS	TWAS	Validation
CXCL6	Yes	Yes	No	No	No	Anders et al.[37]
Dkk-1	Yes	No	No	No	No	Ali et al.[38]
FABP4	Yes	Yes	No	Yes	No	Queipo-Ortuño et al.[39]
FS	Yes	Yes	Yes	Yes	Yes	Perakakis et al.[40]
HB-EGF	Yes	Yes	No	No	No	Matsumoto et al.[41]
LOX-1	Yes	No	No	No	No	Nomata et al.[42]
MPO	Yes	Yes	Yes	Yes	Yes	Qaddoumi et al.[43]
RAGE	Yes	Yes	No	No	No	Gutowska et al.[44]
MMP-12	No	Yes	Yes	Yes	No	Wang et al.[45]
MMP-3	No	Yes	No	No	No	Raman et al.[46]
PSGL-1	No	Yes	No	No	No	
ADM	No	No	Yes	Yes	No	Vila et al.[47]
CHI3L1	No	No	Yes	Yes	No	Ahangari et al.[48]
LEP	No	No	Yes	Yes	Yes	Paracchini et al.[49]

Literature validation sources for the identified proteins are provided.

When considering the effect of sleep duration on BMI, Bidir-SW (AIC and BIC) and Weighted Median demonstrate significant causal effects. MR-Egger also identifies a significant positive effect under p-value threshold of 5×10^{-3}, whereas IVW and CD-cML fail to find any significant effects.

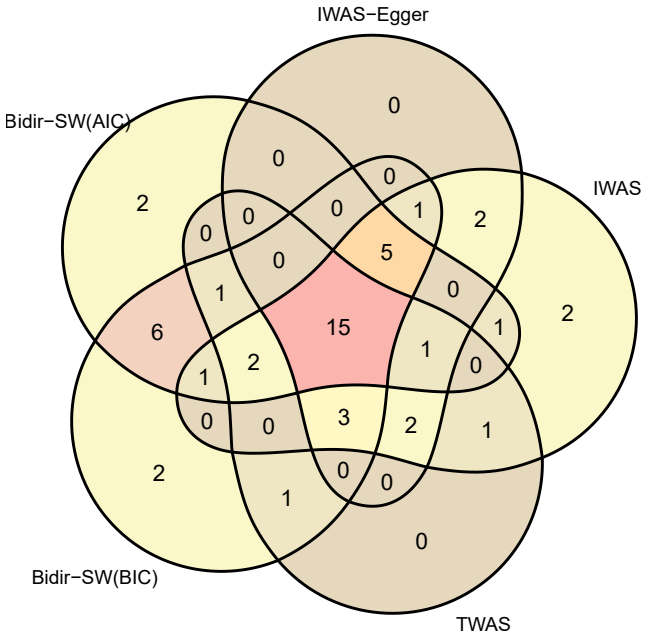

Figure 8. Venn diagram illustrating the significant proteins identified as causal for C-reactive protein (CRP) levels across five different methods, using a Bonferroni-corrected significance threshold
The analysis includes 76 proteins after filtering.

Conclusions and discussions

In this research, we provided a bi-directional MR/TWAS framework for causal direction estimates and inferences between two genetic traits. Our methods can be seen as the extension of our previously developed approach,[18] combined with Steiger's method.[8] We have also demonstrated that our method is robust in the presence of common invalid IVs for both directions through simulation studies. Through both simulation studies and real data applications, our method can always give accurate causal effect estimates as well as statistical power, outperforming traditional methods such as MR-Egger, IVW, and TWAS in the presence of some invalid IVs.

The proposed Bidir-SW can be viewed as a two-stage approach. We first apply stepwise model selection procedures based on BIC or AIC to find valid IVs for both directions; this step is particularly useful in bi-directional MR since the same genetic variant may act as a valid IV for one direction but an invalid IV for the other. Next, we proceed to differentiate the valid IVs for both of the directions, respectively. Our approach can make sure that only valid IVs are included in the final analysis while accounting for correlated IVs, thereby improving statistical power as demonstrated in our simulation studies. This is one of the advantages of the proposed methods, since recent advances in bi-directional MR, such as Xue and Pana,[16] do not take correlations of the SNPs into account.

