

## RESEARCH ARTICLE

## Inferring causal direction between two traits in the presence of horizontal pleiotropy with GWAS summary data

Haoran Xue<sup>1</sup>, Wei Pan<sup>2\*</sup>

**1** School of Statistics, University of Minnesota, Minneapolis, Minnesota, United States of America, **2** Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota, United States of America

\* [panxx014@umn.edu](mailto:panxx014@umn.edu)

## OPEN ACCESS

**Citation:** Xue H, Pan W (2020) Inferring causal direction between two traits in the presence of horizontal pleiotropy with GWAS summary data. *PLoS Genet* 16(11): e1009105. <https://doi.org/10.1371/journal.pgen.1009105>

**Editor:** Michael P. Epstein, Emory University, UNITED STATES

**Received:** May 18, 2020

**Accepted:** September 8, 2020

**Published:** November 2, 2020

**Copyright:** © 2020 Xue, Pan. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** GWAS summary data for CAD and LDL/HDL are publicly available at <http://www.cardiogramplusc4d.org/data-downloads/> and <http://csg.sph.umich.edu/willer/public/lipids2017>, respectively. R code implementing our proposed methods is available at [https://github.com/xue-hr/Causal\\_Direction](https://github.com/xue-hr/Causal_Direction).

**Funding:** HX was supported by NIH grant R01GM126002 and Beverly and Richard Fink Graduate Fellowship at the University of Minnesota. WP was supported by NIH grants R01AG065636, RF1AG067924, R21AG057038, R01HL116720,

## Abstract

Orienting the causal relationship between pairs of traits is a fundamental task in scientific research with significant implications in practice, such as in prioritizing molecular targets and modifiable risk factors for developing therapeutic and interventional strategies for complex diseases. A recent method, called Steiger's method, using a single SNP as an instrument variable (IV) in the framework of Mendelian randomization (MR), has since been widely applied. We report the following new contributions. First, we propose a single SNP-based alternative, overcoming a severe limitation of Steiger's method in simply assuming, instead of inferring, the existence of a causal relationship. We also clarify a condition necessary for the validity of the methods in the presence of hidden confounding. Second, to improve statistical power, we propose combining the results from multiple, and possibly correlated, SNPs as multiple instruments. Third, we develop three goodness-of-fit tests to check modeling assumptions, including those required for valid IVs. Fourth, by relaxing one of the three IV assumptions in MR, we propose several methods, including an Egger regression-like approach and its multivariable version (analogous to multivariable MR), to account for horizontal pleiotropy of the SNPs/IVs, which is often unavoidable in practice. All our methods can simultaneously infer both the existence and (if so) the direction of a causal relationship, largely expanding their applicability over that of Steiger's method. Although we focus on uni-directional causal relationships, we also briefly discuss an extension to bi-directional relationships. Through extensive simulations and an application to infer the causal directions between low density lipoprotein (LDL) cholesterol, or high density lipoprotein (HDL) cholesterol, and coronary artery disease (CAD), we demonstrate the superior performance and advantage of our proposed methods over Steiger's method and bi-directional MR. In particular, after accounting for horizontal pleiotropy, our method confirmed the well known causal direction from LDL to CAD, while other methods, including bi-directional MR, might fail.

R01GM113250 and R01GM126002. This work was supported by the Minnesota Supercomputing Institute at the University of Minnesota. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

In spite of its importance, due to technical challenges, orienting causal relationships between pairs of traits has been largely under-studied. Mendelian randomization (MR) Steiger's method has become increasingly used in the last two years. Here we point out several limitations with MR Steiger's method and propose alternative approaches. First, MR Steiger's method is based on using only one single SNP as the instrument variable (IV), for which we propose a correlation ratio-based method, called Causal Direction-Ratio, or simply CD-Ratio. An advantage of CD-Ratio is its inference of both the existence and (if so) the direction of a causal relationship, in contrast to MR Steiger's prior assumption of the existence and its poor performance if the assumption is violated. Furthermore, CD-Ratio can be extended to combine the results from multiple, possibly correlated, SNPs with improved statistical power. Second, we propose two methods, called CD-Egger and CD-GLS, for multiple and possibly correlated SNPs while allowing horizontal pleiotropy. Third, we propose three goodness-of-fit tests to check modeling assumptions for the three proposed methods. Finally, we introduce multivariable CD-Egger, analogous to multivariable MR, as a more robust approach, and an extension of CD-Ratio to cases with possibly bi-directional causal relationships. Our numerical studies demonstrated superior performance of our proposed methods over MR Steiger and bi-directional MR. Our proposed methods, along with freely available software, are expected to be useful in practice for causal inference.

## Introduction

Mendelian Randomization (MR) has become increasingly popular and widely used to infer causal relationships between pairs of traits, which can be molecular (such as gene expression or DNA methylation), behavioral/educational or clinical [1–3]. MR (or two-sample MR) utilizes genetic variants, almost exclusively single nucleotide polymorphisms (SNPs), as instruments or instrumental variables (IVs) in two independent GWAS (summary) datasets for the two traits respectively. Although the gold standard for causal inference is randomized clinical trials (RCTs), RCTs are often too costly, unethical and infeasible, in which case MR offers a cost-effective approach. The recent years' popularity of MR coincides with the ever-increasing availability of large-scale GWAS data and resources/tools like MR-Base [4]. As reviewed in [5], the potential of MR has been fully demonstrated by several notable examples in the studies of cardiovascular and metabolic diseases, such as MR analyses showing of C-reactive protein's and moderate alcohol consumption's unlikely role in reducing coronary heart disease risk, and replicating RCT findings of several drug targets (e.g. Niemann-Pick C1-like protein). In most MR applications, one often assumes that the causal direction between two traits is known. Otherwise, bi-directional MR can be applied [6, 7]: one first applies MR to infer whether there is a causal relationship from one trait, say  $X$ , to another trait, say  $Y$ ; then one reverses the roles of the two traits to infer whether there is a causal relationship from  $Y$  to  $X$ . If one of the two causal directions does hold, one would hope to detect it with bi-directional MR. Bi-directional MR can be also more intuitively applied graphically [8]: the estimated effect sizes on the two traits from the significant SNPs can be compared in two scatter plots, one from the SNPs significant for each trait; it is expected that, if the causal direction is  $X \rightarrow Y$ , then there should be a strong relationship between the two sets of the effect sizes on the two traits of the SNPs significant for  $X$ , but not necessarily so for the SNPs significant for  $Y$ . However, whether bi-directional MR works depends on one critical, most often unknown, assumption: *two sets of valid SNPs/IVs*

used in the two directions. That is, if an SNP primarily influences  $X$ , but influences  $Y$  only through  $X$ , then it should be used as an IV to infer  $X \rightarrow Y$  in the first step; at the same time, it *cannot be used* as an IV to infer  $Y \rightarrow X$ . In practice, due to limited knowledge, it is almost impossible to always choose valid IVs. Instead, often one just uses all significant (and independent) SNPs for  $X$  to infer  $X \rightarrow Y$ , and vice versa for  $Y \rightarrow X$ . As pointed out by [9], if an SNP primarily influences  $X$  (and influences  $Y$  completely through  $X$ ), it will be associated with *both*  $X$  and  $Y$ . Hence, if the sample sizes are large enough, its associations with both  $X$  and  $Y$  will be detected, and it will be used as an IV in *both* steps or directions in bi-directional MR, leading to the incorrect conclusion that a causal relationship exists in *both* directions. This problem with bi-directional MR will be confirmed in our real data examples later. Relatedly, the result of bi-directional MR (or its graphical implementation) is confounded by possibly varying statistical power for the two traits (e.g., due to different sample sizes and different types of the traits such as one quantitative and the other binary). In the above example that an SNP primarily influences trait  $X$  but not  $Y$ , if the statistical power for detecting the association between the SNP and  $Y$  is large while that for the other pair is small, it is likely that the SNP will be significantly associated with  $Y$  but not with  $X$ , thus incorrectly used as an IV for  $Y$  but not for  $X$ , likely leading to the false conclusion of a causal effect of  $Y$  on  $X$ .

Recently a simple and effective approach, called MR Steiger's method, was proposed [9]. It was discussed in the context of comparison with the main competitor of MR, mediation analysis, which requires one-sample data (i.e. both SNPs and the two traits are observed in the same sample). The proposed method is based on testing the difference of two (possibly correlated) Pearson's correlation coefficients using Steiger's test, and thus named. It was shown that MR Steiger's method is more robust to measurement errors and hidden confounding than the causal inference test (CIT) [10] and likely other mediation analysis approaches. However, MR Steiger's method has been most widely used in 2-sample MR applications. In [9], since MR Steiger was discussed in the context of mediation analysis, its derivation was mostly done in the absence of hidden confounding, though in an appendix the issue was briefly discussed under confounding; in contrast, an appealing advantage of MR is its valid causal inference while allowing the presence of hidden confounding, which is largely expected in practice. Therefore, first, we will formulate the problem in the presence of confounding; along the way we will point out for the first time that a key condition is needed for valid inference. Although the condition is expected to hold often in practice, it would be useful to have a method to check the condition; we propose such a method. Second, since MR Steiger's method is based on using only a single SNP as IV, to improve both its statistical power and applicability, we extend it to allow multiple, and possibly correlated, SNPs. In particular, there is increasing interest in applying transcriptome-wide association studies (TWAS, or PrediXcan) to identify causal genes or other molecular/imaging endophenotypes by integrating GWAS with eQTL/xQTL data [11–20]. In these applications, multiple correlated cis-SNPs near a gene are used as IVs to impute or predict the gene's expression level (or another endophenotype) to infer whether the gene's expression has a causal effect on a trait, say coronary artery disease (CAD); it is assumed that the causal relationship (if existing) is from the gene to the trait. Albeit reasonable, there is possibility of reverse causation: a gene's expression change is not the cause, but the consequence, of having CAD. Hence, it would be useful to apply a method to confirm whether the causal direction is indeed from the gene to the trait (or vice versa). In such a case, due to often a small sample size of an eQTL study, it would be low powered to apply MR Steiger's method with a single SNP; instead, it would be more powerful and thus more desirable to apply a method with all the cis-SNPs used to predict the gene's expression level. Our proposed new methods are more suitable for such an application. The same argument was made for two traits based on large-scale GWAS when single SNP summary statistics-based MR (SMR) [13]

was generalized to GSMR with multiple (possibly correlated) SNPs [21]. Third, we develop corresponding goodness-of-fit tests, similar to Cochran's  $Q$  statistic, to check the IV and other modeling assumptions for our proposed methods. Fourth, most importantly, with multiple SNPs/IVs, we can relax one of the three key assumptions in MR for valid causal inference: allowing the presence of horizontal pleiotropy, which is expected to be commonplace as shown by previous studies [22, 23]. In particular, the likely violation of the IV assumptions for TWAS and its severe consequences on invalid causal conclusions have been increasingly recognized and discussed in the literature [14, 24–26]. We propose a valid approach similar to Egger regression, which has been widely used for exactly the same purpose (i.e. in the presence of horizontal pleiotropy) in MR [27]. As a side product of allowing horizontal pleiotropy, each of our methods can be applied to infer whether there is any causal relationship between two traits *and* if so, what is the causal direction; in contrast, MR Steiger's method can only infer causal direction *after* assuming the existence of one of the two causal directions. Finally, we introduce a multivariable approach, analogous to multivariable MR (MV-MR) [28, 29], as another robust approach; although we focus on uni-directional causal relationships, we also briefly discuss an extension to cases with possibly bi-directional causal relationships.

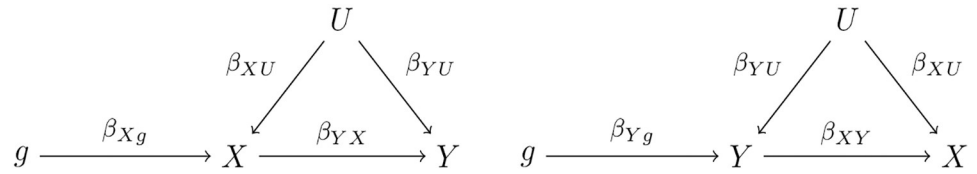
To demonstrate the wide-ranging utility and superior performance of our proposed methods, we applied both the new and existing methods to infer the causal directions between low density lipoprotein (LDL) cholesterol, or high density lipoprotein (HDL) cholesterol, and CAD. Multiple previous MR analyses have clearly established the causal role of LDL cholesterol on CAD [30–33], confirmed by meta-analyses of RCTs with statins and other cholesterol-lowering interventions [34–36]. Hence, the established causal direction of  $LDL \rightarrow CAD$  can serve as a positive control to test various methods. Indeed, after accounting for horizontal pleiotropy, our methods reached the correct conclusion. In contrast, all other methods, including our method without accounting for horizontal pleiotropy and bi-directional MR (with various MR methods either ignoring or robust to/accounting for horizontal pleiotropy), failed to reach the correct conclusion. On the other hand, contrary to observational studies [37], HDL cholesterol does not appear to have a causal role on CAD, as evidenced by several MR analyses and recent RCTs [31–33, 38, 39]. Hence, it can serve as a negative control. Again our proposed and most other methods, as expected, could not establish any causal direction between HDL cholesterol and CAD. We also conducted extensive simulation studies to support the wide applicability and validity of our proposed methods.

## Methods

We will first discuss the ideal case when all chosen SNPs are valid IVs, which means each SNP satisfies the three key IV assumptions: (i) each SNP is associated with the exposure; (ii) conditional on the exposure, each SNP is not associated with the outcome; i.e., there is no direct effect of each SNP on the outcome; (iii) each SNP is not associated with hidden confounders. Then we will consider the case when assumption (ii) is violated when one or more SNPs have direct effects, i.e. horizontal pleiotropic effects, on the outcome. We propose an Egger regression-like approach and an alternative based on the Generalized Least Squares (GLS) estimation method, both performing well and similarly. Finally, we introduce a multivariable extension to the Egger regression method, analogous to (MV-MR), as a more robust approach, and an extension to cases with possibly bi-directional causal relationships.

### Ideal case: With valid IVs (and hidden confounding)

**True model.** If two traits  $X$  and  $Y$  are causally related, we'd like to determine the true causal direction, which can be either  $X \rightarrow Y$  or  $Y \rightarrow X$ . In the ideal case where all the SNPs



**Fig 1. Two possible true models: Left one is  $X \rightarrow Y$ , right one is  $Y \rightarrow X$ .**

<https://doi.org/10.1371/journal.pgen.1009105.g001>

used are valid IVs, two possible true models are shown in Fig 1. We use  $g$  to represent one SNP (but later more SNPs) to be used as IV, and  $U$  is a hidden confounder (or an ensemble of all confounders). Each directed edge represents a direct and causal effect with the effect size above the line. As usual, we also assume that all the relationships in the models are linear.

For  $X \rightarrow Y$ , the data-generative/causal models are:

$$X = \beta_{X0} + \beta_{Xg} \cdot g + \beta_{XU} \cdot U + \epsilon_X, \tag{1}$$

$$Y = \beta_{Y0} + \beta_{YX} \cdot X + \beta_{YU} \cdot U + \epsilon_Y, \tag{2}$$

where  $\epsilon_X$  and  $\epsilon_Y$  are independent error terms.