Our approach represents a data-driven investigation into potential reciprocal influences. As such, the bi-directional effects discovered should be viewed as compelling hypotheses that warrant further examination. Future

Table 13. Proteins identified as significantly causal for C-reactive protein (CRP) levels by different methods

Proteins	Bidir-SW (AIC)	Bidir-SW (BIC)	IWAS-Egger	IWAS	TWAS	Validation
ADM	Yes	Yes	No	No	No	Ehlenz et al.[51]
AGRP	Yes	Yes	Yes	Yes	Yes	
CA-125	Yes	Yes	Yes	Yes	No	Hefler-Frischmuth et al.[52]
CD40	Yes	Yes	No	Yes	No	
CSF-1	Yes	Yes	No	No	No	Lindqvist et al.[53]
CX3CL1	Yes	No	No	No	No	
CXCL16	Yes	Yes	Yes	Yes	Yes	Lehrke et al.[54]
FS	Yes	No	Yes	Yes	Yes	
GDF-15	Yes	Yes	No	Yes	Yes	Hassanzadeh Daloee et al.[55]
HB-EGF	Yes	Yes	Yes	Yes	Yes	
HGF	Yes	Yes	No	No	Yes	Torres et al.[56]
IL-1ra	Yes	Yes	Yes	Yes	Yes	Berger et al.[57]
IL-6RA	Yes	Yes	Yes	Yes	Yes	Del Giudice & Gangestad[58]
IL16	Yes	Yes	No	No	No	
LOX-1	Yes	Yes	Yes	Yes	Yes	Shih et al.[59]
mAmP	Yes	Yes	No	Yes	Yes	
MPO	Yes	Yes	Yes	Yes	Yes	
PAPPA	Yes	Yes	Yes	Yes	No	
PECAM-1	Yes	Yes	Yes	Yes	Yes	
PlGF	Yes	No	No	No	No	Zakiyanov et al.[60]
PTX3	Yes	Yes	Yes	Yes	Yes	Wirestam et al.[61]
RETN	Yes	Yes	No	No	No	
SELE	Yes	Yes	Yes	Yes	Yes	
ST2	Yes	Yes	Yes	Yes	Yes	Ekart et al.[62]
t-PA	Yes	No	No	Yes	No	Singh et al.[63]
TF	Yes	Yes	Yes	Yes	Yes	Fay[64]
TM	Yes	Yes	Yes	Yes	Yes	
TNF-R1	Yes	Yes	Yes	Yes	Yes	Camargo et al.[65]
TRAIL-R2	Yes	Yes	Yes	Yes	Yes	Secchiero et al.[66]
TRAIL	Yes	Yes	Yes	Yes	Yes	Secchiero et al.[66]
TRANCE	Yes	Yes	Yes	Yes	Yes	
U-PAR	Yes	Yes	Yes	Yes	Yes	
VEGF-A	Yes	Yes	No	No	No	Chen et al.[67]
VEGF-D	Yes	Yes	No	No	No	
CCL3	No	Yes	No	No	No	Montecucco et al.[68]
CCL4	No	Yes	Yes	Yes	Yes	Montecucco et al.[68]
Gal-3	No	Yes	No	No	Yes	Synowski et al.[69]
GAL	No	Yes	Yes	Yes	Yes	
IL-27	No	Yes	Yes	Yes	No	Cui et al.[70]
CHI3L1	No	No	Yes	Yes	No	
IL-6	No	No	Yes	Yes	Yes	

(*Continued on next page*)

Table 13. ***Continued***

Proteins	Bidir-SW (AIC)	Bidir-SW (BIC)	IWAS-Egger	IWAS	TWAS	Validation
OPG	No	No	Yes	Yes	Yes	Rhee et al.[71]
PSGL-1	No	No	Yes	Yes	No	
SCF	No	Yes	Yes	Yes	Yes	
MMP-3	No	Yes	No	No	No	Montero et al.[72]
MMP-7	No	No	No	Yes	No	Montero et al.[72]
SPON1	No	No	No	Yes	No	

Detailed findings such as literature validation sources are provided.

studies employing varied analytical methods and replication in independent cohorts will be essential to clarify the intricate relationships among these factors. More broadly, while MR enhances causal inference, it offers robust evidence rather than absolute proof.

While our bi-directional MR framework offers several advantages, it is important to acknowledge some limitations. One key issue is the reliance on stepwise selection methods, which, although commonly used, can be sensitive to the choice of selection criteria. This sensitivity may result in different sets of valid IVs being chosen, ultimately affecting the estimates and interpretations. Future research could focus on developing more robust selection strategies. Additionally, our analysis is limited to a two-sample MR design using summary statistics from independent GWAS datasets. Expanding the method to account for participant overlap in two-sample settings would enhance its generalizability. Finally, the current implementation of the Bidir-SW approach assumes a linear relationship. Given the complexity of biological systems, extending the model to capture potential non-linear effects would be a feasible direction for future work.

However, interpreting potential bi-directional effects identified through MR—ours included—requires careful consideration of pleiotropy. Although our analytical framework includes techniques to account for pleiotropy, correlated pleiotropy still remains a relatively complex challenge. In such cases, observed associations may reflect a common underlying cause rather than a true bi-directional relationship. Addressing this will require continued development of more advanced statistical methodologies, thorough sensitivity analyses to test the robustness of our findings, and complementary evidence from other research approaches. Recognizing the potential for shared etiologies to mimic bi-directional causality is essential for a careful interpretation of results.

Last, as with many MR approaches, our analysis relies on the "NO Measurement Error" (NOME) assumption,[75] which treats the uncertainty in estimating SNP-exposure associations as negligible. However, this assumption may not hold in practice, particularly when the IVs are weak or the exposure sample size is limited. The measurement error can introduce bias into the causal estimate and lead to underestimation of standard errors. While our current research assumes negligible measurement error, we acknowledge this limitation and recommend future work to explore robust estimators that can directly address this source of bias.

Data and code availability

The CVD proteomics GWAS summary data are publicly available at https://zenodo.org/records/2615265. The GWAS summary statistics for BMI, CRP levels, and sleep duration are downloadable from https://gwas.mrcieu.ac.uk/datasets/ieu-a-835/, https://gwas.mrcieu.ac.uk/datasets/ieu-b-35/, and https://gwas.mrcieu.ac.uk/datasets/ebi-a-GCST003839/. The proposed Bidir-SW method is provided on GitHub: https://github.com/chen-siyi7/Bidir-SW.

Acknowledgments

This research was supported by LONI (Louisiana Optical Network Infrastructure) High-Performance Computing. We also thank the reviewers for many helpful comments.

Declaration of interests

The author declares no competing interests.

Table 14. Proteins identified as significantly causal in the CRP-to-proteins direction by Bidir-SW(AIC) and Bidir-SW(BIC)

Proteins	Bidir-SW (AIC)	Bidir-SW (BIC)
CHI3L1	Yes	No
FS	Yes	Yes
IL-6	Yes	Yes

FS and IL-6 were identified as significant by both selection criteria, while CHI3L1 was detected only by Bidir-SW(AIC).

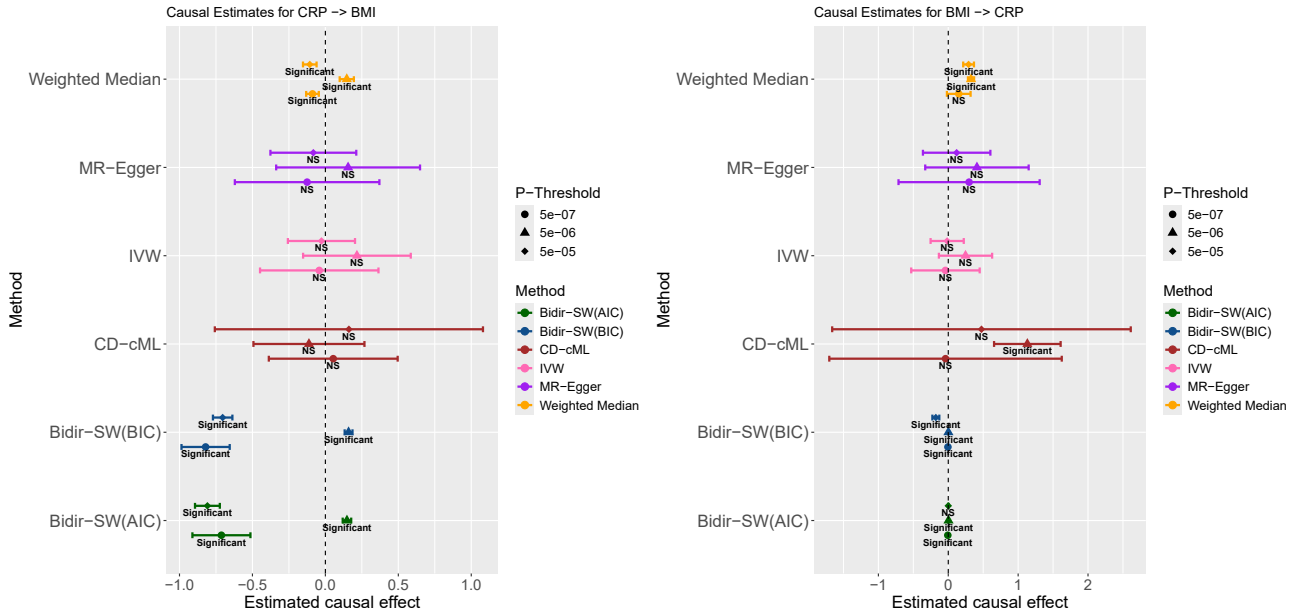

Figure 9. Bi-directional MR analysis between BMI and CRP across multiple methods and *p*-value thresholds

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2025.100449.