For  $Y \rightarrow X$ , we have the corresponding true models.

**Steiger’s method based on a single SNP.** Based on true models (1) and (2), we have the (true) correlation between  $X$  and  $g$ , denoted by  $\rho_{Xg}$ , and correlation between  $Y$  and  $g$ , denoted by  $\rho_{Yg}$ , as

$$\rho_{Xg} = \frac{\beta_{Xg} \cdot \text{var}(g)}{\sqrt{\text{var}(g) \cdot \text{var}(X)}}, \quad \rho_{Yg} = \frac{\beta_{Xg} \cdot \beta_{YX} \cdot \text{var}(g)}{\sqrt{\text{var}(g) \cdot \text{var}(Y)}}. \tag{3}$$

From Eq (3) we obtain the correlation ratio:

$$\frac{\rho_{Yg}}{\rho_{Xg}} = \frac{\beta_{YX} \cdot \sqrt{\text{var}(X)}}{\sqrt{\text{var}(Y)}} := K_{YX}. \tag{4}$$

We note that the correlation ratio  $K_{YX}$  is a constant, independent of  $g$ . It is easy to see that, if there is no confounder  $U$  in the true models, and without loss of generality assuming  $\text{var}(\epsilon_Y) > 0$ , we have

$$\text{var}(Y) > \beta_{YX}^2 \cdot \text{var}(X), \tag{5}$$

implying  $|K_{YX}| < 1$ . On the other hand, if the true model is  $Y \rightarrow X$ , we can similarly define and derive that  $|K_{XY}| := |\rho_{Xg} / \rho_{Yg}| < 1$ .

It is noted that the two conditions,  $|K_{YX}| < 1$  and  $|\rho_{Yg}| < |\rho_{Xg}|$ , are equivalent. While we will use the first one, the second one was used in MR Steiger’s method [9]. In spite of the trivial equivalence between the two, a key insight is that, because  $K_{YX}$  is independent of  $g$ , it can be used to combine over multiple SNPs/IVs; furthermore, the first one can be used to infer whether there is indeed a causal relationship between the two traits by checking whether  $K_{YX} = 0$ , equivalent to  $\beta_{YX} = 0$ , in contrast to simply assuming the existence of a causal relationship in Steiger’s method. These two points here are to be elaborated on later.

From two independent GWAS summary datasets with sample sizes  $n_X$  and  $n_Y$  for traits  $X$  and  $Y$  respectively, for SNP  $g$ , we have its marginal effect estimate on  $X$  and the variance,  $\hat{\beta}_{M,Xg}$  and  $\text{var}(\hat{\beta}_{M,Xg})$ , and the corresponding ones for  $Y$  as  $\hat{\beta}_{M,Yg}$  and  $\text{var}(\hat{\beta}_{M,Yg})$ . Then we can

calculate its sample correlations with  $X$  and  $Y$ , denoted by  $r_{Xg}$  and  $r_{Yg}$ , from the summary data:

$$r_{Xg} = \frac{\hat{\beta}_{M,Xg}}{\sqrt{\hat{\beta}_{M,Xg}^2 + (n_X - 2) \cdot \text{var}(\hat{\beta}_{M,Xg})}}, \quad r_{Yg} = \frac{\hat{\beta}_{M,Yg}}{\sqrt{\hat{\beta}_{M,Yg}^2 + (n_Y - 2) \cdot \text{var}(\hat{\beta}_{M,Yg})}}. \quad (6)$$

MR Steiger’s method [9] first applies Fisher’s Z-transformation to  $|r_{Xg}|$ :

$$Z_{Xg} = F(|r_{Xg}|) = \frac{1}{2} \ln \left( \frac{1 + |r_{Xg}|}{1 - |r_{Xg}|} \right) \sim N \left( \frac{1}{2} \ln \left( \frac{1 + |\rho_{Xg}|}{1 - |\rho_{Xg}|} \right), \frac{1}{n_X - 3} \right),$$

and calculate  $Z_{Yg} = F(|r_{Yg}|)$ . Then it infers the causal direction by comparing  $Z_{Xg}$  and  $Z_{Yg}$  to obtain  $Z_g = Z_{Xg} - Z_{Yg}$  and its variance  $\text{var}(Z_g) = 1/(n_X - 3) + 1/(n_Y - 3)$ . The test statistic  $Z_g / \sqrt{\text{var}(Z_g)}$  is compared to a standard normal distribution  $N(0, 1)$  to yield a p-value. If the p-value is larger than a given significance level, MR Steiger’s method would not make any conclusion; otherwise, depending on if  $Z_g > 0$  or not, it would conclude that the causal direction is from  $X$  to  $Y$ , or from  $Y$  to  $X$ .

It is noted that, to test the suitable hypothesis on the magnitudes of the two correlations, MR Steiger’s method takes the Fisher’s transformation on the absolute values of the two correlations, i.e.,  $|r_{Xg}|$  and  $|r_{Yg}|$ . For finite samples, using the non-negative  $|r_{Xg}|$  and  $|r_{Yg}|$  may lead to poor approximations by their normal distributions, though it is asymptotically valid (under usual regularity conditions, including each correlation is not 0 or  $\pm 1$ ). [S10 Text](#) gives an illustration.

It is clear that the key condition for Steiger’s method as proposed by [9] is inequality (5). It is expected to hold with the following intuitive explanation: based on true model (2),  $\text{var}(Y)$  represents the total variation of  $Y$  while  $\beta_{YX}^2 \cdot \text{var}(X)$  is only its partial variation mediated through  $X$ . However, in the presence of confounder  $U$ , as shown in [Eq \(2\)](#), we have

$$\text{var}(Y) = \beta_{YX}^2 \text{var}(X) + \text{var}(\epsilon_Y) + \beta_{YU}^2 \text{var}(U) + 2\beta_{YX}\beta_{YU} \text{cov}(X, U), \quad (7)$$

from which we cannot conclude that (5) holds, simply because the sum of the last three terms in  $\text{var}(Y)$  can be negative. A sufficient condition for (5) is that  $\beta_{YX}\beta_{YU} \text{cov}(X, U)$  is (i) positive, or (ii) negative but small in magnitude. Condition (i) means that the directions of the correlations between  $X$  and  $U$  and between  $Y$  and  $U$  are consistent with the sign of  $\beta_{YX}$ : for example, if  $\beta_{YX} > 0$ , then the effects of  $U$  on  $X$  and  $Y$  are in the same direction; on the other hand, if  $\beta_{YX} < 0$ , then the effects of  $U$  on  $X$  and  $Y$  are in different directions. A sufficient condition for (ii) is the small confounding and small causal effect sizes. For complex traits, we believe that these assumptions (i) and (ii) are often reasonable, albeit untestable directly due to that the confounder  $U$  is unobserved. Furthermore, [9] showed numerically that, if the proportion of the variance explained between the two traits,  $R_{YX}^2$ , is less than 0.2, then the condition almost always holds. It is noted that for complex human traits  $X$  and  $Y$  that are influenced by many genetic and environmental factors,  $R_{YX}^2 < 0.2$  is expected to hold often.

Nevertheless, as an assumption, it would be desirable to check condition (5) in practice. Next we propose such a method. First, we can obtain an estimate  $\hat{\beta}_{YX}$  of the causal effect  $\beta_{YX}$ , e.g. by MR Egger regression. Second, if we have individual level data, then we can estimate  $\text{var}(X)$  and  $\text{var}(Y)$  directly, thus verify whether the condition holds. It is noted that  $\text{var}(X)$  and  $\text{var}(Y)$  are just the marginal variances of the two traits, which can be consistently estimated by the corresponding sample variances. On the other hand, if we only have GWAS summary data for both traits, we propose the following procedure to estimate  $\text{var}(X)$  and  $\text{var}(Y)$ . For each SNP  $g$ , we have its marginal association estimate  $\hat{\beta}_{M,Xg}$ ,  $\text{var}(\hat{\beta}_{M,Xg})$  and  $\text{var}(g) = 2\text{MAF}(1-\text{MAF})$  in the

marginal model  $X = \beta_{M,0Xg} + \beta_{M,Xg} \cdot g + \epsilon'_X$ , where MAF is the minor allele frequency of the SNP. Then it is easy to verify that

$$var(X) = \hat{\beta}_{M,Xg}^2 var(g) + var(\epsilon'_X) = \hat{\beta}_{M,Xg}^2 var(g) + (n_X - 2) var(\hat{\beta}_{M,Xg}) var(g),$$

where  $n_X$  is the sample size. We can similarly estimate  $var(Y)$ . Since such variance estimates may depend on the SNP  $g$  being used, we can obtain many such estimates based on each of many SNPs, then check the distribution of  $var(Y)/var(X)$  and compare to  $\hat{\beta}_{YX}^2$  to verify condition (5).

In practice, as to be shown in our real data examples, in order to apply MR Steiger’s method and its extensions to be proposed next, we first assume that, if the causal direction is from  $X$  to  $Y$ , condition (5) holds; we apply our above proposed procedure to check it.

**New method with multiple and possibly correlated SNPs.** Steiger’s method is based on using only one SNP as IV. There is a good reason to extend it to multiple and possibly correlated SNPs, e.g. from a genomic locus: to improve statistical efficiency with multiple SNPs. In particular, in some applications such as TWAS [11, 12], it is necessary to use multiple correlated cis SNPs to more effectively predict/impute one gene’s expression level; if we are interested in inferring whether the causal direction is from the gene to a GWAS trait, we have to use multiple correlated SNPs to improve power.

Now we describe the new method for multiple, possibly correlated, SNPs. Due to linkage disequilibrium (LD), it is likely that some SNPs at the same locus are correlated and are significantly associated with both  $X$  and  $Y$ . Suppose we have  $m$  such SNPs denoted in  $g$ ; for each SNP  $i$ , from Eq (6), with summary statistics we can get its sample correlations with  $X$  and  $Y$  respectively as  $r_{Xg_i}$  and  $r_{Yg_i}$ . Denote  $\mathbf{r}_{Xg} = (r_{Xg_1}, \dots, r_{Xg_m})^T$ ,  $\mathbf{r}_{Yg} = (r_{Yg_1}, \dots, r_{Yg_m})^T$ ,  $\boldsymbol{\rho}_{Xg} = (\rho_{Xg_1}, \dots, \rho_{Xg_m})^T$  and  $\boldsymbol{\rho}_{Yg} = (\rho_{Yg_1}, \dots, \rho_{Yg_m})^T$ . As discussed in [40], we obtain the asymptotic distributions of  $\mathbf{r}_{Xg}$  and  $\mathbf{r}_{Yg}$  as

$$\sqrt{n_X} \cdot (\mathbf{r}_{Xg} - \boldsymbol{\rho}_{Xg}) \rightarrow N(0, \mathbf{V}_{Xg}), \quad \sqrt{n_Y} \cdot (\mathbf{r}_{Yg} - \boldsymbol{\rho}_{Yg}) \rightarrow N(0, \mathbf{V}_{Yg}), \tag{8}$$

where the covariance matrices  $\mathbf{V}_{Xg}$  and  $\mathbf{V}_{Yg}$  depend on the correlation matrix of the SNPs in a complicated way, as shown in S1 Text. With GWAS summary data, we use a reference panel, such as one from the 1000 Genomes Project [41], to estimate the correlation matrix of the SNPs. Given the asymptotic distributions of  $\mathbf{r}_{Xg}$  and  $\mathbf{r}_{Yg}$ , and the two independent GWAS datasets,  $\mathbf{r}_{Xg}$  and  $\mathbf{r}_{Yg}$  are also independent, leading to a block-diagonal covariance matrix for their joint distribution. By the delta method, we can obtain:

$$\begin{pmatrix} \mathbf{r}_{Yg} \\ \mathbf{r}_{Xg} \end{pmatrix} = \begin{pmatrix} \frac{r_{Yg_1}}{r_{Xg_1}} \\ \vdots \\ \frac{r_{Yg_m}}{r_{Xg_m}} \end{pmatrix} \sim N \left( \begin{pmatrix} \frac{\rho_{Yg_1}}{\rho_{Xg_1}} \\ \vdots \\ \frac{\rho_{Yg_m}}{\rho_{Xg_m}} \end{pmatrix}, \mathbf{J}_g^T \cdot \begin{pmatrix} \frac{\mathbf{V}_{Yg}}{n_Y} & \mathbf{0}_{m \times m} \\ \mathbf{0}_{m \times m} & \frac{\mathbf{V}_{Xg}}{n_X} \end{pmatrix} \cdot \mathbf{J}_g \right) := N(\mathbf{1} \cdot K_{YX}, \mathbf{V}_{YXg}), \tag{9}$$

where “:=” represents “by definition”,  $\mathbf{1} = (1, \dots, 1)^T$  and

$$\mathbf{J}_g^T = \begin{pmatrix} 1 & \dots & 0 & -\frac{r_{Yg_1}}{r_{Xg_1}^2} & \dots & 0 \\ r_{Xg_1} & & & & & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{r_{Xg_m}} & 0 & \dots & -\frac{r_{Yg_m}}{r_{Xg_m}^2} \end{pmatrix}. \tag{10}$$

As we have already shown,  $\rho_{Yg_1}/\rho_{Xg_1} = \dots = \rho_{Yg_m}/\rho_{Xg_m} = K_{YX}$  is a constant, independent of the SNPs. Hence, the generalized least squares (GLS) method is applied to estimate  $K_{YX}$  as

$$\hat{K}_{YX} = \frac{\mathbf{1}^T \cdot \mathbf{V}_{YXg}^{-1} \cdot \mathbf{r}_{Yg}}{\mathbf{1}^T \cdot \mathbf{V}_{YXg}^{-1} \cdot \mathbf{1}}, \quad \text{var}(\hat{K}_{YX}) = \frac{1}{\mathbf{1}^T \cdot \mathbf{V}_{YXg}^{-1} \cdot \mathbf{1}}. \tag{11}$$

We call this method **Causal Direction-Ratio**, or **CD-Ratio** for short. It is noted that **CD-Ratio** can be applied to special situations with only one SNP, or with multiple independent SNPs. Note that for two independent SNPs, their sample correlations with the same trait are in general non-independent.

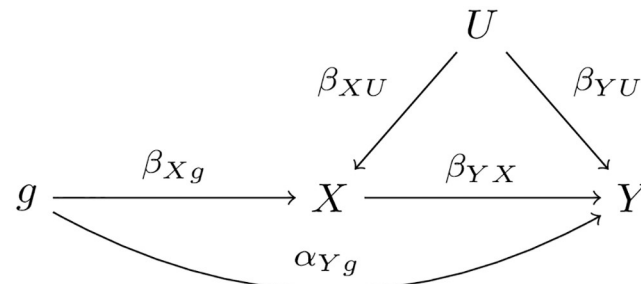
With each individual estimate, or the combined estimate  $\hat{K}_{YX}$  and its  $\text{var}(\hat{K}_{YX})$ , we can construct a confidence interval for  $K_{YX}$ . Similarly, from the other direction we can estimate  $K_{XY}$  as  $\hat{K}_{XY}$  with  $\text{var}(\hat{K}_{XY})$ . Assuming condition (5) is satisfied, we infer the causal direction between  $X$  and  $Y$  by comparing whether the confidence interval of  $K_{XY}$  (or  $K_{YX}$ ) is completely within or outside interval  $[-1, 1]$ , as to be discussed later.