Received: October 30, 2024
Accepted: May 1, 2025

References

1. Sanderson, E., Glymour, M.M., Holmes, M.V., Kang, H., Morrison, J., Munafò, M.R., Palmer, T., Schooling, C.M., Wallace, C., Zhao, Q., and Smith, G.D. (2022). Mendelian randomization. Nat. Rev. Methods Primers *2*, 6.
2. Hariton, E., and Locascio, J.J. (2018). Randomised controlled trials—the gold standard for effectiveness research. BJOG *125*, 1716. Author manuscript; available in PMC 2018 Dec 1. NIHMSID: NIHMS966617.
3. Hernán, M.A., and Robins, J.M. (2006). Estimating causal effects from epidemiological data. Journal of Epidemiology & Community Health *60*, 578–586.
4. Roger, J. (2008). Bowden and Darrell A. Turkington. *Instrumental Variables* (Cambridge University Press).
5. Didelez, V., and Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. Stat. Methods Med. Res. *16*, 309–330.
6. Kang, H., Zhang, A., Cai, T.T., and Small, D.S. (2016). Instrumental variables estimation with some invalid instruments

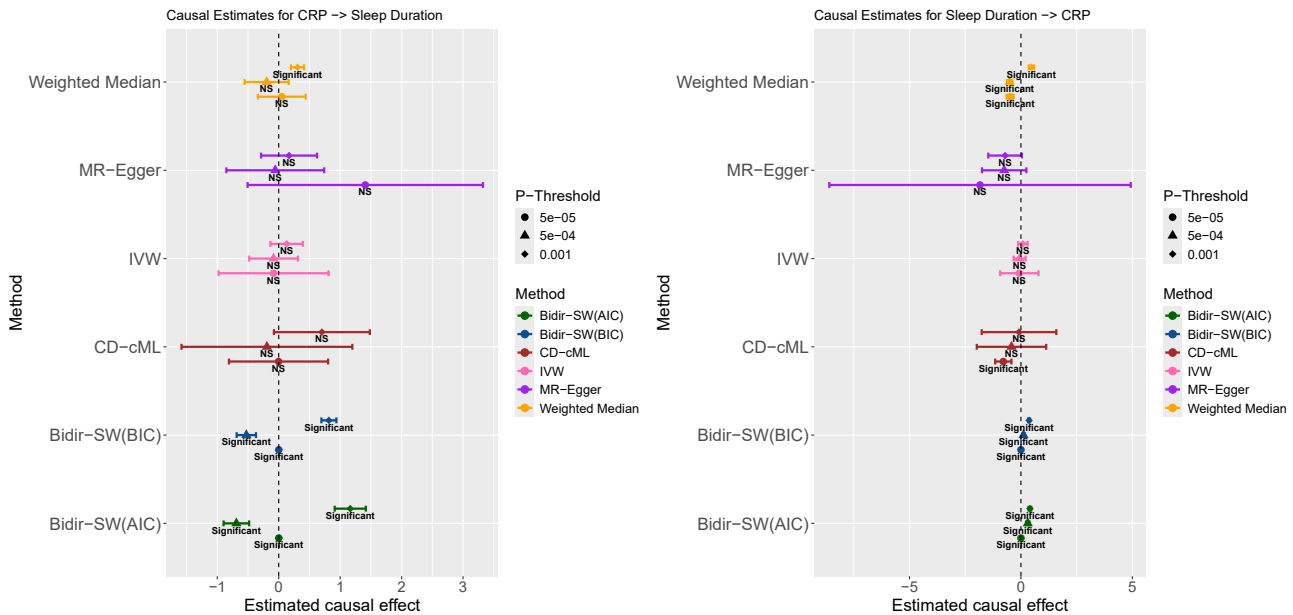

Figure 10. Bi-directional MR analysis between SLP and CRP across multiple methods and *p*-value thresholds

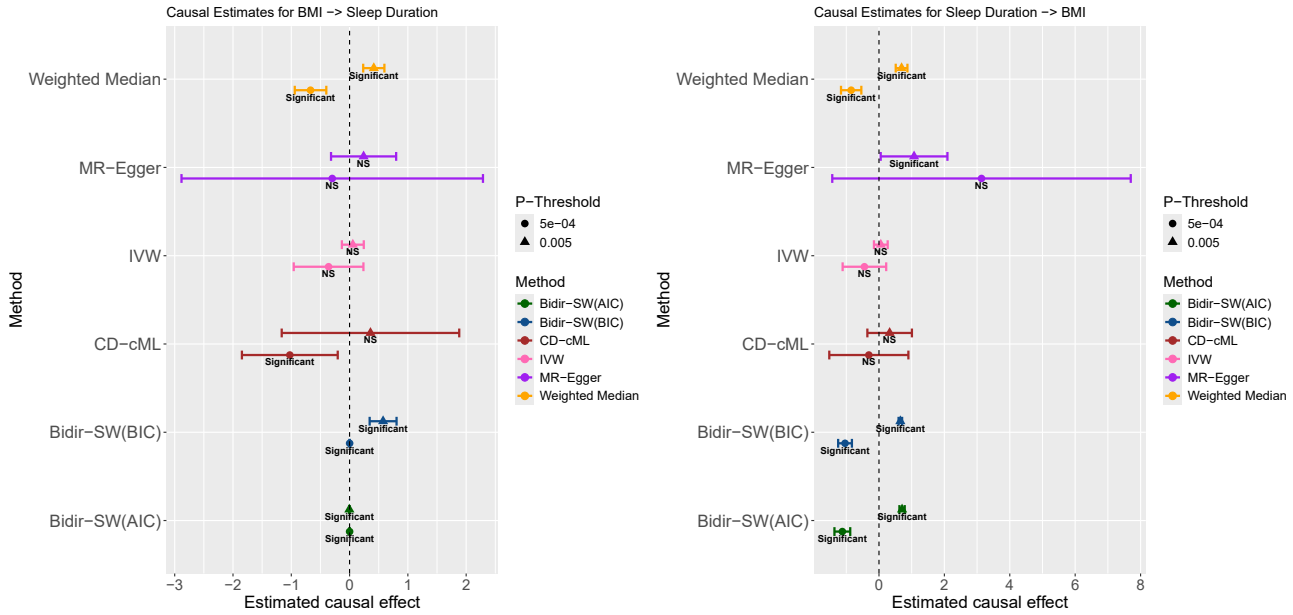

Figure 11. Bi-directional MR analysis between SLP and BMI across multiple methods and *p*-value thresholds

and its application to mendelian randomization. J. Am. Stat. Assoc. *111*, 132–144.

7. Windmeijer, F., Farbmacher, H., Davies, N., and Davey Smith, G. (2019). On the use of the lasso for instrumental variables estimation with some invalid instruments. J. Am. Stat. Assoc. *114*, 1339–1350.
8. Hemani, G., Tilling, K., and Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using gwas summary data. PLoS Genet. *13*, e1007081.
9. Xue, H., and Pan, W. (2020). Inferring causal direction between two traits in the presence of horizontal pleiotropy with gwas summary data. PLoS Genet. *16*, e1009105.
10. Burgess, S., Butterworth, A., and Thompson, S.G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. Genet. Epidemiol. *37*, 658–665.
11. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. Int. J. Epidemiol. *44*, 512–525.
12. Xue, H. (2021). Constrained Likelihood Inference in Instrumental Variable Regression with Invalid Instruments and its Application to GWAS Summary Data. PhD Thesis (university of minnesota).
13. 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 Genet. *17*, e1009922.
14. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., GTEx Consortium, Nicolae, D.L., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098.
15. Knutson, K.A., Deng, Y., and Pan, W. (2020). Implicating causal brain imaging endophenotypes in alzheimer's disease using multivariable iwas and gwas summary data. Neuroimage *223*, 117347.
16. Xue, H., and Pan, W. (2022). Robust inference of bi-directional causal relationships in presence of correlated pleiotropy with gwas summary data. PLoS Genet. *18*, e1010205.
17. Li, S., and Ye, T. (2024). A focusing framework for testing bi-directional causal effects with gwas summary data. J. Roy. Stat. Soc. B *87*, 529–548.
18. Chen, S., Lin, Z., Shen, X., Li, L., and Pan, W. (2023). Inference of causal metabolite networks in the presence of invalid instrumental variables with gwas summary data. Genet. Epidemiol. *47*, 585–599.
19. Schwarz, G.E. (1978). On the bayesian information criterion for model selection. Ann. Stat. *6*, 461–464.
20. Akaike, H. (1974). Information theory and an extension of the maximum likelihood principle. In Proceedings of the 2nd International Symposium on Information Theory, B.N. Petrov and F. Csaki, eds. (Budapest, Hungary: Akademiai Kiado), pp. 267–281.
21. Pacini, D., and Windmeijer, F. (2016). Robust inference for the two-sample 2sls estimator. Econ. Lett. *146*, 50–54.
22. Pattee, J., and Pan, W. (2020). Penalized regression and model selection methods for polygenic scores on summary statistics. PLoS Comput. Biol. *16*, e1008271.
23. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G. A., Abecasis, G.R.; and 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature *526*, 68–74.
24. 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.
25. Folkersen, L., Gustafsson, S., Wang, Q., Hansen, D.H., Hedman, Å.K., Schork, A., Page, K., Zhernakova, D.V., Wu, Y., Peters, J., et al. (2020). Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. Nat. Metab. *2*, 1135–1148.