### Extension to case with invalid IVs with direct effects

Now consider one or more SNPs/IVs with direct effects on the outcome as shown in Fig 2. The true causal models are

$$X = \beta_{X0} + \beta_{Xg} \cdot g + \beta_{XU} \cdot U + \epsilon_X, \tag{12}$$

$$Y = \beta_{Y0} + \beta_{YX} \cdot X + \alpha_{Yg} \cdot g + \beta_{YU} \cdot U + \epsilon_Y, \tag{13}$$



**Fig 2. True model for causal direction from  $X$  to  $Y$  in the presence of pleiotropic/direct effect (i.e.  $\alpha_{Yg}$ ).**

<https://doi.org/10.1371/journal.pgen.1009105.g002>



implying

$$Y = (\beta_{Y0} + \beta_{YX} \cdot \beta_{X0}) + (\alpha_{Yg} + \beta_{YX} \cdot \beta_{Xg}) \cdot g + (\beta_{YU} + \beta_{YX} \cdot \beta_{XU}) \cdot U + (\epsilon_Y + \beta_{YX} \cdot \epsilon_X).$$

Accordingly, we calculate the (Pearson's) correlations between the SNP and X and Y,

$$\rho_{Xg} = \beta_{Xg} \cdot \frac{\sqrt{\text{var}(g)}}{\sqrt{\text{var}(X)}}, \tag{14}$$

$$\rho_{Yg} = (\alpha_{Yg} + \beta_{YX} \cdot \beta_{Xg}) \cdot \frac{\sqrt{\text{var}(g)}}{\sqrt{\text{var}(Y)}} = \alpha_{Yg} \cdot \frac{\sqrt{\text{var}(g)}}{\sqrt{\text{var}(Y)}} + K_{YX} \cdot \rho_{Xg} := b_{Yg} + K_{YX} \cdot \rho_{Xg}. \tag{15}$$

Again by definition  $K_{YX} := \beta_{YX} \sqrt{\text{var}(X)} / \sqrt{\text{var}(Y)}$ ; assuming condition (5) is satisfied, we have  $|K_{YX}| < 1$ . Furthermore,  $K_{YX} = 0$  is the sufficient and necessary condition for  $\beta_{YX} = 0$ , i.e. no causal relationship from X to Y. Different from that in the Ideal Model, as pleiotropic effects are allowed now, even there is no causal relationship between X and Y, there could still be significant SNPs  $g$  for both traits due to pleiotropy. By testing whether  $K_{YX} = 0$  or not, we can determine the existence of a causal relationship.

It is noted that  $b_{Yg} \propto \alpha_{Yg} \neq 0$ . We next develop an estimation method by treating  $b_{Yg}$ 's as random effects, similar to that in MR Egger regression [27]. For simplicity we assume  $g$ 's are standardized to have mean 0 and variance 1, and the covariance matrix of  $g$ 's is  $\Sigma$ . Since  $g$ 's are possibly correlated, we may have correlated  $b_{Yg}$ 's. Assuming the direct effects from  $g$  to  $Y$  are independently from a normal distribution, we would have

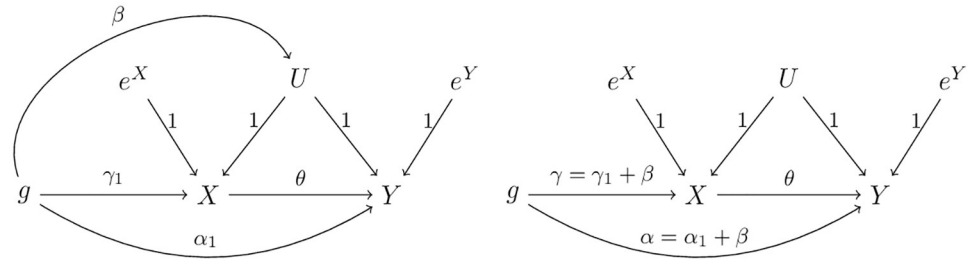
$$b_{Yg} \sim N\left(\frac{b_0 \cdot \Sigma \mathbf{1}}{\sqrt{\text{var}(Y)}}, \frac{\sigma_0^2 \Sigma^2}{\text{var}(Y)}\right). \tag{16}$$

See S2 Text for the derivation. For simplicity, in the following we denote  $\mathbf{v} = \Sigma \mathbf{1} = (v_1, \dots, v_m)^T$ ,  $b_0$  for  $b_0 / \sqrt{\text{var}(Y)}$ , and  $\sigma_0^2$  for  $\sigma_0^2 / \text{var}(Y)$ . Using the sample correlation coefficients to approximate the population correlation coefficients in Eq (16) (while ignoring the errors in  $r_{Xg}$ ), we have

$$r_{Yg} = b_0 \cdot \mathbf{v} + K_{YX} \cdot r_{Xg} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N\left(0, \frac{\mathbf{V}_{Yg}}{n_Y} + \sigma_0^2 \Sigma^2 := \mathbf{V}_{YX}\right), \tag{17}$$

in which the InSIDE assumption (i.e. the independence between  $\beta_{Xg}$  and  $\alpha_{Yg}$ ) is imposed (as for MR Egger regression). Given  $\sigma_0^2$ , estimating  $(b_0, K_{YX})$  is a GLS problem, thus we estimate  $(b_0, K_{YX})$  and  $\sigma_0^2$  iteratively, and obtain their standard errors  $se(\hat{b}_0)$ ,  $se(\hat{K}_{YX})$ ; see S2 Text for details. Briefly, we simply regress the two sample correlations with an intercept term to estimate the slope parameter  $K_{YX}$  using the GLS method (which reduces to the weighted least squares with the weight inversely proportional to  $\text{var}(e_{Yg})$  if the SNPs are independent). This is similar to Egger regression in MR, for which the effect size estimates, instead of the sample correlation coefficients, are regressed against each other. We call this method **Causal Direction-Egger**, or **CD-Egger** for short.

A possible downside of the regression approach is its ignoring the variability of  $r_{Xg}$  while account for that of  $r_{Yg}$ , similar to that in MR Egger regression [42]. To overcome this shortcoming, we propose a new method called **Causal Direction-GLS**, or simply **CD-GLS**. Since it performs similarly to CD-Egger (due to small variability of  $r_{Xg}$  with a usually large sample size of GWAS), we relegate its detail to S2 Text.



**Fig 3. The case where the SNP/IV is directly associated with confounder  $U$  (left panel), which, by absorbing the effect of  $SNP \rightarrow U$  into  $SNP \rightarrow X$  and  $SNP \rightarrow Y$ , leads to the case where the SNP/IV has a direct effect on  $Y$  (right panel).**

<https://doi.org/10.1371/journal.pgen.1009105.g003>

**Remarks on extension to case with invalid IVs associated with confounders**

Now we consider an SNP with a direct effect on the hidden confounder, as shown in the DAG in the left panel of Fig 3. This is the case when the third key IV assumption is violated. Note that, as in [27], for simplicity we assume an effect size 1 of the confounder  $U$  on both  $X$  and  $Y$ , while it can be easily generalized. The case can be converted to that as if the second key IV assumption is violated, as shown in the right panel of Fig 3. It may appear that the two new methods developed in the previous subsection can be applied. However, the InSIDE assumption (i.e. the independence assumption of the IV effect on the exposure  $X$ ,  $\gamma = \gamma_1 + \beta$ , and that on the outcome  $Y$ ,  $\alpha = \alpha_1 + \beta$ ) is clearly violated because of their shared component  $\beta$ . Hence, applying the two new methods developed in the previous section would lead to biased inference, unless the direct effect  $\beta$  is small and thus negligible. We will not pursue this challenging topic further, which is beyond the scope of this paper.

**Decision rules**

For each of three methods **CD-Ratio**, **CD-Egger**, and **CD-GLS**, we would have  $\hat{K}_{XY}$  and  $\hat{K}_{YX}$ . Along with their standard errors (SEs), we construct the confidence intervals at any significance level, say  $\alpha$ , for  $K_{XY}$  and  $K_{YX}$  respectively as  $CI_{XY} = \hat{K}_{XY} \pm Z_{\alpha/2}SE(\hat{K}_{XY})$  and  $CI_{YX} = \hat{K}_{YX} \pm Z_{\alpha/2}SE(\hat{K}_{YX})$ . Assuming Key Condition (5) holds, then our decision rules are: (i) If  $CI_{XY}$  is completely outside interval  $[-1, 1]$ , while  $CI_{YX}$  is inside  $[-1, 0]$  or  $(0, 1]$ , we conclude that the causal direction is from  $X$  to  $Y$ ; (ii) If  $CI_{YX}$  is completely outside interval  $[-1, 1]$ , while  $CI_{XY}$  is inside  $[-1, 0]$  or  $(0, 1]$ , we conclude that the causal direction is from  $Y$  to  $X$ ; (iii) Otherwise, we cannot make a conclusion.

In the above, we have incorporated inferring whether there is indeed any causal relationship between the two traits  $X$  and  $Y$ , in contrast to Steiger’s method [9] in simply *assuming* the existence of a causal relationship. As to be shown in simulations, this is advantageous over Steiger’s method and expected to be useful in practice. Our proposal again exploits the use of our formulated estimand  $K_{YX}$  (or  $K_{XY}$ ): the sufficient and necessary condition for no causal relationship with  $\beta_{XY} = \beta_{YX} = 0$  is by definition equivalent to  $K_{YX} = K_{XY} = 0$ . Hence, we propose using their confidence intervals to determine the existence of a causal relationship.

**Goodness-of-fit tests**

In MR the intercept estimate and its SE from Egger regression can be used as a goodness-of-fit (GOF) test to check the assumptions of the valid IVs [27]; some extensions have also been developed [43]. Accordingly, we can also use our estimate  $\hat{b}_0$  and its CI for such a purpose. In

addition, as argued in [44], it is likely better and more powerful to conduct homogeneity testing using Cochran’s  $Q$  statistics (for the IVW method) and Rucker’s  $Q'$  statistics (for MR-Egger regression). We develop the corresponding three tests in the current context, similar to but different from those for IVW or Egger regression in MR. Using the asymptotic distribution (9), we can construct a chi-squared GOF test statistic for CD-Ratio:

$$Q_{Ratio}^{YX} = \begin{pmatrix} \mathbf{r}_{Yg} \\ \mathbf{r}_{Xg} \end{pmatrix} - \mathbf{1} \cdot \hat{K}_{YX} \Big)^T \mathbf{V}_{YXg}^{-1} \begin{pmatrix} \mathbf{r}_{Yg} \\ \mathbf{r}_{Xg} \end{pmatrix} - \mathbf{1} \cdot \hat{K}_{YX}, \tag{18}$$

where  $\hat{K}_{YX}$  is the CD-Ratio (or any other consistent) estimate. Under the null hypothesis that all SNPs are valid IVs, the (asymptotic) null distribution is  $Q_{Ratio}^{YX} \sim \chi_{m-1}^2$ , a chi-squared distribution with degrees of freedom  $m - 1$ , where  $m = \dim(\mathbf{r}_{Xg})$  is the number of the SNPs/IVs being used. Rejection of the null hypothesis may imply that the SNPs may have direct effects on the outcome, i.e. there is horizontal pleiotropy.

Based on the asymptotic distribution (17), we can construct a similar GOF test for CD-Egger after accounting for the possible presence of horizontal pleiotropy:

$$Q_{Egger}^{YX} = (\mathbf{r}_{Yg} - \hat{b}_0 \cdot \mathbf{V} - \hat{K}_{YX} \cdot \mathbf{r}_{Xg})^T \mathbf{V}_{YX}^{-1} (\mathbf{r}_{Yg} - \hat{b}_0 \cdot \mathbf{V} - \hat{K}_{YX} \cdot \mathbf{r}_{Xg}), \tag{19}$$

where  $\hat{b}_0$  and  $\hat{K}_{YX}$  are the CD-Egger estimates of  $b_0$  and  $K_{YX}$  respectively. Under the null hypothesis that the model assumptions hold, the (asymptotic) null distribution is  $Q_{Egger}^{YX} \sim \chi_{m-2}^2$ . We can construct a similar GOF test based on CD-GLS, which we found performed similarly to  $Q_{Egger}^{YX}$  and thus is relegated to [S3 Text](#). Rejection of the null hypothesis by either test may imply that the SNPs/IVs are possibly correlated with some hidden confounders.

For the other direction  $Y$  to  $X$ , similarly we can perform corresponding chi-squared GOF tests with  $Q_{Ratio}^{XY}$ ,  $Q_{Egger}^{XY}$  and  $Q_{GLS}^{XY}$ .

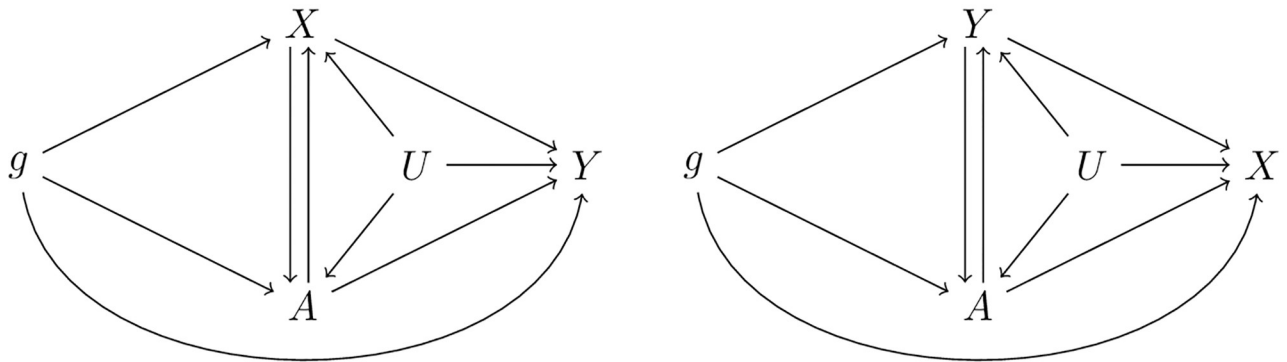
### Further extensions

**Multivariable approach.** Analogous to the extension of MR to multivariable MR (MV-MR) [28, 29], we extend our CD-Egger method to its multivariable version, denoted **MV-CD-Egger**, to explicitly account for pleiotropic/direct effects through a third (or more) trait. Specifically, we introduce the third trait  $A$ , which may be an observed confounder for the two traits of interest,  $X$  and  $Y$ . When the true causal direction is from  $X$  to  $Y$ , the true causal model is:

$$\begin{aligned} A &= \beta_{Ag} \cdot g + \beta_{AX} \cdot X + \beta_{AU} \cdot U + e_A, \\ X &= \beta_{Xg} \cdot g + \beta_{XA} \cdot A + \beta_{XU} \cdot U + e_X, \\ Y &= \alpha_{Yg} \cdot g + \beta_{YX} \cdot X + \beta_{YA} \cdot A + \beta_{YU} \cdot U + e_Y. \end{aligned} \tag{20}$$

Here  $e_A$ ,  $e_X$ ,  $e_Y$  are independent random error terms. In model (20), we allow a general relationship between  $A$  and  $X$ , and between  $A$  and  $Y$ . When the true causal direction is from  $Y$  to  $X$ , we have a corresponding true model. The two possible true models are illustrated in [Fig 4](#).