26. Jones, S.E., Tyrrell, J., Wood, A.R., Beaumont, R.N., Ruth, K.S., Tuke, M.A., Yaghootkar, H., Hu, Y., Teder-Laving, M., Hayward, C., et al. (2016). Genome-wide association analyses in 128,266 individuals identifies new morningness and sleep duration loci. PLoS Genet. *12*, e1006125.
27. Smith, J., Doe, J., and Patel, R. (2024). Impact of inflammatory markers on sleep quality in patients with obstructive sleep apnea. Journal of Sleep Research *33*, 39288581.
28. Zhu, L., Verbeek, K.A., Roussel, M., Shepherd, M., Yang, M., and Smith, G.E. (2022). Obstructive sleep apnea and cardiovascular risk: a mechanistic perspective. Sleep *39*, 1361–1372.
29. Díaz-García, E., García-Sánchez, A., Alfaro, E., López-Fernández, C., Mañas, E., Cano-Pumarega, I., López-Collazo, E., García-Río, F., and Cubillos-Zapata, C. (2023). Psgl-1: a novel immune checkpoint driving t-cell dysfunction in obstructive sleep apnea. Front. Immunol. *14*, 1277551.
30. Theorell-Haglöw, J., Hammar, U., Lind, L., Elmståhl, S., Lindberg, E., and Fall, T. (2021). Sleep duration is associated with protein biomarkers for cardiometabolic health: A large-scale population study. Journal of Sleep Research *30*, e13284.
31. Gozal, D., Khalyfa, A., Qiao, Z., Smith, D.L., Philby, M.F., Koren, D., and Kheirandish-Gozal, L. (2017). Angiopoietin-2 and soluble tie-2 receptor plasma levels in children with obstructive sleep apnea and obesity. Obesity *25*, 1083–1090.
32. Wang, Y., Meagher, R.B., Ambati, S., Ma, P., and Phillips, B.G. (2020). Patients with obstructive sleep apnea have suppressed levels of soluble cytokine receptors involved in neurodegenerative disease, but normal levels with airways therapy. Sleep Breath. *25*, 1641–1653.
33. Franczak, A., Skomro, R., and Sawicki, G. (2020). Matrix metalloproteinases as possible biomarkers of obstructive sleep apnea severity - a systematic review. Sleep Med. Rev. *46*, 9–16.
34. Olejarz, W., Głuszko, A., Cyran, A., Bednarek-Rajewska, K., Proczka, R., Smith, D.F., Ishman, S.L., Migacz, E., and Kukwa, W. (2020). Tlrs and rage are elevated in carotid plaques from patients with moderate-to-severe obstructive sleep apnea syndrome. Sleep Breath. *24*, 741–752.
35. Català, R., Texidó, A., Hernández-Flix, S., Ferré, R., Sangenís, S., Plana, N., Texidó, A., and Masana, L. (2013). Circulating fabp4 and fabp5 levels are differently linked to osa severity and treatment. Sleep *36*, 1833–1840.
36. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206.
37. Anders, L., Carlsson, L., Lind, A.-L., Gordh, T., Bodolea, C., Kamali-Moghaddam, M., and Thulin, M. (2015). The body mass index (bmi) is significantly correlated with levels of cytokines and chemokines in cerebrospinal fluid. Cytokine *76*, 514–518.
38. Ali, H., Zmuda, J.M., Cvejkus, R.K., Kershaw, E.E., Kuipers, A. L., Oczypok, E.A., Wheeler, V., Bunker, C.H., and Miljkovic, I. (2019). Wnt pathway inhibitor dkk1: A potential novel biomarker for adiposity. J. Endocr. Soc. *3*, 488–495.
39. Queipo-Ortuño, M.I., Escoté, X., Ceperuelo-Mallafré, V., Garrido-Sanchez, L., Miranda, M., Clemente-Postigo, M., Pérez-Pérez, R., Peral, B., Cardona, F., Fernández-Real, J.M., et al. (2012). Fabp4 dynamics in obesity: Discrepancies in adipose tissue and liver expression regarding circulating plasma levels. PLoS One *7*, e48605.