Under the true model of the causal direction from  $X$  to  $Y$ , since the IV  $g$  is independent of the hidden confounder  $U$  and error term  $e_Y$ , we can calculate the covariance between  $g$  and  $Y$



**Fig 4. Two possible true models: The left one is for causal direction from X to Y, and the right one is for Y to X.**

<https://doi.org/10.1371/journal.pgen.1009105.g004>

as:

$$cov(Y, g) = \alpha_{Yg} \cdot var(g) + \beta_{YX} \cdot cov(X, g) + \beta_{YA} \cdot cov(A, g) \tag{21}$$

Let  $\rho_{Yg}$  be the correlation between Y and g, and similarly define  $\rho_{Xg}$  and  $\rho_{Ag}$ , then we have

$$\rho_{Yg} = \frac{cov(Y, g)}{\sqrt{var(Y) \cdot var(g)}} = \alpha_{Yg} \cdot \frac{\sqrt{var(g)}}{\sqrt{var(Y)}} + \beta_{YX} \cdot \frac{cov(X, g)}{\sqrt{var(Y) \cdot var(g)}} + \beta_{YA} \cdot \frac{cov(A, g)}{\sqrt{var(Y) \cdot var(g)}} \tag{22}$$

Let  $K_{YX} = \beta_{YX} \cdot \frac{\sqrt{var(X)}}{\sqrt{var(Y)}}$ ,  $K_{YA} = \beta_{YA} \cdot \frac{\sqrt{var(A)}}{\sqrt{var(Y)}}$ , and  $b_{Yg} = \alpha_{Yg} \cdot \frac{\sqrt{var(g)}}{\sqrt{var(Y)}}$ . Paralleling with the key condition (5), we assume  $|K_{YX}| < 1$  and  $|K_{YA}| < 1$ . Hence we have

$$\rho_{Yg} = b_{Yg} + K_{YX} \cdot \rho_{Xg} + K_{YA} \cdot \rho_{Ag} \tag{23}$$

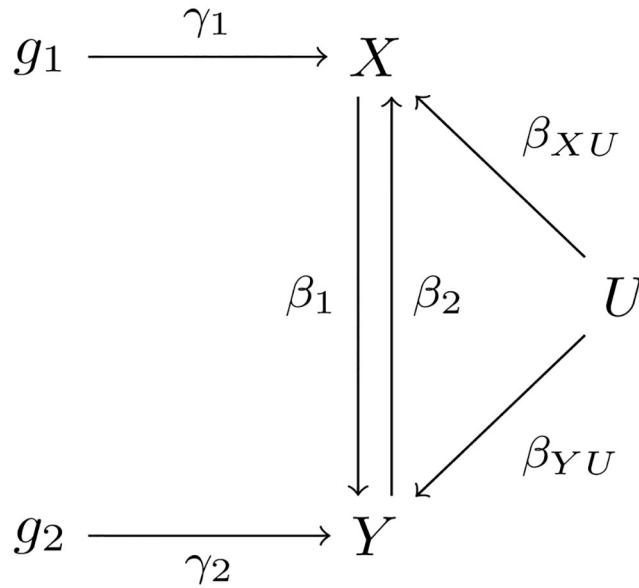
As for CD-Egger, we model the distribution of  $b_{Yg}$  similarly, and apply an iterative method to obtain the maximum likelihood estimates of parameters  $K_{YA}$  and  $K_{YX}$  with their standard errors as shown in S2 Text; the only change is to redefine  $\mathbf{X} = (\mathbf{v}, \mathbf{r}_{Xg}, \mathbf{r}_{Ag})$ . At the end, we construct the normal-based confidence intervals for  $K_{YX}$  and  $K_{XY}$ , and apply the same decision rules as before.

**Bi-directional causal effects.** So far we have restricted our discussion on uni-directional causal effects. Inspired by [45], as suggested by a reviewer, we now briefly discuss an extension of our CD-Ratio method to **bi-CD-Ratio** that works regardless whether there is no, a uni- or bi-directional causal relationship between two traits. In the presence of bi-directional causal effects, as discussed by [45], for the model identifiability, at least one valid IV is required for each trait as shown in Fig 5, which is the case we consider here.

The true causal model is:

$$\begin{aligned} X &= \beta_2 \cdot Y + \gamma_1 \cdot g_1 + \beta_{XU} \cdot U + \epsilon_1, \\ Y &= \beta_1 \cdot X + \gamma_2 \cdot g_2 + \beta_{YU} \cdot U + \epsilon_2, \end{aligned} \tag{24}$$

where  $\epsilon_1$  and  $\epsilon_2$  are independent errors for X and Y respectively. The reduced form of the true



**Fig 5. The true model for two traits  $X$  and  $Y$  with a bi-directional causal relationship:  $g_1$  and  $g_2$  are valid IVs for  $X$  and  $Y$  respectively, and  $U$  is unobserved confounder.**

<https://doi.org/10.1371/journal.pgen.1009105.g005>

model is

$$\begin{aligned}
 X &= \frac{1}{1 - \beta_1\beta_2} (\gamma_1 \cdot g_1 + \beta_2\gamma_2 \cdot g_2 + \beta_{XU} \cdot U + \beta_2\beta_{YU} \cdot U + \epsilon_1 + \beta_2 \cdot \epsilon_2), \\
 Y &= \frac{1}{1 - \beta_1\beta_2} (\gamma_2 \cdot g_2 + \beta_1\gamma_1 \cdot g_1 + \beta_{YU} \cdot U + \beta_1\beta_{XU} \cdot U + \epsilon_2 + \beta_1 \cdot \epsilon_1),
 \end{aligned}
 \tag{25}$$

from which it is easy to calculate the correlation between  $X$  and  $g_1$ , denoted by  $\rho_{Xg_1}$ , and similarly  $\rho_{Xg_2}$ ,  $\rho_{Yg_1}$  and  $\rho_{Yg_2}$  as:

$$\begin{aligned}
 \rho_{Xg_1} &= \frac{\gamma_1}{1 - \beta_1\beta_2} \cdot \frac{\sqrt{\text{var}(g_1)}}{\sqrt{\text{var}(X)}}, & \rho_{Xg_2} &= \frac{\beta_2\gamma_2}{1 - \beta_1\beta_2} \cdot \frac{\sqrt{\text{var}(g_2)}}{\sqrt{\text{var}(X)}}, \\
 \rho_{Yg_1} &= \frac{\beta_1\gamma_1}{1 - \beta_1\beta_2} \cdot \frac{\sqrt{\text{var}(g_1)}}{\sqrt{\text{var}(Y)}}, & \rho_{Yg_2} &= \frac{\gamma_2}{1 - \beta_1\beta_2} \cdot \frac{\sqrt{\text{var}(g_2)}}{\sqrt{\text{var}(Y)}},
 \end{aligned}$$

and thus

$$K_1 = \frac{\rho_{Yg_1}}{\rho_{Xg_1}} = \frac{\beta_1 \cdot \sqrt{\text{var}(X)}}{\sqrt{\text{var}(Y)}}, \quad K_2 = \frac{\rho_{Xg_2}}{\rho_{Yg_2}} = \frac{\beta_2 \cdot \sqrt{\text{var}(Y)}}{\sqrt{\text{var}(X)}}.
 \tag{26}$$

As before, since  $\beta_1^2 \cdot \text{var}(X)$  represents a part of the variance in  $Y$  due to the direct effect of  $X$  on  $Y$ , we assume  $\beta_1^2 \cdot \text{var}(X) < \text{var}(Y)$ , and thus  $|K_1| < 1$ . Similarly we assume  $|K_2| < 1$ . These two assumptions parallel with our key condition (5).

From the GWAS summary data for traits  $X$  and  $Y$ , we calculate the sample correlations  $r_{Xg_1}$ ,  $r_{Xg_2}$ ,  $r_{Yg_1}$  and  $r_{Yg_2}$ . Then we apply our CD-Ratio method with  $g_1$  to estimate  $K_1$  as  $\hat{K}_1$  with its standard error  $\text{SE}(\hat{K}_1)$ , and construct its  $100(1 - \alpha)\%$  confidence interval

$CI_{K_1} = \hat{K}_1 \pm Z_{\alpha/2} \text{SE}(\hat{K}_1)$ . If  $CI_{K_1}$  is completely inside  $[-1, 0)$  or  $(0, 1]$ , we conclude that  $X$  has a

causal effect on  $Y$ ; if  $CI_{K_1}$  covers 0, then we conclude that  $X$  has no causal effect on  $Y$ . Similarly we construct a confidence interval for  $K_2$  and infer whether there is a causal effect from  $Y$  to  $X$ . In this way, we may conclude with bi-directional causal effects between  $X$  and  $Y$ .

As to be shown in simulations, in the case with valid IVs, both our bi-CD-Ratio and bi-directional MR performed well, being able to distinguish among no, uni- and bi-directional causal relationships between two traits. In contrast, Steiger's method did not work because it cannot determine whether there is any causal relationship between two traits.

## Other methods

In addition to Steiger's method, bi-directional MR, also called reciprocal MR, is popular. For the latter, it consists of two steps: first, an MR analysis is applied to infer the causal effect of  $X$  on  $Y$ , often using a set of (approximately) independent SNPs significantly associated with  $X$  as IVs; second, a separate MR analysis is applied to infer the causal effect of  $Y$  on  $X$  using a possibly different set of independent SNPs that are significantly associated with  $Y$ . The decision rules are similar to ours: 1) if the result is significant in Step 1, but not in Step 2, we conclude that the causal direction is from  $X$  to  $Y$ ; 2) If the result is significant in Step 2, but not in Step 1, we conclude that the causal direction is from  $Y$  to  $X$ ; 3) Otherwise, we cannot determine the causal direction. However, an issue with bi-directional MR is that the differing sample sizes, or generally differing statistical power, between the two GWAS for traits  $X$  and  $Y$  respectively, would possibly confound the result. An extreme example for an ideal case is that, suppose that both GWAS are sufficiently powered to detect all associated SNPs; it is easy to see that, if there is indeed a causal relation from  $X$  to  $Y$  and that  $Y$  is completely determined by  $X$  genetically, then all the SNPs associated with  $X$  would be associated with  $Y$  too, leading to that both Steps 1 and 2 would give significant results.

We will apply a wide range of MR methods, representing both well-established and new ones as implemented in R package **TwoSampleMR**, for bi-directional MR. These include the inverse-variance-weighted (IVW) method [46], MR-Egger regression [27], the weighted median estimator [47], the weighted mode-based estimation [48, 49], and the Robust Adjusted Profile Score (RAPS) method [42]. To save space, we will only present the results of RAPS with the set-up of overdispersion and a squared-error loss; using other set-ups led to qualitatively similar results. The IVW is perhaps the most popular, requiring all the three key assumptions on valid IVs to hold, under which it is most efficient; the other ones either explicitly account for or are robust to horizontal pleiotropy. Since all these MR methods require using independent SNPs, we selected the most significant SNP from each locus if it contained more than one significant SNPs, then applying MR to only independent SNPs in any simulated or real data.

## GWAS data and reference panel

To draw inference on the causal direction between two traits using either our proposed methods or bi-directional MR, we only need two independent GWAS summary datasets for the two traits respectively, and a reference panel to estimate LD-matrices for correlated SNPs.

We showcase an application to infer causal directions between LDL and CAD, and between HDL and CAD. The large-scale GWAS summary data for traits LDL/HDL and CAD were drawn from two year 2017 publications [50, 51] with sample sizes up to 295826 for LDL, 316391 for HDL and 336860 for CAD of primary European ancestry respectively. We used the European samples of 489 unrelated individuals in the 1000 Genomes Project [41] as our reference panel to calculate the LD-matrix.

For our proposed methods and Steiger's method, we chose the IVs from the SNPs significantly associated with both traits, denoted as before as  $X$  and  $Y$ . Based on [52], we partitioned

the whole genome into 1703 approximately LD-independent loci and assigned the significant SNPs (associated with both  $X$  and  $Y$ ) to one of these 1703 independent loci. We pruned out highly correlated SNPs falling inside each locus in the following two steps. Step 1, we ordered the significant SNPs in each locus according to their combined significance level with  $X$  and  $Y$ . Suppose there were  $m$  significant SNPs with their p-values  $\mathbf{p}_X = (p_{X1}, \dots, p_{Xm})$  and  $\mathbf{p}_Y = (p_{Y1}, \dots, p_{Ym})$  for two traits  $X$  and  $Y$  respectively. We sorted  $\mathbf{p}_X$  and  $\mathbf{p}_Y$  in an ascending order. For each SNP  $i$ ,  $i = 1, \dots, m$ , we obtained the ranks of its p-values  $p_{Xi}$  and  $p_{Yi}$ , denoted as  $I_{Xi}$  and  $I_{Yi}$ . Then we calculated its rank sum  $I_i = I_{Xi} + I_{Yi}$ ,  $i = 1, \dots, m$ . After that, we reordered the SNPs in the ascending order of  $I_i$ . Step 2, we pruned out some highly correlated (and significant) SNPs in each locus. We kept SNP 1, and for each  $i = 2, \dots, m$ , if  $|cor(SNP_i, SNP_j)| \leq 0.8$  for all  $j = 1, \dots, i - 1$ , we kept SNP  $i$ ; otherwise we removed it. At the end, we had a set of the significant SNPs with pairwise absolute correlations less than 0.8.

### Main simulation set-ups

For our main simulations, to mimic the real data analysis for LDL and CAD, in which 22 SNPs from 12 loci were used, we generated simulated data with the same 22 SNPs. We extracted the genotype data of 489 European individuals from the 1000-Genomes Project, then replicated the genotype data of the 22 SNPs from the 489 individuals 200 times, obtaining a sample of size  $n = 97800$ .