40. Perakakis, N., Kokkinos, A., Peradze, N., Tentolouris, N., Ghaly, W., Tsilingiris, D., Alexandrou, A., and Mantzoros, C.S. (2019). Follistatins in glucose regulation of healthy and obese subjects. Metabolism: Clinical and Experimental *93*, 149–159.
41. Matsumoto, S., Kishida, K., Shimomura, I., Maeda, N., Nagaretani, H., Matsuda, M., Nishizawa, H., Kihara, S., Funahashi, T., Matsuzawa, Y., et al. (2002). Increased plasma hb-egf associated with obesity and coronary artery disease. Biochem. Biophys. Res. Commun. *292*, 781–786.
42. Nomata, Y., Kume, N., Sasai, H., Katayama, Y., Nakata, Y., Okura, T., and Tanaka, K. (2009). Weight reduction can decrease circulating soluble lectin-like oxidized low-density lipoprotein receptor-1 levels in overweight middle-aged men. Metabolism *58*, 1209–1214.
43. Qaddoumi, M.G., Alanbaei, M., Hammad, M.M., Al Khairi, I., Cherian, P., Channanath, A., Thanaraj, T.A., Al-Mulla, F., Abu-Farha, M., and Abubaker, J. (2020). Investigating the role of myeloperoxidase and angiopoietin-like protein 6 in obesity and diabetes. Sci. Rep. *10*, 6170.
44. Gutowska, K., Czajkowski, K., and Kuryłowicz, A. (2023). Receptor for the advanced glycation end products (rage) pathway in adipose tissue metabolism. Int. J. Mol. Sci. *24*, 10982.
45. Wang, C.-yao, Zhang, C.-ping, Li, B.-jie, Jiang, S.-su, He, W.-he, Long, S.-yin, and Tian, Y. (2020). Mmp-12 as a potential biomarker to forecast ischemic stroke in obese patients. Med. Hypotheses *136*, 109524.
46. Raman, A., Peiffer, J.J., Hoyne, G.F., Lawler, N.G., Currie, A., and Fairchild, T.J. (2023). Exercise-induced responses in matrix metalloproteinases and osteopontin are not moderated by exercise format in males with overweight or obesity. Eur. J. Appl. Physiol. *123*, 1115–1124.
47. Vila, G., Riedl, M., Maier, C., Struck, J., Morgenthaler, N.G., Handisurya, A., Prager, G., Ludvik, B., Clodi, M., and Luger, A. (2009). Plasma mr-proadm correlates to bmi and decreases in relation to leptin after gastric bypass surgery. Obesity *17*, 1184–1188.
48. Ahangari, F., Sood, A., Ma, B., Takyar, S., Schuyler, M., Qualls, C., Dela Cruz, C.S., Chupp, G.L., Lee, C.G., and Elias, J.A. (2015). Chitinase 3-like-1 regulates both visceral fat accumulation and asthma-like th2 inflammation. Am. J. Respir. Crit. Care Med. *191*, 746–757.
49. Paracchini, V., Pedotti, P., and Taioli, E. (2005). Genetics of leptin and obesity: A huge review. Am. J. Epidemiol. *162*, 101–114.
50. Ligthart, S., Vaez, A., Võsa, U., Stathopoulou, M.G., de Vries, P.S., Prins, B.P., Van der Most, P.J., Tanaka, T., Naderi, E., Rose, L.M., et al. (2018). Genome analyses of >200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. Am. J. Hum. Genet. *103*, 691–706.
51. Ehlenz, K., Koch, B., Preuss, P., Simon, B., Koop, I., and Lang, R.E. (2009). High levels of circulating adrenomedullin in severe illness: correlation with c-reactive protein and evidence against the adrenal medulla as site of origin. Exp. Clin. Endocrinol. Diabetes *105*, 156–162.
52. Hefler-Frischmuth, K., Hefler, L.A., Heinze, G., Paseka, V., Grimm, C., and Tempfer, C.B. (2009). Serum c-reactive protein in the differential diagnosis of ovarian masses. Eur. J. Obstet. Gynecol. Reprod. Biol. X. *147*, 65–68.
53. Lindqvist, D., Hall, S., Surova, Y., Nielsen, H.M., Janelidze, S., Brundin, L., and Hansson, O. (2013). Cerebrospinal fluid inflammatory markers in parkinson's disease – associations