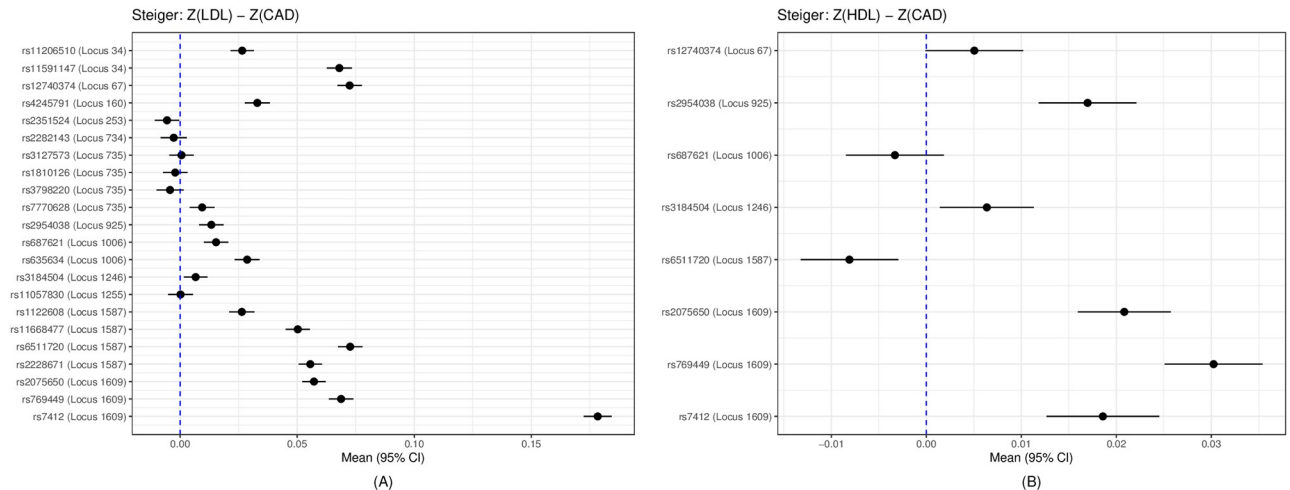
Following [53], in each simulation, we generated the two traits from true models (12-13), where  $\mathbf{g}$  was the vector of the genotype scores of the 22 SNPs,  $\beta_{Xg}$  was the (row) vector of the effect sizes of the 22 SNPs on trait  $X$  with each element independently drawn from a truncated standard normal distribution (truncated at 0.5 to ensure  $|\beta_{Xg_i}| \geq 0.5$ ), the confounder  $U \sim N(0, 1)$ ; the causal effect size  $\beta_{YX}$  from  $X$  to  $Y$  was fixed at one value, the direct effects of the SNPs on  $Y$ , the elements of (row) vector  $\alpha$ , were generated independently  $\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$ ; and both error terms were independently from  $N(0, 4)$ . With various values of  $\beta_{YX}$ ,  $\mu_\alpha$  and  $\sigma_\alpha$  we considered a total of 75 simulation set-ups, covering a wide range of situations from no pleiotropy (i.e.  $\mu_\alpha = \sigma_\alpha = 0$ ) to balanced pleiotropy (i.e.  $\mu_\alpha = 0$  and  $\sigma_\alpha > 0$ ) and direction pleiotropy (i.e.  $\mu_\alpha \neq 0$  and  $\sigma_\alpha > 0$ ).

After generating each simulated dataset, we performed single SNP-based univariate/marginal analysis for each trait, obtaining the corresponding two GWAS summary datasets. The proposed methods, CD-Ratio, CD-Egger and CD-GLS, were then applied to the GWAS summary data, and for each method one of three possible decisions was made according to the decision rule. For comparison, we also applied MR Steiger's method to each of the 22 SNPs; we then calculated the proportions among the 22 SNPs of the three possible decisions reached, denoted by "Steiger-Prop". In addition, we also pooled together the individual results from each of the 22 SNPs by majority voting to reach a final decision; this method is denoted as "Steiger-MV".

## Results

### Real data examples

**LDL and CAD.** For LDL and CAD, at the significance cutoff of  $p < 5 \times 10^{-8}$ , out of 1703 (approximately) independent loci [52], we identified 39 significant loci for CAD, 102 for LDL, and 12 common ones for both traits from the two GWAS summary datasets. There were 40 SNPs in these 12 common loci significant for both LDL and CAD. After pruning out highly correlated ones, we obtained 22 SNPs in the 12 loci. We applied MR Steiger's method and our proposed methods with these 22 SNPs as IVs. For each of the 22 SNPs, we applied MR Steiger's



**Fig 6. The results by Steiger's method: 95% CI of the test statistics  $Z_g$  for each SNP.** (A) Results of 22 SNPs for LDL and CAD. (B) Results of 8 SNPs for HDL and CAD.

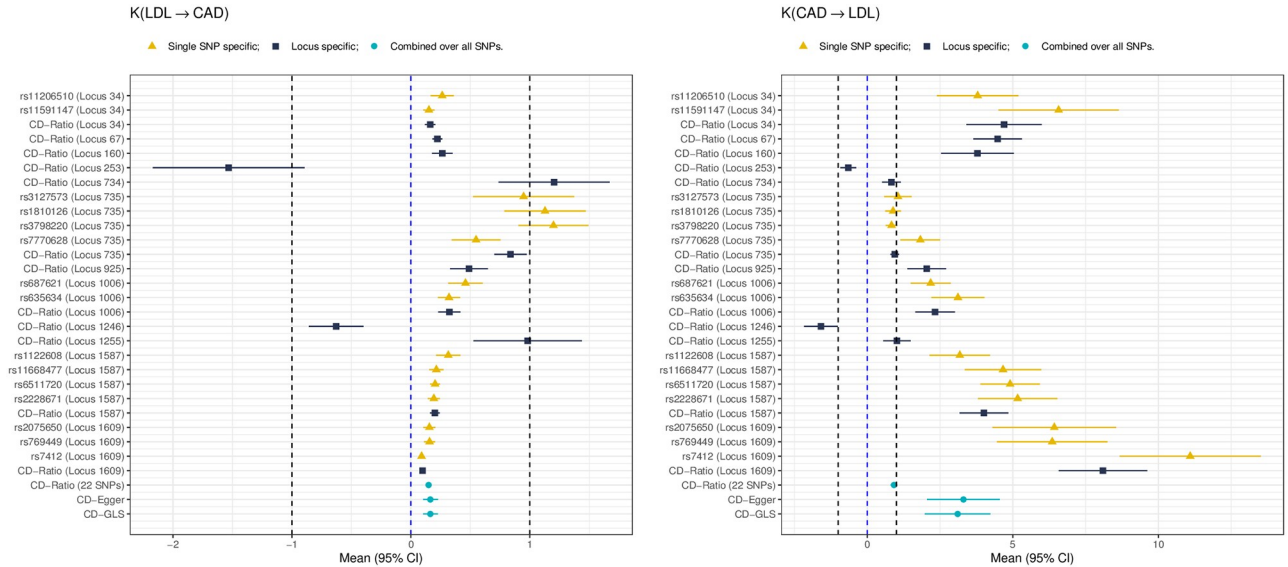
<https://doi.org/10.1371/journal.pgen.1009105.g006>

method to obtain a 95% CI for  $Z_g$ , as shown in Fig 6(A). If a CI covers 0, no conclusion would be made; otherwise, depending on whether it is in the right or left side of the zero point, we conclude the causal direction of LDL to CAD or the opposite. A problem with single SNP-based analysis is its lower power and thus possibly inconsistent results: as shown in Fig 6(A), although we'd conclude the causal direction of LDL driving CAD based on most of the SNPs, we would fail to conclude based on each of five SNPs while concluding CAD driving LDL on one SNP.

As an alternative, under the Ideal Model, at each locus, we first applied CD-Ratio to all individual SNPs; if there were more than one SNP, we applied CD-Ratio to combine the SNP-specific results at the locus; then we applied CD-Ratio to all 22 SNPs from the 12 loci. For the Pleiotropy Model, we applied CD-Egger and CD-GLS directly to all the SNPs across all the loci. Fig 7 shows the results. First, we emphasize that, perhaps as expected, the single SNP-based results varied with the specific SNP being used as the IV, presumably due to that some SNPs were not valid IVs. For example, for the majority of the SNPs, CD-Ratio would give each a 95% confidence interval (CI) of  $K_{LDL \rightarrow CAD}$  completely inside (0, 1], but a few outside, supporting the causal direction from LDL to CAD. Also note that none of the CIs covered 0, supporting the existence of a causal relationship between the two traits. Second, for the several loci with multiple SNPs, by combining the results over the SNPs, our proposed CD-Ratio improved statistical efficiency by giving shorter locus-specific CIs than the individual SNP-specific ones. Finally, most importantly, under the Ideal Model, after combining over all 22 SNPs from 12 loci, the CD-Ratio method could not conclude on the causal direction, because its CI for  $K_{LDL \rightarrow CAD}$  was inside (0,1] while that for  $K_{CAD \rightarrow LDL}$  covered 1, as given in Table 1. In contrast, after accounting for possible horizontal pleiotropy, our proposed methods CD-Egger and CD-GLS would both conclude the causal direction from LDL to CAD, in agreement with the previous MR analysis and RCT results reported in the literature [30, 34–36].

It is noted that, Locus 253 and Locus 1246 showed a negative (i.e. beneficial) effect of LDL on CAD, in contrary to what is known in the literature. Treating each as an outlier, after removing it, we applied the new methods, and all the three (including CD-Ratio) reached the same conclusion on the causal direction from LDL to CAD; see S4 Text. On the other hand, if we only selected and used one SNP from each locus, CD-Ratio could not conclude while both CD-Egger and CD-GLS reached the same conclusion of LDL to CAD, as shown in S4 Text.





**Fig 7. Forest plots for inferring the causal direction between LDL and CAD, showing the differently colored SNP- and locus-specific results and the combined results across all SNPs and loci, respectively.**

<https://doi.org/10.1371/journal.pgen.1009105.g007>

It is emphasized that the necessity of accounting for horizontal pleiotropy in this example was supported by the statistically significant non-zero estimates of  $b_0$  in either direction, which was interpreted as representing the mean of the pleiotropic effects; it was notably larger when considering causal direction  $CAD \rightarrow LDL$  than that in the other direction (Table 1). This latter point was also confirmed by the results of MR Egger regression as to be shown. Furthermore, we conducted the GOF test with  $Q_{Ratio} = 337.1$ , leading to a  $p$ -value almost 0, rejecting the null hypothesis that the Ideal Model was adequate. The GOF test  $Q_{Ratio} = 527.2$  for  $CAD \rightarrow LDL$ , leading to a  $p$ -value almost 0 too. For CD-Egger, for the direction  $LDL \rightarrow CAD$  we got  $Q_{Egger} = 22.2$  and  $p$ -value = 0.329; for the other direction, we obtained  $Q_{Egger} = 22.0$  and  $p$ -value = 0.340. The similar results were obtained from  $Q_{GLS}$ , supporting the adequacy of the Pleiotropic Model and the good performance of CD-Egger and CD-GLS.

For comparison, we also applied bi-directional MR. In each of 39 loci significant for CAD, we picked up the most significant SNP (i.e. with the smallest  $p$ -value) as the IV, then applied various MR methods to estimate and test the causal effect of CAD on LDL. Similarly, in each of 102 significant loci for LDL, we used the most significant SNP as the IV, then applied MR to estimate and test the causal effect of LDL on CAD. The results are shown in Table 2. We can see that, at the significance level 0.05, IVW, Egger and RAPS all showed some significant causal effects of CAD on LDL, and of LDL on CAD. Although Weighted Median and Weighted Mode concluded that CAD had no causal effect on LDL, and that LDL had causal effect on CAD, it could be due to the low statistical efficiency of the two methods; one evidence is that the two methods gave a negative effect of CAD on LDL, contrary to the known positive association between the two. Overall, across all the methods, there was much stronger statistical evidence to support the causal direction  $LDL \rightarrow CAD$  than the other direction. Besides, MR-Egger gave an estimate of the mean pleiotropic effect for direction LDL to CAD as -0.003 with standard error 0.003, and -0.054 with standard error 0.028 for CAD to LDL. In summary, as expected, due to use of possibly invalid IVs, bi-directional MR tended to conclude causal effects for both directions.

**Table 1. The results for inferring the causal direction between LDL and CAD after combining across all 22 SNPs in 12 loci.**

Method	LDL → CAD		CAD → LDL	
	$\hat{K}$ [95% CI]	$\hat{b}_0$ (SE)	$\hat{K}$ [95% CI]	$\hat{b}_0$ (SE)
CD-Ratio	0.151 [0.137, 0.165]	NA	0.919 [0.807, 1.031]	NA
CD-Egger	0.166 [0.103, 0.229]	0.006 (0.002)	3.303 [2.051, 4.555]	-0.025 (0.010)
CD-GLS	0.165 [0.103, 0.227]	0.006 (0.002)	3.100 [1.971, 4.229]	-0.023 (0.008)

<https://doi.org/10.1371/journal.pgen.1009105.t001>

Finally, since our proposed CD-Egger and CD-GLS suggested the causal direction from LDL to CAD, we checked the key condition (5). Using all the 40 significant SNPs in the 12 loci, we obtained the distribution of  $var(CAD)/var(LDL)$  ranging from 7.62 to 13.95 with mean 9.44, much larger than any MR  $\hat{\beta}^2$  for LDL → CAD listed in Table 2, supporting that condition (5) held here.

**HDL and CAD.** For HDL and CAD, at the significance cutoff  $5 \times 10^{-8}$ , out of 1703 (approximately) independent loci, we identified 39 significant loci for CAD, 119 for HDL, and 6 common ones for both traits. There were 19 SNPs significant for both HDL and CAD from the 6 common loci; after pruning out highly correlated ones, we obtained 8 SNPs. We applied the MR Steiger's method and our proposed methods to the 8 SNPs in the 6 common and significant loci. Fig 6(B) shows the results from Steiger's method, while Fig 8 showing those for CD-Ratio, CD-Egger and CD-GLS.

Again, similar to that for LDL and CAD, the single SNP results varied with different SNPs being used. Combining multiple SNPs in a locus showed gains with higher statistical estimation efficiency: for example, at Locus 1609 with 3 SNPs, CD-Ratio gave a shorter CI than the 3 individual SNP-specific ones. As shown in Table 3, we cannot conclude on the causal direction using CD-Ratio under the Ideal Model, because its CIs for  $K_{HDL \rightarrow CAD}$  and  $K_{CAD \rightarrow HDL}$  were both inside  $[-1, 0)$ . After accounting for possible pleiotropy, both CD-Egger and CD-GLS gave a stronger but not statistically significant result of the causal direction from HDL to CAD; both reached the same conclusion: although each gave a CI of  $K_{HDL \rightarrow CAD}$  within  $(-1, 0)$ , each gave an estimate of  $K_{CAD \rightarrow HDL} < -1$  but with a CI covering  $-1$ , failing to determine the causal direction.

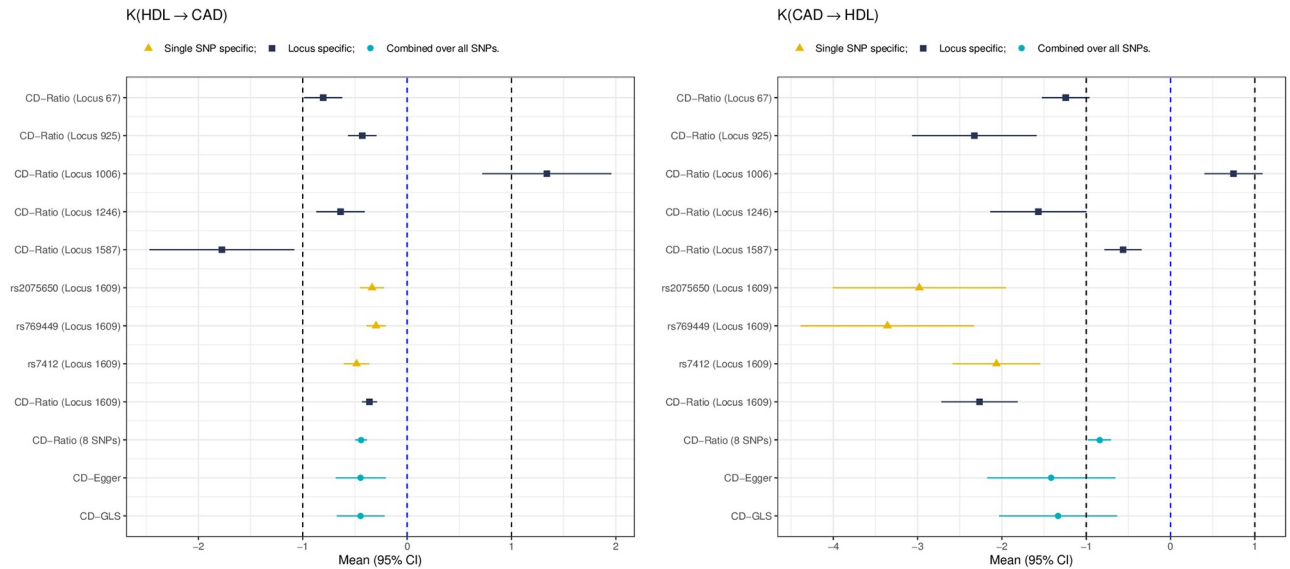
We also conducted GOF testing. First, we obtained  $Q_{Ratio} = 78.1$ , leading to a  $p$ -value almost 0, thus rejecting the null hypothesis that the Ideal Model for  $HDL \rightarrow CAD$  was adequate. Similarly, we reached the same conclusion for  $CAD \rightarrow HDL$  with  $Q_{Ratio} = 163.9$  and a  $p$ -value almost 0. Next, for the Pleiotropic Model with CD-Egger, for the direction  $HDL \rightarrow CAD$ , we obtained  $Q_{Egger} = 7.99$  and  $p$ -value = 0.239; for the other direction, we had  $Q_{Egger} = 8.07$  and  $p$ -value 0.233. Similar results were obtained with CD-GLS. Hence, we concluded the adequacy of the Pleiotropic Model.

We applied bi-directional MR for comparison. The results are shown in Table 4. At the significance level 0.05, IVW and RAPS showed significant causal effects of both HDL to CAD

**Table 2. Results of bi-directional MR for the causal relationship between LDL and CAD by various MR methods. Each cell gives " $\hat{\beta}$  (SE),  $p$ -value".**

Causality	IVW	Egger	Weighted Median	Weighted Mode	RAPS
LDL→CAD	0.51 (0.048),	0.57 (0.070),	0.47 (0.044),	0.54 (0.031),	0.52 (0.023),
	$3.6 \times 10^{-26}$	$1.5 \times 10^{-12}$	$3.0 \times 10^{-27}$	$1.0 \times 10^{-31}$	$2.3 \times 10^{-113}$
CAD→LDL	0.37 (0.14),	0.98 (0.34),	-0.011 (0.017),	-0.020 (0.013),	1.8 (0.12),
	0.0076	0.0063	0.52	0.14	$6.9 \times 10^{-54}$

<https://doi.org/10.1371/journal.pgen.1009105.t002>



**Fig 8. Forest plots for inferring the causal direction between HDL and CAD, showing the differently colored SNP- and locus-specific results and the combined results across all SNPs and loci, respectively.**

<https://doi.org/10.1371/journal.pgen.1009105.g008>

and CAD to HDL. MR Egger concluded with a significant causal effect of HDL to CAD but not CAD to HDL, while Weighted Median gave the opposite result. Weighted Mode showed that a causal effect from CAD to HDL was more significant than that from HDL to CAD, though neither was significant at level 0.05. In summary, different bi-directional MR methods gave different and contradictory results.

Finally, given the inconclusive results by our methods, we checked the key condition (5) for both directions. Using all the 19 significant SNPs in the 6 loci, we obtained the distribution of  $var(CAD)/var(HDL)$  ranging from 8.37 to 11.77 with mean 9.84, much larger than all MR  $\hat{\beta}^2$  for HDL  $\rightarrow$  CAD; on the other hand,  $var(HDL)/var(CAD)$  ranged from 0.0849 to 0.1195 with mean 0.1032, larger than any MR  $\hat{\beta}^2$  for CAD  $\rightarrow$  HDL. Hence there was evidence supporting condition (5) (or its analog) in either direction.

**LDL/HDL and CAD: Multivariable approaches.** Since MV-MR has been shown to be more robust than MR, in particular in estimating causal effects of LDL and HDL on CAD [54], we also applied bi-directional MV-MR and MV-CD-Egger to infer the causal direction between LDL (or HDL) and CAD while adjusting for HDL (or LDL) as a potential confounder. As shown in S5 Text, our proposed MV-CD-Egger concluded with the causal direction from LDL to CAD, but no causal relationship between HDL and CAD. In contrast, both MV-MR-IVW and MV-MR-Egger concluded with LDL and HDL causal to CAD; in addition, MV-MR-Egger also identified a significant causal effect of CAD on LDL.

**Table 3. Results of inferring the causal direction between HDL and CAD after combining across all 8 SNPs in 6 loci.**

Method	HDL $\rightarrow$ CAD		CAD $\rightarrow$ HDL	
	$\hat{K}$ [95% CI]	$\hat{b}_0$ (SE)	$\hat{K}$ [95% CI]	$\hat{b}_0$ (SE)
CD-Ratio	-0.441 [-0.498, -0.385]	NA	-0.842 [-0.979, -0.706]	NA
CD-Egger	-0.445 [-0.685, -0.205]	-0.001 (0.003)	-1.414 [-2.174, -0.654]	-0.001 (0.006)
CD-GLS	-0.444 [-0.674, -0.215]	-0.001 (0.003)	-1.334 [-2.035, -0.634]	-0.001 (0.006)

<https://doi.org/10.1371/journal.pgen.1009105.t003>

**Table 4. Results of bi-directional MR for the causal relationship between HDL and CAD by various MR methods.** Each cell gives " $\hat{\beta}$  (SE), p-value".

Causality	IVW	Egger	Weighted Median	Weighted Mode	RAPS
HDL→CAD	-0.281(0.058),	-0.197(0.096),	-0.085(0.052),	-0.083(0.061),	-0.301(0.053),
	$1 \times 10^{-6}$	0.041	0.100	0.176	$1.4 \times 10^{-8}$
CAD→HDL	-0.0823(0.033),	-0.104(0.084),	-0.032(0.015),	-0.027(0.014),	-0.089(0.033),
	0.013	0.225	0.029	0.061	0.006

<https://doi.org/10.1371/journal.pgen.1009105.t004>

## Main simulation results

**Estimation of  $K_{YX}$ : CD-Ratio performed well with no pleiotropy while CD-Egger and CD-GLS always performed well.** First we show the simulation results in terms of estimating  $K_{YX}$  by CD-Ratio, CD-Egger and CD-GLS. For each simulation set-up and each method, we show the mean and standard deviation (sd) of its estimates of  $K_{YX}$  from 500 simulated datasets, which are compared to the true value of  $K_{YX}$  and the mean standard error (se) respectively; we would like them to be as close as possible. From [Table 5](#), we can see that CD-Egger and CD-GLS always gave almost unbiased estimates (with good SE estimates too) in all situations, suggesting their correct (and similar) inference on the causal direction. In contrast, while the CD-Ratio estimator was almost unbiased in the absence of pleiotropy, it became more and more biased as the mean pleiotropic effect (i.e.  $\mu_\alpha$ ) increased. We obtained similar results and reached the same conclusion from many other simulation set-ups as shown in [S6 Text](#).

**Inferring causal direction: CD-Ratio outperformed with no pleiotropy while CD-Egger and CD-GLS always outperformed Steiger's and bidirectional MR methods.** We next compare CD-Ratio, CD-Egger and CD-GLS with MR Steiger's and bi-directional MR in terms of their decisions made on the causal direction. [Fig 9](#) shows some representative results; more are given in [S6 Text](#). First, we consider the case of no causal relationship between the two traits (i.e.  $\beta_{YX} = \beta_{XY} = 0$ ). It was confirmed that MR Steiger's does not work if there is no causal relationship between two traits: it often incorrectly concluded with one of the two directions; in contrast, our CD-Ratio based on single SNPs overcame this shortcoming if there was no pleiotropy, and performed similarly to Steiger's method otherwise (see [S6 Text](#)). In general, bi-directional MR performed reasonably well, though most of them might make incorrect conclusions. For example, in the presence of balanced pleiotropy, all the five MR methods considered made incorrect decisions at least 10% time. In contrast, our proposed three new methods almost always reached the correct decision. Second, when the causal direction was from X to Y, our methods CD-Egger and CD-GLS had high (and similar) power to detect it across all the scenarios, regardless of the presence of balanced or directional pleiotropy. As expected, CD-Ratio performed well only when there was no pleiotropy or only balanced pleiotropy, but it completely lost power in the presence of directional pleiotropy. Steiger's method was high powered except that it would often conclude with the incorrect causal direction in the presence of directional pleiotropy. On the other hand, all the five MR methods were low powered, and would often conclude with the incorrect causal direction in the presence of directional pleiotropy. Perhaps it is surprising to see so low power of bi-directional MR even in the ideal case of no pleiotropy. The reason is that bi-directional MR tends to detect both directions being significant (due to the use of invalid IVs), leading to no final conclusions by the decision rules (since we assume either no or a uni-directional causal relationship); this is confirmed by [Fig 10](#), in which the significant results in either direction are shown separately (without being combined by the decision rules).

**Table 5. Simulation results for estimating  $K_{YX}$ .** The means, standard deviations (sd) and SEs (se) from each method are shown for each set-up based on 500 simulated datasets.  $\mu_\alpha$  represents the mean of the pleiotropic/direct effects (with standard deviation  $\sigma_\alpha = 0.1$ ).

No pleiotropy										
$\mu_\alpha$	$K_{YX}$	CD-Ratio			CD-Egger			CD-GLS		
		Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))
0	0	0.0001	0.0034	0.0035	0.0001	0.0034	0.0036	0.0001	0.0034	0.0036
0	0.2257	0.2256	0.0033	0.0036	0.2256	0.0033	0.0036	0.2256	0.0033	0.0036
0	-0.2344	-0.2341	0.0032	0.0036	-0.2341	0.0032	0.0036	-0.2341	0.0032	0.0036
Balanced pleiotropy										
$\mu_\alpha$	$K_{YX}$	CD-Ratio			CD-Egger			CD-GLS		
		Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))
0	0	0.0013	0.0563	0.0037	-0.0007	0.0431	0.0410	-0.0005	0.0436	0.0419
0	0.2257	0.2181	0.0531	0.0037	0.2201	0.0397	0.0393	0.2203	0.0402	0.0401
0	-0.2344	-0.2233	0.0545	0.0037	-0.2296	0.0409	0.0408	-0.2295	0.0414	0.0416
Directional pleiotropy										
$\mu_\alpha$	$K_{YX}$	CD-Ratio			CD-Egger			CD-GLS		
		Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))	Mean( $\hat{K}$ )	sd( $\hat{K}$ )	Mean (se( $\hat{K}$ ))
0.2	0	-0.0112	0.0538	0.0039	-0.0005	0.0386	0.0368	-0.0004	0.0392	0.0375
0.4	0	-0.0183	0.0448	0.0041	-0.0004	0.0306	0.0291	-0.0003	0.0311	0.0297
0.6	0	-0.0317	0.0357	0.0042	-0.0003	0.0240	0.0228	-0.0002	0.0244	0.0233
1	0	-0.0506	0.0241	0.0043	-0.0002	0.0161	0.0153	-0.0002	0.0164	0.0156
0.2	0.2027	0.1898	0.0514	0.0039	0.1987	0.0369	0.0355	0.1989	0.0375	0.0362
0.4	0.1614	0.1438	0.0448	0.0041	0.1593	0.0301	0.0285	0.1595	0.0307	0.0290
0.6	0.1272	0.0944	0.0362	0.0042	0.1261	0.024	0.0225	0.1263	0.0244	0.0230
1	0.0856	0.0313	0.0244	0.0043	0.0851	0.0162	0.0152	0.0852	0.0165	0.0155
0.2	-0.2093	-0.2172	0.0517	0.0039	-0.2061	0.037	0.0366	-0.206	0.0373	0.0373
0.4	-0.1649	-0.1796	0.0435	0.0041	-0.1636	0.0298	0.0290	-0.1635	0.0300	0.0296
0.6	-0.1290	-0.1552	0.0351	0.0042	-0.1285	0.0236	0.0228	-0.1285	0.0238	0.0233
1	-0.0862	-0.1301	0.0238	0.0043	-0.0862	0.0159	0.0153	-0.0862	0.0161	0.0156

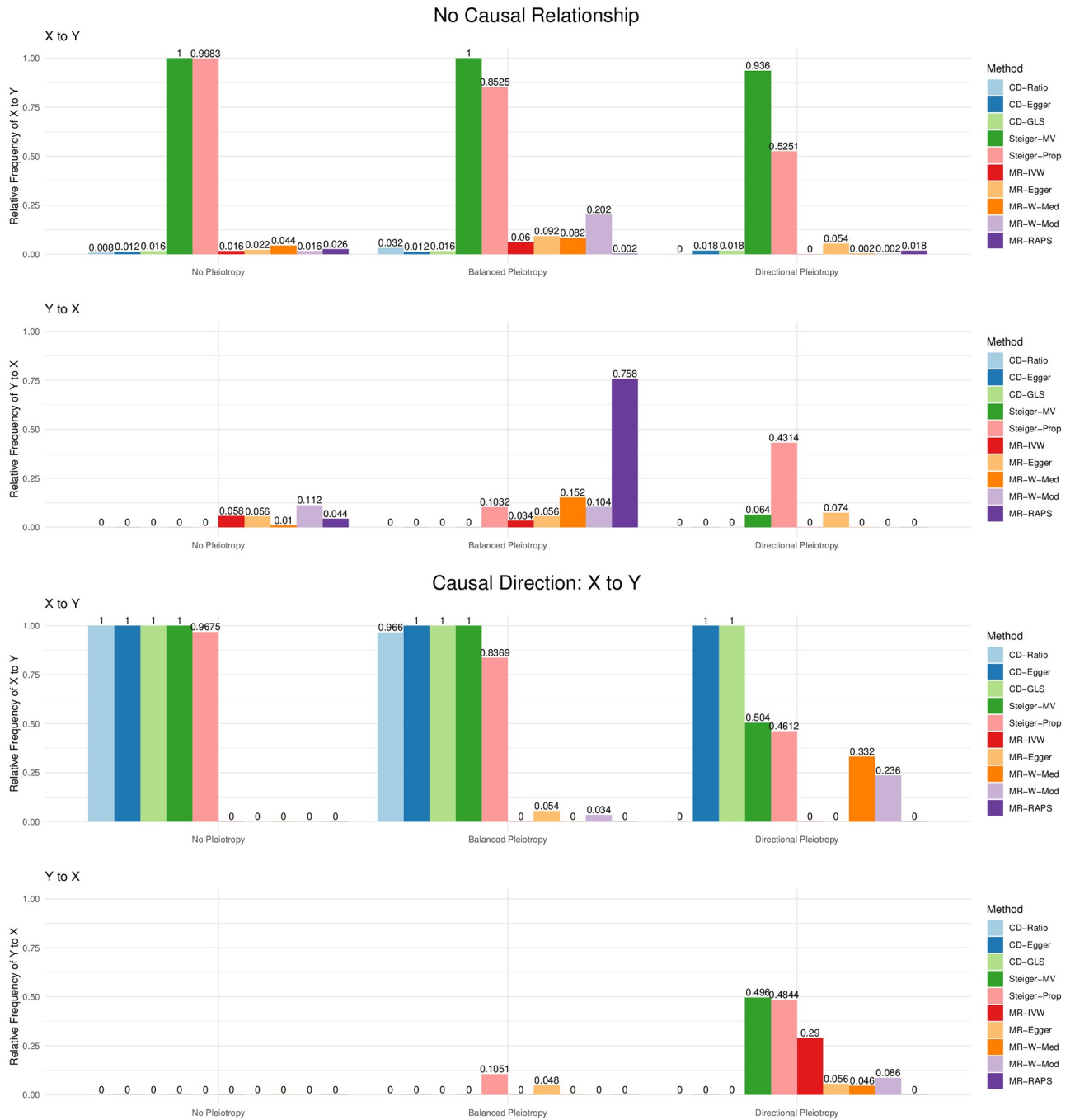
<https://doi.org/10.1371/journal.pgen.1009105.t005>

**The GOF tests performed well.** In *S6 Text*, we show the results of our proposed GOF tests: all performed well with controlled type I error rates and high power in our simulations.