with depression, fatigue, and cognitive impairment. Brain Behav. Immun. *33*, 183–189.
54. Lehrke, M., Millington, S.C., Lefterova, M., Cumaranatunge, R.G., Szapary, P., Wilensky, R., Rader, D.J., Lazar, M.A., and Reilly, M.P. (2007). Cxcl16 is a marker of inflammation, atherosclerosis, and acute coronary syndromes in humans. J. Am. Coll. Cardiol. *49*, 442–449.
55. Hassanzadeh Daloee, S., Nakhaei, N., Hassanzadeh Daloee, M., Mahmoodi, M., Barzegar-Amini, M., and Soltani, T.G. (2021). Evaluation of growth differentiation factor-15 in patients with or without coronary artery disease. J. Cardiovasc. Thorac. Res. *13*, 87–92.
56. Torres, L., Klingberg, E., Nurkkala, M., Carlsten, H., and Forsblad-d'Elia, H. (2019). Hepatocyte growth factor is a potential biomarker for osteoproliferation and osteoporosis in ankylosing spondylitis. Osteoporos. Int. *30*, 441–449.
57. Berger, P., McConnell, J.P., Nunn, M., Kornman, K.S., Sorrell, J., Stephenson, K., and Duff, G.W. (2002). C-reactive protein levels are influenced by common il-1 gene variations. Cytokine *17*, 171–174.
58. Del Giudice, M., and Gangestad, S.W. (2018). Rethinking il-6 and crp: Why they are more than inflammatory biomarkers, and why it matters. Brain Behav. Immun. *70*, 61–75.
59. Shih, H.H., Zhang, S., Cao, W., Hahn, A., Wang, J., Paulsen, J. E., and Harnish, D.C. (2009). Crp is a novel ligand for the oxidized ldl receptor lox-1. Am. J. Physiol. Heart Circ. Physiol. *296*, H1643–H1650.
60. Zakiyanov, O., Kriha, V., Vachek, J., Zima, T., Tesar, V., and Kalousova, M. (2013). Placental growth factor, pregnancy-associated plasma protein-a, soluble receptor for advanced glycation end products, extracellular newly identified receptor for receptor for advanced glycation end products binding protein and high mobility group box 1 levels in patients with acute kidney injury: a cross-sectional study. BMC Nephrol. *14*, 245.
61. Wirestam, L., Pihl, S., Saleh, M., Wetterö, J., and Sjöwall, C. (2021). Plasma c-reactive protein and pentraxin-3 reference intervals during normal pregnancy. Front. Immunol. *12*, 722118.
62. Ekart, R., Homsak, E., Gorenjak, M., Bevc, S., Hojs, N., and Hojs, R. (2015). Soluble st2 is associated with ntprobnp and c-reactive protein in ckd patients on hemodialysis. Nephrol. Dial. Transplant. *30*, iii566–iii567.
63. Singh, U., Devaraj, S., and Jialal, I. (2005). C-reactive protein decreases tissue plasminogen activator activity in human aortic endothelial cells: evidence that c-reactive protein is a procoagulant. Arterioscler. Thromb. Vasc. Biol. *25*, 2216–2221.
64. Fay, W.P. (2010). Linking inflammation and thrombosis: Role of c-reactive protein. N. Engl. J. Med. *363*, 1959–1971.
65. Camargo, J.F., Pallikkuth, S., Moroz, I., Natori, Y., Alcaide, M. L., Rodriguez, A., Guerra, G., Burke, G.W., and Pahwa, S. (2019). Pretransplant levels of c-reactive protein, soluble tnf receptor-1, and cd38+hladr+ cd8 t cells predict risk of allograft rejection in hiv+ kidney transplant recipients. Kidney Int. Rep. *4*, 1705–1716.
66. Secchiero, P., Rimondi, E., di Iasio, M.G., Agnoletto, C., Melloni, E., Volpi, I., and Zauli, G. (2013). C-reactive protein downregulates trail expression in human peripheral monocytes via an egr-1–dependent pathway. Clin. Cancer Res. *19*, 1949–1959.
67. Chen, J.Y., Gu, Z.J., Wu, M.X., Yang, Y., Zhang, J.H., Ou, J.S., Zuo, Z.Y., Wang, J.F., and Chen, Y.X. (2016). C-reactive protein can upregulate vegf expression to promote adsc-induced angiogenesis by activating hif-1α via cd64/pi3k/akt and mapk/erk signaling pathways. Oncotarget *7*, 114–55487.
68. Montecucco, F., Steffens, S., Burger, F., Pelli, G., Monaco, C., and Mach, F. (2008). C-reactive protein (crp) induces chemokine secretion via cd11b/icam-1 interaction in human adherent monocytes. J. Leukoc. Biol. *84*, 1109–1119.
69. Synowski, S.J., Kop, W.J., Christenson, R.H., Gottdiener, J.S., Marshall, J.M., Marshall, J.P., Jerome, S.D., and Gottlieb, S.S. (2012). Abstract 16594: Elevated levels of galectin-3 are associated with higher levels of c-reactive protein, increased disease severity and poorer prognosis in patients with chronic heart failure. Circulation *126*, A16594.
70. Cui, X., Jiao, C., Wang, D., Ye, Z., Ma, J., Tang, N., and Zhang, H. (2021). Elevated levels of il-27 are associated with disease activity in patients with crohn's disease. Mediat. Inflamm. *2021*, 5527627. Article ID 8564213, 2021.
71. Rhee, E.-J., Lee, W.-Y., Kim, S.-Y., Kim, B.-J., Sung, K.-C., Kim, B.-S., Kang, J.-H., Oh, K.-W., Oh, E.-S., Baek, K.-H., et al. (2005). Relationship of serum osteoprotegerin levels with coronary artery disease severity, left ventricular hypertrophy and c-reactive protein. Clin. Sci. *108*, 237–243.
72. Montero, I., Orbe, J., Varo, N., Beloqui, O., Monreal, J.I., Rodríguez, J.A., Díez, J., Libby, P., and Páramo, J.A. (2006). C-reactive protein induces matrix metalloproteinase-1 and -10 in human endothelial cells: Implications for clinical and subclinical atherosclerosis. J. Am. Coll. Cardiol. *47*, 1369–1378.
73. Visser, M., Bouter, L.M., McQuillan, G.M., Wener, M.H., and Harris, T.B. (1999). Elevated c-reactive protein levels in overweight and obese adults. JAMA *282*, 2131–2135.
74. Aronson, D., Bartha, P., Zinder, O., Kerner, A., Markiewicz, W., Avizohar, O., Brook, G.J., and Levy, Y. (2004). Obesity is the major determinant of elevated c-reactive protein in subjects with the metabolic syndrome. Int. J. Obes. *28*, 674–679.
75. Bowden, J., Greco M, F.D., Minelli, C., Smith, G.D., Sheehan, N.A., and Thompson, J.R. (2016). Assessing the suitability of summary data for two-sample mendelian randomization analyses using mr-egger regression: the role of the i2statistic. Int. J. Epidemiol. *45*, 1961–1974.