### Secondary simulation results

**Bi-CD-Ratio could detect bi-directional causal relationships.** In *S7 Text*, we also show the results for inferring no, uni- and bi-directional relationships between two traits with valid IVs. It is confirmed that our proposed bi-CD-Ratio and (bi-directional) MR-Wald-Ratio performed well, while Steiger’s method did not work due to its inability to determine whether there is any causal relationship between two traits as discussed earlier.

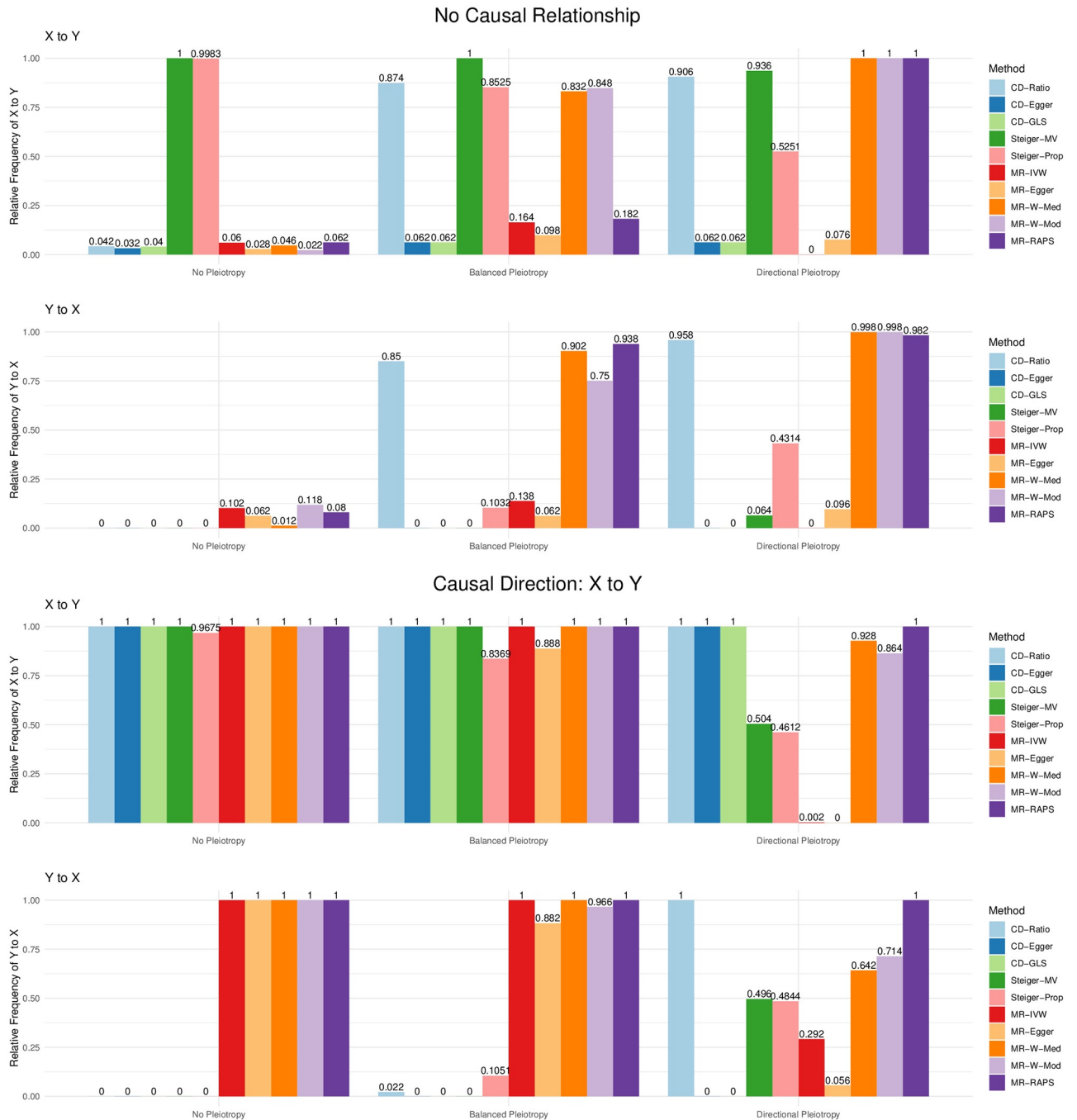
**CD-Ratio could perform in TWAS with relatively small sample sizes.** We also did a simulation study to investigate how the asymptotic theory for the sample correlations and their ratios performs with smaller sample sizes as typical with molecular endophenotypes as in TWAS. We generated simulated data under realistic set-ups mimicking real eQTL datasets with sample sizes from around two hundreds to about a thousand, much smaller than those of typical GWAS as used in our main simulations. We also considered the presence of some highly correlated SNPs/IVs. As shown in *S8 Text*, our proposed CD-ratio performed very well: it might be a bit conservative with empirical type I error rates smaller than the nominal level



**Fig 9. Relative frequencies of the causal directions reached by each method (from 500 simulations in each set-up) when there is no causal relationship (top two panels) or a causal direction is from X to Y ( $\beta_{YX} = 0.2$ ; bottom two) with no pleiotropy, balanced pleiotropy ( $\mu_\alpha = 0, \sigma_\mu = 0.1$ ) and directional pleiotropy ( $\mu_\alpha = 0.4, \sigma_\mu = 0.1$ ) respectively.**

<https://doi.org/10.1371/journal.pgen.1009105.g009>

when the sample size was small at a few hundreds; but as the sample size increased to a few thousands, its type I error rates approached the nominal level. Importantly, it was confirmed that by combining the results across the SNPs, even including some highly correlated ones, CD-Ratio improved the statistical power over that based on single SNPs.



**Fig 10. Relative frequencies of having significant results in each direction, in contrast to that of the final conclusions shown in Fig 9.**

<https://doi.org/10.1371/journal.pgen.1009105.g010>

**Selection bias.** We did simulations to study possible selection bias due to the double use of the same sample for SNP/IV selection and estimation. The simulation setups are the same as described in Main Simulation Set-ups section. We used 22 SNPs to generate simulated data, and in Main Simulation Results section we applied our proposed three methods to all 22 SNPs with the results shown in Table 5. Now for each simulated dataset, we used a p-value cutoff to

select exactly  $N$  SNPs that were most significantly associated with both  $X$  and  $Y$ , then we applied a method with these  $N$  selected SNPs. We tried with  $N = 10, 15$  or  $22$ ; when  $N = 22$ , we used all 22 SNPs as before. [S9 Text](#) shows the simulation results with no pleiotropy, balanced pleiotropy, and directional pleiotropy respectively. We summarize the results below. First, when there was no pleiotropy, all three methods gave almost unbiased estimates of  $K_{YX}$  with any of the three values of  $N$ , suggesting no or negligible selection bias. Second, with balanced pleiotropy, there were small biases. Third, with directional pleiotropy, there could be large biases. However, interestingly, under the null hypothesis of  $K_{YX} = 0$ , CD-Egger and CD-GLS yielded only small biases, leading to barely any inflation of their Type I error rates; see [S9 Text](#).

## Discussion

We have proposed several extensions to the increasingly popular MR Steiger's method to infer the causal direction between two traits. MR Steiger's method is based on using a single SNP as a valid IV; if there is strong biological knowledge to justify the use of the SNP as a valid IV, it is both useful and simple to apply the method. However, in practice, due to incomplete knowledge, one often just uses genome-wide significant SNPs as IVs, in which case there are several potential problems. Some of the SNPs are likely invalid IVs due to wide-spread horizontal pleiotropy. In addition, the results may be low-powered and inconsistent, each critically depending on the SNP being used. Our proposed multiple SNP-based approaches would alleviate the above two problems. First, by using multiple SNPs, we allow and thus explicitly model possible pleiotropic effects, leading to two new methods, CD-Egger and CD-GLS, that are robust to and more powerful in the presence of pleiotropy. Second, combining the results over multiple SNPs is both more robust and more powerful than a single SNP-based method. In applications to inferring/confirming causal directions between molecular endophenotypes and complex diseases or quantitative traits, such as in TWAS, one would typically use multiple SNPs in LD from a local region to predict an endophenotype (e.g. a gene's expression level) with relatively small sample sizes; in these applications, it may be necessary to apply a multiple correlated SNPs-based method like ours to achieve as high power as possible. Third, we can apply some goodness-of-fit tests to check the modeling assumptions. Furthermore, our methods can simultaneously infer both the existence and (if so) the direction of a causal relationship between two traits. For this reason, even for single SNPs, our CD-Ratio is a useful alternative to Steiger's method because of the former's ability to avoid incorrect conclusions of a causal direction when there is no causal relationship (and no pleiotropy). In addition, our methods have some advantages, i.e. higher statistical power and lower false positives, over, and are thus useful alternatives to, multiple SNP-based bi-directional MR. These advantages over bi-directional MR are perhaps due to that our proposed methods take advantage of and thus exploit some specific relationships between the Pearson correlations between each SNP/IV and the two traits implicitly implied by the true causal models. These points were confirmed in both extensive simulations and an application to infer the causal direction between LDL cholesterol and CAD. In particular, bi-directional MR was found to be non-robust for inferring causal directions in the presence of invalid IVs even when equipped with a robust MR method. In summary, our proposed methods are expected to be useful in the toolbox for causal inference, for which application of multiple complementary methods, called triangulation [55], is strongly advocated.

Our methods have some limitations. First and foremost, our methods are based on the standard asymptotic theory for sample correlations under the assumption of a large sample size (say  $n$ ) and a small number of SNPs/IVs (say  $m$ ). It has some important implications. Although including more relevant SNPs/IVs is expected to improve the statistical estimation efficiency



and power in the ideal case, in practice there are some arguments against doing so: 1) including too many SNPs, i.e. making  $m$  large relative to  $n$ , the standard asymptotic theory may break down, leading to incorrect inference; 2) in particular, including some highly correlated SNPs may gain little while being at the risk of inducing a covariance matrix nearly singular, leading to bad estimates and inference with numerical problems; 3) our methods require estimating a large covariance matrix of dimension in the order of  $m^2 \times m^2$  for a sample correlation matrix in the order of  $m \times m$ , demanding too much computer memory and computing time even if  $m$  is only in hundreds or thousands; 4) including more SNPs will increase the likelihood of violating IV assumptions; 5) it might be harder to accurately impute with high LD between any two SNPs for GWAS summary data based on a not-too-large reference panel. For the above reasons, in our applications we pruned out those significant SNPs with pairwise correlations larger than 0.8. We also emphasize that we require the selected SNPs/IVs to be statistically significant for both traits. One reason is similar to avoiding biases due to weak IVs in MR. Our methods are based on the ratios of sample correlations, say  $r_{Yg}/r_{Xg}$ . If the mean of the denominator  $r_{Xg}$  is 0, regardless of whether the normality (approximately) holds for each sample correlation, the ratio cannot be well approximated by a normal approximation; our methods are based on the asymptotic joint normal distribution for the correlation ratios. Second, we have mainly restricted to inferring uni-directional causal relationships. Although we have outlined an extension of our CD-Ratio method (along with bi-directional MR) to the case with bi-directional causal relationships, and shown some promising simulation results under the ideal situation with valid IVs, it remains to see how to extend and apply the methods in practice without knowing which IVs are valid, in which case there may be issues of non-identifiable models [45]. We note by passing that [45] considered a mediating analysis of how to estimate direct/indirect effects given the presence of a bi-directional causal relationship, differing from our task of estimating what causal direction if any. Third, as often done in applications with MR, we have used the same two GWAS samples both for selecting SNPs as IVs and for subsequent estimation and inference. As expected, this double use of the data could lead to selection bias. It was confirmed in our simulations in the presence of directional pleiotropy, though no or only small biases were observed in cases of no or only balanced pleiotropy, possibly due to our recommendation of selecting SNPs/IVs that are significant for both traits, alleviating the problem as compared to selecting SNPs/IVs based on only one trait as commonly adopted in MR. Nevertheless, this issue needs to be further studied; if possible, we may use a third independent GWAS sample for such selection to avoid selection bias as in the three-sample MR design [56]. Fourth, it is noted that our proposed methods are developed in the framework of two-sample MR, which is more common and less restrictive than the one-sample set-up (where both traits and SNPs are obtained from the same sample). Nevertheless, it will be useful to extend our methods to the one-sample set-up. In particular, such methods may be useful for mediation analysis [57, 58], in which methods to determine causal directions (i.e. to determine between two traits which is the mediator and which is the outcome) lack and are urgently needed while a bi-directional mediation analysis would be problematic as recently reported [59]. In addition, we have not compared with structure equation modeling [60–62], Bayesian network [63] and latent causal variable [64] approaches, mainly because of their lack of statistical significance testing or handling hidden confounding. Finally, we have pointed out a strong connection between our proposed CD-Egger method and MR Egger regression in dealing with horizontal pleiotropy. Taking advantage of this connection, one may be able to develop other extensions as developed for MR [22, 29, 53, 65, 66] in addition to those used here (e.g. weighted median and mode methods in MR) for inference of causal direction. These interesting topics warrant future investigation.

## Supporting information

**S1 Text. Asymptotic covariance matrix of  $r_{Xg}$  and  $r_{Yg}$ .**  
(PDF)

**S2 Text. Details of CD-Egger and CD-GLS.**  
(PDF)

**S3 Text. Degrees of freedom for the goodness-of-fit tests.**  
(PDF)

**S4 Text. More results for LDL and CAD.**  
(PDF)

**S5 Text. Results for LDL/HDL and CAD with multivariable approaches.**  
(PDF)

**S6 Text. More simulation results.**  
(PDF)

**S7 Text. Simulations for bi-directional causal effects.**  
(PDF)

**S8 Text. Simulations for TWAS with small sample sizes.**  
(PDF)

**S9 Text. Simulations to study selection bias.**  
(PDF)

**S10 Text. Fisher's transformation of the absolute value of a correlation.**  
(PDF)

## Author Contributions

**Conceptualization:** Haoran Xue, Wei Pan.

**Data curation:** Haoran Xue.

**Formal analysis:** Haoran Xue.

**Funding acquisition:** Wei Pan.

**Investigation:** Haoran Xue, Wei Pan.

**Methodology:** Haoran Xue, Wei Pan.

**Project administration:** Wei Pan.

**Resources:** Wei Pan.

**Software:** Haoran Xue.

**Supervision:** Wei Pan.

**Validation:** Haoran Xue.

**Visualization:** Haoran Xue.

**Writing – original draft:** Haoran Xue.

**Writing – review & editing:** Haoran Xue, Wei Pan.

## References

1. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003 Feb; 32(1):1–22. <https://doi.org/10.1093/ije/dyg070> PMID: 12689998
2. Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol*. 2004 Feb; 33(1):30–42. <https://doi.org/10.1093/ije/dyh132> PMID: 15075143
3. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014 Sep 15; 23(R1):R89–98. <https://doi.org/10.1093/hmg/ddu328> PMID: 25064373
4. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018 May 30; 7:e34408. <https://doi.org/10.7554/eLife.34408> PMID: 29846171
5. Holmes MV, Ala-Korpela M, Davey Smith G. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol*. 2017 Oct; 14(10):577–590. <https://doi.org/10.1038/nrcardio.2017.78> PMID: 28569269
6. Timpson NJ, Nordestgaard BG, Harbord RM, Zacho J, Frayling TM, Tybjaerg-Hansen A, et al. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int J Obes (Lond)*. 2011 Feb; 35(2):300–8. <https://doi.org/10.1038/ijo.2010.137> PMID: 20714329
7. Richmond RC, Davey Smith G, Ness AR, den Hoed M, McMahon G, Timpson NJ. Assessing causality in the association between child adiposity and physical activity levels: a Mendelian randomization analysis. *PLoS Med*. 2014 Mar 18; 11(3):e1001618. <https://doi.org/10.1371/journal.pmed.1001618> PMID: 24642734
8. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet*. 2016 Jul; 48(7):709–17. <https://doi.org/10.1038/ng.3570> PMID: 27182965
9. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017 Nov 17; 13(11):e1007081. <https://doi.org/10.1371/journal.pgen.1007081> PMID: 29149188
10. Millstein J, Zhang B, Zhu J, Schadt EE. Disentangling molecular relationships with a causal inference test. *BMC Genet*. 2009 May 27; 10:23. <https://doi.org/10.1186/1471-2156-10-23> PMID: 19473544
11. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet*. 2015 Sep; 47(9):1091–8. <https://doi.org/10.1038/ng.3367> PMID: 26258848
12. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016 Mar; 48(3):245–52. <https://doi.org/10.1038/ng.3506> PMID: 26854917
13. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016 May; 48(5):481–7. <https://doi.org/10.1038/ng.3538> PMID: 27019110
14. Xu Z, Wu C, Wei P, Pan W. A Powerful Framework for Integrating eQTL and GWAS Summary Data. *Genetics*. 2017 Nov; 207(3):893–902. <https://doi.org/10.1534/genetics.117.300270> PMID: 28893853
15. Xu Z, Wu C, Pan W; Alzheimer's Disease Neuroimaging Initiative. Imaging-wide association study: Integrating imaging endophenotypes in GWAS. *Neuroimage*. 2017 Oct 1; 159:159–169. <https://doi.org/10.1016/j.neuroimage.2017.07.036> PMID: 28736311
16. Han S, Lin Y, Wang M, Goes FS, Tan K, Zandi P, et al. Integrating brain methylome with GWAS for psychiatric risk gene discovery. *bioRxiv*. 2018 Jan 1:440206. <https://doi.org/10.1101/440206>
17. Su YR, Di C, Bien S, Huang L, Dong X, Abecasis G, et al. A Mixed-Effects Model for Powerful Association Tests in Integrative Functional Genomics. *Am J Hum Genet*. 2018 May 3; 102(5):904–919. <https://doi.org/10.1016/j.ajhg.2018.03.019> PMID: 29727690
18. Cai M, Chen L, Liu J, Yang C. Quantifying the impact of genetically regulated expression on complex traits and diseases. *bioRxiv*. 2019 Jan 1:546580. <https://doi.org/10.1101/546580>
19. Hu Y, Li M, Lu Q, Weng H, Wang J, Zekavat SM, et al. A statistical framework for cross-tissue transcriptome-wide association analysis. *Nat Genet*. 2019 Mar; 51(3):568–576. <https://doi.org/10.1038/s41588-019-0345-7> PMID: 30804563
20. Yang T, Wu C, Wei P, Pan W. Integrating DNA sequencing and transcriptomic data for association analyses of low-frequency variants and lipid traits. *Hum Mol Genet*. 2020 Feb 1; 29(3):515–526. <https://doi.org/10.1093/hmg/ddz314> PMID: 31919517

21. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun.* 2018 Jan 15; 9(1):224. <https://doi.org/10.1038/s41467-017-02317-2> PMID: 29335400
22. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* 2018 May; 50(5):693–698. <https://doi.org/10.1038/s41588-018-0099-7> PMID: 29686387
23. Watanabe K, Stringer S, Frei O, Umičević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet.* 2019 Sep; 51(9):1339–1348. <https://doi.org/10.1038/s41588-019-0481-0> PMID: 31427789
24. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet.* 2019 Apr; 51(4):592–599. <https://doi.org/10.1038/s41588-019-0385-z> PMID: 30926968
25. Mancuso N, Freund MK, Johnson R, Shi H, Kichaev G, Gusev A, et al. Probabilistic fine-mapping of transcriptome-wide association studies. *Nat Genet.* 2019 Apr; 51(4):675–682. <https://doi.org/10.1038/s41588-019-0367-1> PMID: 30926970
26. Wu C, Pan W. A powerful fine-mapping method for transcriptome-wide association studies. *Hum Genet.* 2020 Feb; 139(2):199–213. <https://doi.org/10.1007/s00439-019-02098-2> PMID: 31844974
27. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015 Apr; 44(2):512–25. <https://doi.org/10.1093/ije/dyv080> PMID: 26050253
28. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol.* 2015 Feb 15; 181(4):251–60. <https://doi.org/10.1093/aje/kwu283> PMID: 25632051
29. Burgess S, Dudbridge F, Thompson SG. Re: “Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects”. *Am J Epidemiol.* 2015 Feb 15; 181(4):290–1. <https://doi.org/10.1093/aje/kwv017> PMID: 25660081
30. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *J Am Coll Cardiol.* 2012 Dec 25; 60(25):2631–9. <https://doi.org/10.1016/j.jacc.2012.09.017> PMID: 23083789
31. Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J.* 2015 Mar 1; 36(9):539–50. <https://doi.org/10.1093/eurheartj/ehv571> PMID: 24474739
32. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet.* 2012 Aug 11; 380(9841):572–80. [https://doi.org/10.1016/S0140-6736\(12\)60312-2](https://doi.org/10.1016/S0140-6736(12)60312-2) PMID: 22607825
33. White J, Swerdlow DI, Preiss D, Fairhurst-Hunter Z, Keating BJ, Asselbergs FW, et al. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol.* 2016 Sep 1; 1(6):692–9. <https://doi.org/10.1001/jamacardio.2016.1884> PMID: 27487401
34. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet.* 2010 Nov 13; 376(9753):1670–81. [https://doi.org/10.1016/S0140-6736\(10\)61350-5](https://doi.org/10.1016/S0140-6736(10)61350-5) PMID: 21067804
35. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. *Lancet.* 2016 Nov 19; 388(10059):2532–2561. [https://doi.org/10.1016/S0140-6736\(16\)31357-5](https://doi.org/10.1016/S0140-6736(16)31357-5) PMID: 27616593
36. Silverman MG, Ference BA, Im K, Wiviott SD, Giugliano RP, Grundy SM, et al. Association Between Lowering LDL-C and Cardiovascular Risk Reduction Among Different Therapeutic Interventions: A Systematic Review and Meta-analysis. *JAMA.* 2016 Sep 27; 316(12):1289–97. <https://doi.org/10.1001/jama.2016.13985> PMID: 27673306
37. Emerging Risk Factors Collaboration, Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, et al. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA.* 2009 Nov 11; 302(18):1993–2000. <https://doi.org/10.1001/jama.2009.1619> PMID: 19903920
38. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJ, Komajda M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med.* 2007 Nov 22; 357(21):2109–22. <https://doi.org/10.1056/NEJMoa0706628> PMID: 17984165
39. Schwartz GG, Olsson AG, Abt M, Ballantyne CM, Barter PJ, Brumm J, et al. Effects of dalcetrapib in patients with a recent acute coronary syndrome. *N Engl J Med.* 2012 Nov 29; 367(22):2089–99. <https://doi.org/10.1056/NEJMoa1206797> PMID: 23126252

40. Neudecker H, Wesselman AM. The asymptotic variance matrix of the sample correlation matrix. *Linear Algebra and its Applications*. 1990 Jan 1; 127:589–99. [https://doi.org/10.1016/0024-3795\(90\)90363-H](https://doi.org/10.1016/0024-3795(90)90363-H)
41. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015 Oct 1; 526(7571):68–74. <https://doi.org/10.1038/nature15393> PMID: 26432245
42. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Annals of Statistics*. 2020; 48(3):1742–69. <https://doi.org/10.1214/19-AOS1866>
43. Dai JY, Peters U, Wang X, Kocarnik J, Chang-Claude J, Slattery ML, et al. Diagnostics for Pleiotropy in Mendelian Randomization Studies: Global and Individual Tests for Direct Effects. *Am J Epidemiol*. 2018 Dec 1; 187(12):2672–2680. <https://doi.org/10.1093/aje/kwy177> PMID: 30188971
44. Bowden J, Hemani G, Davey Smith G. Invited Commentary: Detecting Individual and Global Horizontal Pleiotropy in Mendelian Randomization—A Job for the Humble Heterogeneity Statistic? *Am J Epidemiol*. 2018 Dec 1; 187(12):2681–2685. PMID: 30188969
45. Talluri R, Shete S. An approach to estimate bidirectional mediation effects with application to body mass index and fasting glucose. *Ann Hum Genet*. 2018 Nov; 82(6):396–406. <https://doi.org/10.1111/ahg.12261> PMID: 29993118
46. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013 Nov; 37(7):658–65. <https://doi.org/10.1002/gepi.21758> PMID: 24114802
47. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016 May; 40(4):304–14. <https://doi.org/10.1002/gepi.21965> PMID: 27061298
48. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017 Dec 1; 46(6):1985–1998. <https://doi.org/10.1093/ije/dyx102> PMID: 29040600
49. Burgess S, Zuber V, Gkatzionis A, Foley CN. Modal-based estimation via heterogeneity-penalized weighting: model averaging for consistent and efficient estimation in Mendelian randomization when a plurality of candidate instruments are valid. *Int J Epidemiol*. 2018 Aug 1; 47(4):1242–1254. <https://doi.org/10.1093/ije/dyy080>
50. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet*. 2017 Dec; 49(12):1758–1766. <https://doi.org/10.1038/ng.3977> PMID: 29083408
51. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet*. 2017 Sep; 49(9):1385–1391. <https://doi.org/10.1038/ng.3913> PMID: 28714975
52. Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics*. 2016 Jan 15; 32(2):283–5.
53. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol*. 2020 Jun; 44(4):313–329. <https://doi.org/10.1002/gepi.22295>
54. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, et al. Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep*. 2017; 4(4):330–345. <https://doi.org/10.1007/s40471-017-0128-6> PMID: 29226067
55. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016 Dec 1; 45(6):1866–1886.
56. Zhao Q, Chen Y, Wang J, Small DS. Powerful three-sample genome-wide design and robust statistical inference in summary-data Mendelian randomization. *Int J Epidemiol*. 2019 Oct 1; 48(5):1478–1492. <https://doi.org/10.1093/ije/dyz142>
57. Wang K. Understanding Power Anomalies in Mediation Analysis. *Psychometrika*. 2018 Jun; 83(2):387–406. <https://doi.org/10.1007/s11336-017-9598-1>
58. Schaid DJ, Sinnwell JP. Penalized models for analysis of multiple mediators. *Genet Epidemiol*. 2020 Jul; 44(5):408–424. <https://doi.org/10.1002/gepi.22296>
59. Lutz SM, Sordillo JE, Hokanson JE, Chen Wu A, Lange C. The effects of misspecification of the mediator and outcome in mediation analysis. *Genet Epidemiol*. 2020 Jun; 44(4):400–403. <https://doi.org/10.1002/gepi.22289>
60. Ainsworth HF, Shin SY, Cordell HJ. A comparison of methods for inferring causal relationships between genotype and phenotype using additional biological measurements. *Genet Epidemiol*. 2017 Nov; 41(7):577–586. <https://doi.org/10.1002/gepi.22061>

61. Yuan Y, Shen X, Pan W, Wang Z. Constrained likelihood for reconstructing a directed acyclic Gaussian graph. *Biometrika* 2019 March; 106(1):109–125. <https://doi.org/10.1093/biomet/asy057>
62. Li C, Shen X, Pan W. Likelihood ratio tests for a large directed acyclic graph. *Journal of the American Statistical Association* 2020; 115(531):1304–1319. <https://doi.org/10.1080/01621459.2019.1623042>
63. Howey R, Shin SY, Relton C, Davey Smith G, Cordell HJ. Bayesian network analysis incorporating genetic anchors complements conventional Mendelian randomization approaches for exploratory analysis of causal relationships in complex data. *PLoS Genet.* 2020 Mar 2; 16(3):e1008198. <https://doi.org/10.1371/journal.pgen.1008198> PMID: 32119656
64. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet.* 2018 Dec; 50(12):1728–1734. <https://doi.org/10.1038/s41588-018-0255-0> PMID: 30374074
65. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med.* 2017 Dec 20; 36(29):4705–4718. <https://doi.org/10.1002/sim.7492> PMID: 28960498
66. Qi G, Chatterjee N. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nat Commun.* 2019 Apr 26; 10(1):1941. <https://doi.org/10.1038/s41467-019-09432-2> PMID: 31028273